RECENT ADVANCES on APPLIED MATHEMATICS and COMPUTATIONAL METHODS in ENGINEERING

Proceedings of the International Conference on Applied Mathematics and Computational Methods in Engineering (AMCME 2015)

Barcelona, Spain
April 7-9, 2015

RECENT ADVANCES on APPLIED MATHEMATICS and COMPUTATIONAL METHODS in ENGINEERING

Proceedings of the International Conference on Applied Mathematics and Computational Methods in Engineering (AMCME 2015)

Barcelona, Spain
April 7-9, 2015

All papers of the present volume were peer reviewed by no less than two independent reviewers. Acceptance was granted when both reviewers' recommendations were positive.

Series: Mathematics and Computers in Science and Engineering Series | 43

ISSN: 2227-4588
ISBN: 978-1-61804-292-7

RECENT ADVANCES on APPLIED MATHEMATICS and COMPUTATIONAL METHODS in ENGINEERING

Proceedings of the International Conference on Applied Mathematics and Computational Methods in Engineering (AMCME 2015)

Barcelona, Spain
April 7-9, 2015

Organizing Committee

Additional Reviewers

Francesco Zirilli	Sapienza Universita di Roma, Italy
Sorinel Oprisan	College of Charleston, CA, USA
Xiang Bai	Huazhong University of Science and Technology, China
Philippe Dondon	Institut polytechnique de Bordeaux, France
Yamagishi Hiromitsu	Ehime University, Japan
Frederic Kuznik	National Institute of Applied Sciences, Lyon, France
George Barreto	Pontificia Universidad Javeriana, Colombia
Takuya Yamano	Kanagawa University, Japan
Imre Rudas	Obuda University, Budapest, Hungary
Tetsuya Shimamura	Saitama University, Japan
M. Javed Khan	Tuskegee University, AL, USA
Eleazar Jimenez Serrano	Kyushu University, Japan
Valeri Mladenov	Technical University of Sofia, Bulgaria
Jon Burley	Michigan State University, MI, USA
Andrey Dmitriev	Russian Academy of Sciences, Russia
Moran Wang	Tsinghua University, China
Jose Flores	The University of South Dakota, SD, USA
Hessam Ghasemnejad	Kingston University London, UK
Santoso Wibowo	CQ University, Australia
Kazuhiko Natori	Toho University, Japan
Konstantin Volkov	Kingston University London, UK
Kei Eguchi	Fukuoka Institute of Technology, Japan
Abelha Antonio	Universidade do Minho, Portugal
Tetsuya Yoshida	Hokkaido University, Japan
Matthias Buyle	Artesis Hogeschool Antwerpen, Belgium
Deolinda Rasteiro	Coimbra Institute of Engineering, Portugal
Masaji Tanaka	Okayama University of Science, Japan
Bazil Taha Ahmed	Universidad Autonoma de Madrid, Spain
Zhong-Jie Han	Tianjin University, China
James Vance	The University of Virginia's College at Wise, VA, USA
Angel F. Tenorio	Universidad Pablo de Olavide, Spain
Genqi Xu	Tianjin University, China
João Bastos	Instituto Superior de Engenharia do Porto, Portugal
Miguel Carriegos	Universidad de Leon, Spain
Shinji Osada	Gifu University School of Medicine, Japan
Ole Christian Boe	Norwegian Military Academy, Norway
Lesley Farmer	California State University Long Beach, CA, USA
Dmitrijs Serdjuks	Riga Technical University, Latvia
Alejandro Fuentes-Penna	Universidad Autónoma del Estado de Hidalgo, Mexico
Francesco Rotondo	Polytechnic of Bari University, Italy
Stavros Ponis	National Technical University of Athens, Greece
José Carlos Metrôlho	Instituto Politecnico de Castelo Branco, Portugal
Minhui Yan	Shanghai Maritime University, China

Table of Contents

ISBN: 978-1-61804-292-7

ISBN: 978-1-61804-292-7

Fall Avoidance Using Temporal Bayesian Networks and Wireless Sensors in Soft Computing

Darren Seifert
Department of Computer Science
University of North Dakota
Grand Forks, ND 58202-9015, U.S.A.
Darren.Seifert@my.und.edu

Eunjin Kim*
Department of Computer Science
University of North Dakota
Grand Forks, ND 58202-9015, U.S.A.
ejkim@cs.und.edu

***Abstract*— The assisted living center uses a distributed system of sensors to monitor potential slips and falls of the residents. In our study, a Temporal Dynamic Bayesian Network is employed in monitoring of the sensors. Through simulation, we show that it's effective to use TDBN and may help improve the time management of staff.**

Keywords—Dynamic Bayesian Network, wireless sensor networks; filtering, smoothing, healthcare;

I. INTRODUCTION

Wireless sensor networks have been changing the way of our living significantly in our lives. Researchers are investigating everything from microscopic sensors that traverse the bloodstream and wirelessly report health conditions to intelligent household devices that can interact with each other wirelessly [2]. Some of this research has already made it to our lives in the form of modern home security systems. These systems have become somewhat advanced in that many have the ability to be controlled over the internet or a mobile device, performing tasks such as control of lights and appliances, locking doors and monitoring of various sensors, etc. They are becoming more common in our lives.

One of the major challenges we will face over the next 20 years is the aging of our population. The US Census Bureau estimates the percentage of population over the age of 65 will grow from 12.9% in 2010 to 16.1% in 2020 and on to 19.3% by 2030 [4]. Over time, this ever growing population of seniors will likely put a strain on the staff and resources of assisted living centers. The Alzheimer's Society of Canada estimates that the number of hours of informal care required by people with dementia will have tripled from 231 million hours to 756 million hours by the year 2038. [1] Wireless sensor networks can be seen, in part, as a solution to this problem.

A sensor network deployed in an assisted living facility can take advantage of many of the same technologies that are used at the in-home system. However, one key difference is the user of the network is typically also the person being monitored in the in-home system. In an assisted living facility, a network needs to be designed for staff to monitor the wellbeing of the residents. As such, more sophisticated monitoring software is needed. Since staffs of assisted living centers have many responsibilities, the system will become quickly unused if a sensor system provides them a false alert or fail to alert a problem.

These sorts of technologies are not entirely new to assisted living facilities. Doors are often alarmed and equipped with keypad entry systems. If a resident failed to stand repeatedly without assistance, they typically are supplied with a pressure sensor on his/her bed or/and chair(s). If (s)he moved off of this sensor, an alarm is triggered to alert a staff. While these types of sensors are seen to be effective and reliable in many situations, they do have some drawbacks.

Existing pressure sensors consist of a simple switch inside of a large pad. If a person moves off of the pad, it makes a very loud and high pitched noise until pressure returns or it is switched off. If a resident is unaware of a cause of this noise, they can often become agitated. This is often the case with patients with dementia. In addition, these pads frequently come dislodged accidentally. Residents may find sitting or lying on them uncomfortable, or may shift in their chair or roll over and dislodge them. While this situation usually is not an emergency, staff members have to always respond to it.

One approach to address this problem is through the use of a wireless sensor network. A simple solution that would be of great help is to connect the pressure sensor to a network and alert staff through the aid of a computer that the sensor had become dislodged. This would reduce the agitation that comes from the audible noise in the current pad, but would do little to help the overburdened staff.

With the goal of eliminating resident agitation while simultaneously easing the burden on staff, we propose to address this problem through a combination of wireless sensors and a Dynamic Bayesian Network (DBN). We will first use a networked *pressure sensor* in combination with a *motion sensor* to predict a person's attempt to stand. These predictions will be combined over time by the DBN to better estimate the probability of a person's attempt to get up. This will help to eliminate false alerts from a single sensor reading as well as provide a more informed picture of exact situation. This model will then be updated to include readings from a Radio-frequency identification (RFID) sensor. This will allow the system to react to staffs and adjust the alert level appropriately.

ISBN: 978-1-61804-292-7

This study achieves the following objectives:

1. Clearly distinguishable levels of convergence will be identified for an appropriate alert to staff regarding the problems of residents.

2. The inclusion of the DBN will reduce *false alert cases.*

3. Response delays incurred as a result of the inclusion of the DBN will be held to within appropriate levels.

II. Survey of the Related Works

Activity recognition (AR) has been a popular research area and can take on many forms. The types of recognition can vary from recognizing simple touch gestures as the modern smart devices do, to more complex systems that allow individuals with cognitive impairments to transition through vocational tasks using in-air gestures [10]. Some of these systems may depend on monitoring video and inferred sensor readings of a situation while others may rely on wearable sensors [3].

Using sensors and artificial intelligence techniques to assist seniors and people with disabilities with daily tasks has been a fast growing field in recent years. One such system, called PROACT, uses body-worn Radio-frequency identification (RFID) sensors and a probability engine that infers activities given sensor observations to create probabilistic models of a person's activities [5]. The COACH (Cognitive Orthosis for Assistive aCtivities in the Home) system uses computer vision and a Partially Observable Markov Decision Process (POMDP) to learn characteristics about an individual over time. These learned traits are then used to build custom plans for an individual to accomplish tasks in their everyday lives [6].

Much of the research on the use of sensors in fall prevention has centered on detection of the moment of fall. Such research relies on the use of accelerometers or gyroscopes attached to the resident [7]. Other techniques may rely on either video or still images of a person to determine a person's physical positioning. Hidden Markov Models are often used in conjunction with these sensors to determine when the likelihood of a sequence of events has fallen below some threshold at which point an alert is signaled [8][9].

While these systems are of great value, the potential residents who fall and possibly cause injury are typically confined to a wheelchair when staffs are unavailable in the assisted living center. As a result, very specific models that can detect a moment of a person's attempt to stand up are needed.

III. Temporal Dynamic Bayesian Models and Their Simulation

Our initial model relies on a pressure sensor and a motion sensor that operate independently of each other. Both sensors can make a probabilistic assertion about whether a person is attempting to stand based on their readings. These independent probabilities are then combined with previous probabilistic evidence of the person's standing at time *t-1* to make a decision for the current time slice *t*.

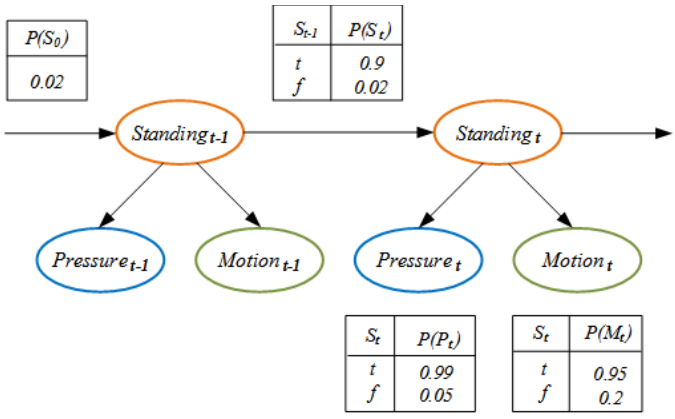


Fig. 1. Initial Dynamic Bayesian Network Model

A. Initial Model

For our model, we started with the assumption that given the average senior with dementia was standing previously, there was about a 90% chance that they would continue standing. If they were previously not standing, there would be a 2% chance that they would attempt to stand. Our pressure sensor would alert 99% of the time given that the person occupying it had stood up. It also would alert about 5% of the time if the occupant had not stood up. This error would most likely be due to a person shifting in their seat and dislodging the sensor. Our motion sensor would alert to a person standing up 95% of the time. However, 20% of the time it would see various kinds of motion even if the person was still seated. This initial model is depicted with the given probabilities in Figure 1.

For each time slice, our initial algorithm first calculates the prediction probability that a person is standing at time *t* based on the standing probability at time *t-1* given the previous sensor evidence. In this example, we seek for the prediction probability of standing (*S*) at *t* with an active pressure sensor (*p*) and an inactive motion sensor (¬*m*) at time *t-1*:

$$P(S_t|p_{t-1}, \neg m_{t-1}) \tag{1}$$

$$= \sum_{s_{t-1}} P(S_t|s_{t-1})P(s_{t-1}|p_{t-1}, \neg m_{t-1})$$

$$= P(S_t|s_{t-1})P(s_{t-1}|p_{t-1}, \neg m_{t-1})$$

$$+P(S_t|\neg s_{t-1})P(\neg s_{t-1}|p_{t-1}, \neg m_{t-1})$$

This probability is then used in conjunction with the current sensor readings to determine the current probability of a person standing at time *t*:

$$P(S_t|p_{t-1:t}, \neg m_{t-1:t}) \tag{2}$$
$$= \alpha P(p_t, \neg m_t|S_t)P(S_t|p_{t-1}, \neg m_{t-1})$$

The process was repeated over several time slices to predict what would happen with the proposed model under various conditions. In addition, we chose to examine how the predictions were affected when future sensor evidence was known through smoothing. The smoothing process begins by

trying to determine what the probability of the future sensor reading occurring is, given the current standing probability:

$$P(p_t, \neg m_t | S_{t-1}) = \sum_{S_t} P(p_t, \neg m_t | S_t) P(S_t) P(S_t | S_{t-1}) \quad (3)$$
$$= P(p_t, \neg m_t | s_t) P(s_t) P(s_t | S_{t-1})$$
$$+ P(p_t, \neg m_t | \neg s_t) P(\neg s_t) P(\neg s_t | S_{t-1})$$

This value is then combined with the current standing probability, given the current sensor readings:

$$P(S_{t-1} | p_{t-1:t}, \neg m_{t-1:t}) = \quad (4)$$
$$\alpha P(S_{t-1} | p_{t-1}, \neg m_{t-1}) P(p_t, \neg m_t | S_{t-1})$$

For our simulations we performs smoothing backwards for 10 time slices, *t-10*. Assuming a sensor reading occurred once a second, we felt that this was the maximum possible delay that could be expected if the smoothed value were to be used to alert staff to a problem. To conduct these simulations, a C++ program was written to model the Dynamic Bayesian Network and the filtering and smoothing processes. These processes were then fed various sequences of sensor readings that were thought to simulate real world readings.

Let us first investigate what would happen if the model was fed constant sensor readings for a number of time slices to try and decide a reasonable point at which staff should be alerted using the wireless sensor network. This was done by running an extended simulation with sensor values locked at either true or false to determine convergence points. These simulations clearly show our modeled probabilities level off at three distinct levels around 0%, 60%, and 100%. As a result, we address our cases using two alert levels. When the probability of a person standing rises above 50%, a *warning* is issued to staff members through the wireless network. When the probability of a person standing rises above 90%, an *alert* is issued to staff members. A warning is an indication that a situation needs to be dealt with soon and an alert can be seen as an emergency.

B. Simulation of the Initial Model

The first case we investigate was a simulation of likely sensor readings that would occur when a resident attempted to stand up to get out of his/her chair or bed. The series of simulated readings with no pressure or motion activity for a period of 20 slices were executed to show a person at rest. This was then followed by 20 time slices of motion readings, but without pressure readings to show someone struggling to get up. Next, 20 time slices were simulated with both motion and pressure sensors alerting to show both a person rising off of his/her chair or bed and someone responding to the problem to help them get settled back down again. This was followed by 20 more slices of motion only as a person calms down and staff members leave the room and then once again 20 time slices of no motion or pressure disturbances.

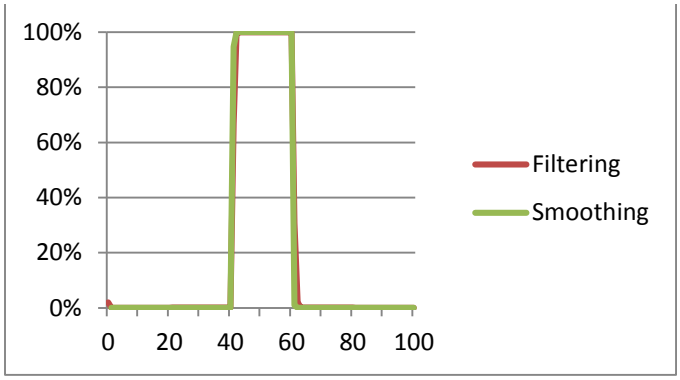


Fig. 2. A simulation of sensor readings from a resident attempting to stand (Model 1).

Figure 2 shows the good performance of the model. The probability of a person standing remained extremely low until the moment both the pressure and motion sensors were activated (t=40). Within 1 time slice the probability that someone was standing had risen above our warning level to over 66%. After an additional time slice, the probability that someone was standing had risen above our alert level to 99%. Similar response times can be seen after the person is returned to his/her chair or bed with it taking only 1 time slice for the probability to drop from above 99% to about 30%. Assuming sensor readings were occurring once a second, this seems very much in line with what would occur with an audio alarm based sensor system. The smoothed data model shows little difference in its probability predictions producing virtually identical readings over the entirety of the use case.

The second case we examined was the situation where a person shifted in his or her chair or bed and dislodged the pressure sensor. This common occurrence in assisted living centers can occur accidentally, or occasionally a resident will find the pressure pad uncomfortable and pull it out not realizing why they are sitting on it. For this use case, we started again with a 20 time slice period of no activity from either our motion or pressure sensor. This was again followed by 20 slices of activity registering from the motion sensor without activity from the pressure sensor. Then, 5 time slices of activity from both sensors to indicate that a person has successfully dislodged the pressure sensor from underneath him or herself. This is then followed by a 40 slice period where the pressure sensor is active, but there is no motion to indicate that a person has settled back in his or her chair, but without repositioning the pressure sensor. At this point, a staff member arrives to help the person reposition the pressure sensor. As a result, both sensors are registering activity. Once the resident is positioned back on his or her pressure pad, there is another 10 slice period of only motion activity as the staff member leaves the room and then 20 slices of no activity from either sensor.

As can be seen in Figure 3, this case starts off very similar to our first case and once the resident dislodges the pressure sensor (t=20) the probability again exceeds our alert level of 90% within 2 time slices. However as the resident has not actually attempted to stand, the probability quickly recedes to a warning level as they settle back into their chair or bed. Once a staff member arrives to assist in repositioning the pressure sensor (t=65), the probability again shoots up over our alert

ISBN: 978-1-61804-292-7

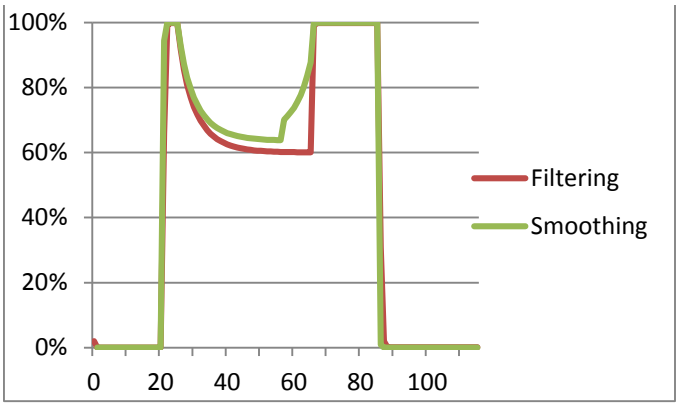


Fig. 3. A simulation of sensor readings from a resident shifting in his or her chair (Model 1).

level of 90% and stays there until the person is properly repositioned on his or her pressure sensor. The probability drops back down to near zero as the staff member finishes moving about the room and leaves. The smoothed data model again closely follows the filtered model with the one exception being from t=57 to 65. Since our smoothing period is 10 slices, this indicates that integrating future observations indicates a higher probability of the resident standing up.

In both of these use cases, the proposed system performed as well as an audible pressure sensor would have alone. In fact, in the case of a person misaligning the pressure sensor without standing up, the performance was quite a bit better. The proposed system was responsive, while eliminating an unnecessary alert condition for the middle period (t=25 to 65). However, in our first case we had an extended time period where after staff had entered the room an alert condition was sustained unnecessarily. In the second case, we returned to an alert condition unnecessarily after a staff member had entered the room. In each of these cases, the extended alert periods may cause additional staff to arrive once a situation was under control. As a result, we looked at implementing a second model and simulation that would account for a staff member being present in the room.

C. Incorporating RFID

Our modified model continues to rely on a pressure sensor and a motion sensor that operate independently of each other. However, the condition they are checking for to be standing without assistance was slightly modified. This is because the third sensor was added to the model, called Radio-frequency identification (RFID). A RFID sensor which scans staff ID badges as they walk through the main door of a resident's room.

The additional sensor was determined not to impact the probabilities for our pressure sensor as whether a staff member has entered the room or not has no effect on whether the sensor is triggered or not. However, the motion sensor probabilities are impacted. As a result, we have adjusted this probability table so that if a resident is standing without assistance and the RFID sensor has signaled that a staff member is in the room, there is a 99% chance that the motion sensor will signal. If the resident is standing without assistance and a staff member is not in the room then the chance that the motion sensor will signal is our original value of 95%. If the RFID sensor has signaled that a staff member has entered the room but the resident is not standing without assistance it is still highly likely that our motion sensor will be activated, so this probability was set at 90%, and if the resident is not standing unassisted and no staff member is present then our original probability of detecting motion applies at 20%.

We also need to establish probabilities of our RFID sensor detecting a staff member in the room given that a resident is standing without assistance. Given that a resident is standing without assistance, the probability that this sensor has detected a staff member would be extremely low. We've set this at 0.1% as if a staff member was in a resident's room it is almost a guarantee that they would be there assisting resident. The probability that a staff member would be in a resident's room if they were not standing without assistance has been set to 10% as it is a reasonable estimate of the percentage of a day a staff member spends with each resident.

It could be argued that a more accurate representation would be to use a second unknown value representing a staff member in the room which the RFID sensor and Motion sensor would monitor. This staff variable could then be used as a sensor variable for the unassisted standing unknown value. However, as the RFID sensor is an almost perfect predictor of the presence of a staff member we chose to use this simplified model.

For each time slice, our modified algorithm follows much the same process as our initial algorithm while incorporating the additional information from the new sensor. The algorithm first calculates the probability that a person is standing based on the previous time slice's standing probability, given the previous sensor evidence. In this example, we are looking at time slice t with an active pressure sensor, an inactive motion sensor, and an inactive RFID sensor:

$$P(US_t|(p_{t-1}, \neg m_{t-1}|\neg r_{t-1}, \neg r_{t-1})) \quad (5)$$
$$= \sum_{s_{t-1}} P(US_{t-1}|(p_{t-1}, \neg m_{t-1}|\neg r_{t-1}, \neg r_{t-1}))$$
$$= P(US_t|us_{t-1})P(us_{t-1}|(p_{t-1}, \neg m_{t-1}|\neg r_{t-1}, \neg r_{t-1}))$$
$$+P(US_t|\neg us_{t-1})P(\neg us_{t-1}|(p_{t-1}, \neg m_{t-1}|\neg r_{t-1}, \neg r_{t-1}))$$

This probability is then used in conjunction with the current sensor readings to determine the current probability of a person standing:

$$P(US_t|(p_{t-1:t}, \neg m_{t-1:t}|\neg r_{t-1:t}, \neg r_{t-1:t})) \quad (6)$$
$$= \alpha P\big((p_t, \neg m_t|\neg r_t, \neg r_t)\big|US_t\big)$$
$$P(US_t|(p_{t-1}, \neg m_{t-1}|\neg r_{t-1}, \neg r_{t-1}))$$

The process was repeated over several time slices to predict what would happen with the modified model under various conditions. Much like our first algorithm, we chose to examine how the predictions were affected when future sensor evidence was known through smoothing. The smoothing process begins by trying to determine what the probability of the future sensor reading occurring is, given the current standing probability:

$$P((p_t, \neg m_t|\neg r_t, \neg r_t)|US_{t-1}) \quad (7)$$

ISBN: 978-1-61804-292-7

$$= \sum_{S_t} P((p_t, \neg m_t | \neg r_t . \neg r_t) | US_t) P(US_t) P(US_t | US_{t-1})$$
$$= P((p_t, \neg m_t | \neg r_t, \neg r_t) | us_t) P(us_t) P(us_t | US_{t-1})$$
$$+ P((p_t, \neg m_t | \neg r_t, \neg r_t) | \neg us_t) P(\neg us_t) P(\neg us_t | US_{t-1})$$

This value is then combined with the current probability of standing without assistance, given the current sensor readings:

$$P(US_{t-1} | (p_{t-1:t}, \neg m_{t-1:t} | \neg r_{t-1:t}, \neg r_{t-1:t})) \quad (8)$$
$$= \alpha P\left(US_{t-1} | (p_{t-1}, \neg m_{t-1} | \neg r_{t-1}, \neg r_{t-1})\right)$$
$$P((p_t, \neg m_t | \neg r_t, \neg r_t) | US_{t-1})$$

D. Simulation Incorporating RFID

As in our initial simulation, we performed smoothing backwards for 10 time slices. Assuming a sensor reading occurred once a second, we felt that this was the maximum possible delay that could be expected if the smoothed value were to be used to alert staff to a problem. A second C++ program was written to accommodate the modified sensor model.

We again attempted to establish points of convergence in our model by running each of the three sensors locked in either true or false state. Although there is a slight variation, these simulations again show three distinct probability levels around 0%, 70%, and 100%. As a result, we will keep using the same two alert levels from our initial model. When the probability of a person standing without assistance rises above 50% we will issue a warning to staff members through the wireless network. When the probability of a person standing rises above 90%, we will issue an alert to staff members. As in our first model, a warning is an indication a situation needs to be dealt with soon and an alert can be seen as an emergency.

In our first case that simulates a person attempting to stand, we see a slightly different result. As before this simulation starts with 20 time slices with the pressure and motion sensors inactive. As this situation is meant to model a person standing up when staff is not around, we have set the initial RFID sensor reading to false as well. At this point, our model predicts a filtered probability of 0.0015%. At t=20 motion is detected as

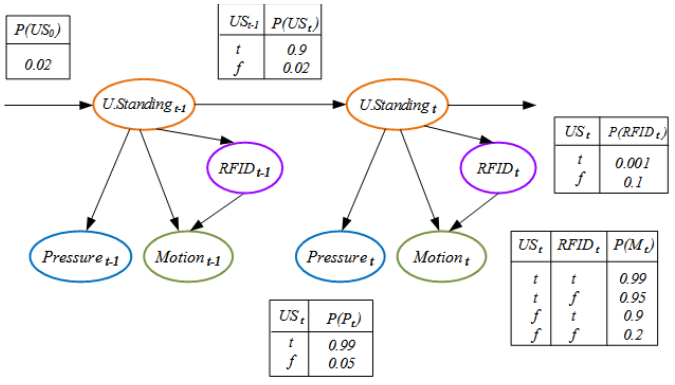


Fig. 4. Modified Dynamic Bayesian Network Model

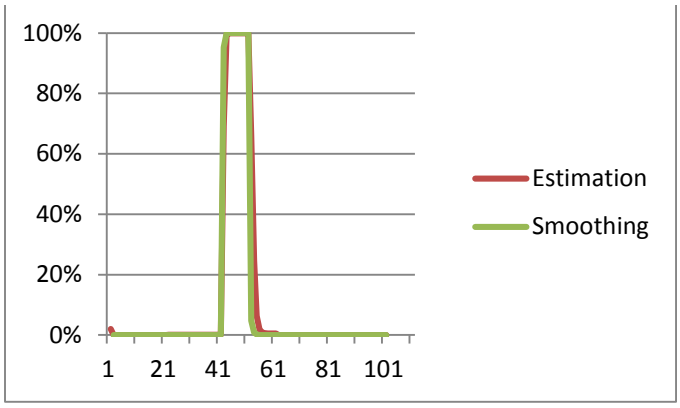


Fig. 5. A simulation of sensor readings from a resident attempting to stand (Model 2).

the resident attempts to stand up. This causes a small uptick in the probability the resident is standing without assistance and the filtered probability rises to 0.11%. At t=40, both the pressure and motion sensors are activated indicating that the person has risen out of his or her chair or bed. Our model quickly adjusts to the situation and within 1 time slice the probability that the resident is standing without assistance has risen above the warning probability to 69.18% and within 2 time slices it as risen above our alert level to 99.43%.

At t=50, the RFID reading changes to true as a staff member responds to the alert condition. This causes an immediate and drastic change in the probability that the resident is standing without assistance. By t=51 the probability has dropped back down to 65.98% and by t=55 it has dropped to 0.84%. At t=60, the pressure sensor becomes inactive. This is to indicate that the staff member has successfully assisted the resident to return to his or her chair or bed. Note that the RFID sensor is still active since the staff member has not yet left the room. This change in sensor readings pushes the probability that the resident is standing without assistance down to 0.00024%. At t=80, the staff member leaves the residents room and the RFID sensor becomes inactive. As expected, this causes a slight increase in the monitored probability to 0.0015%.

As shown in Figure 5, this model does appear to perform much as we expected. Staff are warned of the problem within 1 time slice and alerted to the emergency within 2 time slices. In addition, the alert state is nullified within 2 time slices of a staff member arriving on the scene. This is 9 slices earlier than the initial model without a loss in responsiveness.

The second case takes another look at what would happen when a resident dislodges his or her pressure sensor without actually attempting to stand up out of his or her chair or bed. Much like the previous simulation, we start this simulation with all three sensors off for 20 time slices. At t=20, the probability that the resident is standing without assistance is 0.0015%. At this time, both the pressure and motion sensors are activated for 5 time slices as the resident shifts in his or her chair and dislodges the sensor. As in our first use case, within 1 time slice, we have exceeded our warning level as the

ISBN: 978-1-61804-292-7

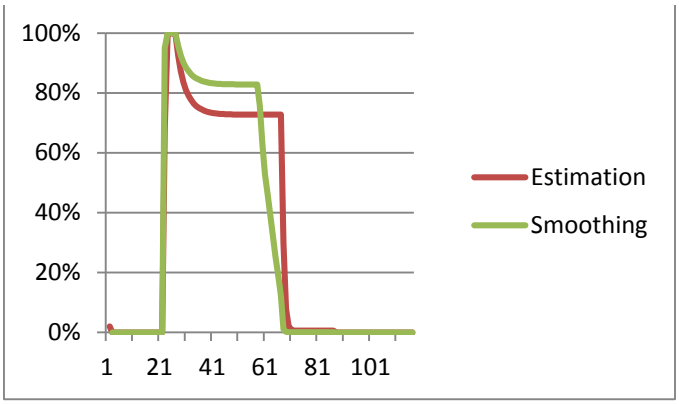


Fig. 6. A simulation of sensor readings from a resident shifting in their chair (Model 2).

probability of the resident standing without assistance has reached 68.07% and by t=22 the probability has exceeded the alert level at 99.41%. At t=25, the motion sensor starts reading false as the resident has settled back in his or her chair or bed.

By t=27, the probability has dropped down to the warning level at 87.3%. The probability that the resident is standing without assistance will continue dropping for several more time slices, but will level off at roughly 72.7%. At t=65, the RFID sensor is activated as a staff member responds to the warning condition. This causes an immediate drop in the probability and at t=66 it has dropped below the warning level to 29.71% and by t=77, the probability has converged on 0.55%. At t=85, the staff member has finished assisting the resident to properly adjust the pressure sensor and is preparinng to leave the room. As a result, the motion sensor and RFID sensors are still active. This causes an additional decrease and at t=95, the probability has reached 0.00024%. As the staff member leaves the room, the motion and RFID readings cease, as in the previous model, which causes a slight increase in the probability to 0.0015%

As can be seen in Figure 6, this second model completely eliminates the second peak in probability that occurred in the initial model for this same use case. Looking back to the initial model in Figure 3 when the staff member entered the residents room at t=65, the probability that the resident was standing shot back up above the error condition and stayed there until t=85 when the resident was again situated back on the pressure sensor. Using our second DBN at t=65, the RFID sensor is activated as the staff member enters the room.

This immediately suppresses the probability down to a negligable level and not only eliminates the occurance of the error condition, but also deactivates the warning condition a full 20 time slices sooner. This does not lead to a loss in responsiveness as a staff member has already arrived in the residents room to assist them when the warning condition is terminated.

IV. Conclusion

Through this study, we studied 3 major questions. First, it was possible to reduce the number of instances where staff members were errantly alerted to an emergency situation thought to be a physically unstable resident attempting to get out of his or her wheelchair or bed. This situation is commonly caused by misaligned pressure sensors. Although it is hard to determine this conclusively from a simulation, our second model suggests that this can be accomplished. Figure 6 shows this combination of sensors and a DBN were able to alert staff within a single time slice that an erratic accident occurs in the resident's room. Although there was a slight adjustment period at t=20 where an emergency was signaled, the signal was quickly adjusted to the warning level allowing staff members to prioritize their current tasks before arriving and adjusting the pressure sensor. Alerting staff members using the wireless network rather than using an audible noise could also keep the resident from becoming agitated in a situation where it is clearly unnecessary.

In addition, Figure 5 demonstrates that a simulated situation where a physically unstable resident was genuinely trying to stand up without assistance could also be properly identified and staff could be alerted and when a staff member arrived to alleviate the situation the sensor network was able to adjust completely out of both the alert and warning levels 2 time slices. Assuming time slices are equal to seconds these delays would be in line with the amount of time it would typically take a staff member to respond to an audible alert.

Lastly, simulating long sequences of static sensor readings for our models allowed us to identify a warning level of 50% and an alert level of 90% probability of a physically unstable resident attempting to stand. If the DBN produces a probability above our warning level it can be seen as a situation that needs to be dealt with soon and production of a probability exceeding the alert level can be seen as an emergency that needs to be dealt with immediately.

References

[1] Alzheimer Society of Canada, "Rising Tide: The Impact of Dementia on Canadian Society", 2010 .

[2] D.Estrin et al., "Embedded, Everywhere: A Research Agenda for Networked Systems of Embedded Computers," National Research Council Report, 2001.

[3] D. Guan et al., "Review of Sensor-based Activity Recognition Systems," IETE Technical Review 28.5, September 2011, p. 418.

[4] G. Vincent, V. Velkoff, "The Next Four Decades The Older Population in the United States 2010 to 2050," U.S. Department of Commerce Economics and Statistics Administration, 2010.

[5] J. Hoey et al, "People, Sensors, Decisions: Customizable and Adaptive Technologies for Assistance in Healthcare", ACM Transaction on Interactive Intelligent Systems Volum 2 Issue 4, December 2012

[6] M. Grzes et al, "Relational approach to knowledge engineering for POMDP-based assistance systems as a translation of a psychological model", International Journal of Approximate Reasoning, March 2013, pp. 36-58

[7] M. Lustrek, B. Kaluza, "Fall Detection and Activity Recognition with Machine Learning," Informatica 33, 2009, pp. 197-204

[8] N. Thome, S. Miguet, S. Ambellouis, "A Real-Time, Multiview Fall Detection System: A LHMM-Based Approach," IEEE Transactions on Circuits and Systems for Video Technology Vol. 18, November 2008 pp. 1522-1532.

[9] S. Khan, M. Karg, J Hoey, D. Kulic, "Towards the Detection of Unusual Temporal Events during Activities Using HMMs" Proceedings of the 2012 ACM Conference on Ubiquitous Computing, 2012, pp. 1075-1084

[10] Y. Chang, C. Shu-Fang, A. Chuang, "A gesture recognition system to transition autonomously through vocational tasks for individuals with cognitive impairments," Research in Developmental Disabilities, 2011 pp. 2064-2068.

UPCEO, connecting statistics and people using R

Pau Fonseca i Casas, Raül Tormos, Josep Casanovas

***Abstract*—** A methodology and a tool that implements this methodology are developed using R to construct a web site that allows a lay user to consult statistical information owned by an institution and stored in a cloud database. This methodology was developed following the open-data philosophy and was implemented with open-source software using R as a key element. The proposed methodology was applied successfully to develop a tool to manage the data of the Centre d'Estudis d'Opinió, but it can be applied to another statistical center to enable open access to its data. The system is deployed on a cloud infrastructure that scales according to demand, implementing a 24/7 solution. A user (or a computer program) can access the information on the website using the R language as a communication channel or using a programming application interface. Additionally, in the R language, a common framework can be defined to structure the various processes involved in any statistical operation.

***Keywords*—** Web; Cloud; R language; R-Serve; API; Surveys

I. Introduction

THE primary goal of the project is to develop a methodology that leads to the implementation of a tool to analyze statistical information online. This research has various facets. First, a mechanism must be defined to manage the large amount of data generated by the surveys and the studies, ensuring that the information remains safe and that the analysts can work with it. Second, a mechanism is required to define what information can be published on the web and what information is not ready to be published (e.g., information that must be anonymized). Finally, a mechanism is required to allow mass media, other research institutions, and the general public to work with the data to obtain new information. To solve these problems, a methodology was defined with the aim of simplifying the interaction with the data of all the actors involved.

The tool that implements the proposed methodology (named UPCEO) addresses all of these various aspects; the last feature described in this paper allows the interaction of the users with the data.

This project pursues the idea of open data, i.e., certain data should be freely available to everyone who desires to use them and republish them, as they wish. The concept of data open to everyone is not new. It was established with the formation of the World Data Center system (WDC) during the International Geophysical Year in 1957 – 1958 [1]. In the beginning, the WDC had centers in the United States, Europe, the Soviet Union and Japan, now it includes 52 centers in 12 countries. The Science Ministers of the Organization for Economic Co-operation and Development (OECD) signed a declaration stating that all the information created or found by the public must be freely available [2]. Following this direction, certain legal tools, such as Open Data Commons [3] came into existence to simplify the use of Open Data over the Internet. In that sense, several tools exist that allow the final user to access information, such as the system in [4], a website devoted to the representation of information on a map, or the Socrata® system [5], a system that supports some interesting applications, such as Data.gov [6] that has the primary mission "*.. to improve access to Federal data and expand creative use of those data beyond the walls of government by encouraging innovative ideas (e.g., web applications).*"

There not only exist several websites and tools to access information but also several applications that allow the reuse and sharing of code related to the access of public information, such as [7] or [8]. The next step is to allow users without technical knowledge to access the information and perform easy tasks with it. To do this, the user must be able to execute tasks on a remote server that stores both remote information and certain statistical functions.

The possibility to allow end-users to execute certain statistical functions to obtain new information from the data were described by [9]. Several different tools exist to show information over the web and allow the execution of statistical functions by the end users, e.g., the NESSTAR system [10]. In parallel with these proprietary solutions, several efforts are focused to develop APIs to access statistical information. As an example, Data.org is preparing an API that allows users to interact with the system data to build their own applications and mash-ups; the [11] has also implemented an API to interact with its data. However, the question of how to develop and use these APIs remains. Every infrastructure that develops this type of solution implements a new API, and the developers must be able to address all of them.

Another problem is related to the data preparation; several alternatives exist to define the surveys, e.g., [12] or [13]. These tools allow the user to export the data to various formats to

Pau Fonseca i Casas, Universitat Politècnica de Catalunya-BarcelonaTech, Barcelona, CA 80034 ESP (corresponding author, phone: (+34) 93 401 7732; fax: (+34) 93 401 5855; e-mail: pau@fib.upc.edu).

Raül Tormos, Centre d'Estudis d'Opinió, Barcelona, CA 08009 ESP (e-mail: rtormos.ceo@gencat.cat).

Josep Casanovas, Universitat Politècnica de Catalunya-BarcelonaTech, Barcelona, CA 80034 ESP (e-mail: josepk@fib.upc.edu).

ISBN: 978-1-61804-292-7

perform posterior analyses (a well-known format is the Triple-S, an XML for survey software that enables the user to import and export surveys between different software). The main issue with this approach is that manual operations are required to process the data. In our proposed approximation, once the surveys are completed by the users, they can easily be uploaded in the system, and all of the answers can be related directly to the historical representation of each of the proposed questions.

II. The Proposed Solution

The statistical institutions that desire to publish complex studios often deal with complex and unstructured data. For this, we propose a methodology based on the R language [14] [15] that simplifies the CRUD (create, read, update and delete) operations that can be performed over the data. To be capable to interact with the data, it is necessary to define a flow for the statistical studies that a statistical institution wants to publish. To do so, it is first necessary to categorize the data that we own in the system. We have the surveys that are the elements that lead to obtaining information from the representative sample of the population of study. These surveys must also be managed by the system. In our proposal, they are represented by an initial matrix of data, containing the questions (and the answers to these questions). Because a survey can be related with other surveys (to obtain information over time), it is necessary to define a superstructure to relate the various initial matrixes between them at two levels: at the matrix level, and at the table-field level.

Additionally, often the data obtained from the survey cannot be published (maybe some information contained in the data are not anonymous), and hence some transformations to the data must be performed to assure the perfect anonymity of the data. After this is performed, several versions of a study can be published, for example, to correct errors detected in the data. The public must have access only to those matrixes of data that pass the necessary quality control, and the other matrixes are stored on the system as working matrixes but are not accessible to the general public. Every study has descriptors to identify the nature of the study and an identification number. For each one of the studies, at least one matrix representing the survey exists. All of the versions obtained from this work are stored in the study structure. Usually, this implies modifying the matrix structures or adding new information. For that, a working matrix exists, representing the last up-to-date matrix related to the studies. The definitive matrix is the matrix that the users can operate using R operations.

Because various matrixes exist, different roles must be defined. Table 1 presents the minimum roles we propose to achieve with this approach. Each one of these roles has different privileges in the final application. For example, an *analyst* can add new studies, add new matrixes to the system, and modify *working* matrixes, whereas an *external* user can only perform the statistical operations allowed by the system with the *definitive* matrix.

Table 1. System roles.

Role	Description
Administrator:	Controls access to the system and defines the roles of the other users.
Analyst:	Manages the information related to the studies (matrix, documentation, etc.)
External:	Can access the system to perform specific operations.

To manage the matrices of data and allow a modification of these data over a cloud infrastructure, worldwide organizations are developing approaches to share statistical information over the Web using an API. From our point of view, this is not enough to address statistical information and data because of the inherent complexity of its nature, and this approach requires continuous modifications of the API functions to accommodate them to the new requirements of the users and institutions that use these data. In our approach, a statistical language is used, to provide a common mechanism to access all the information. The data contained in the proposed platform can be published over the internet using the statistical language itself. The result is that the user can interact with the system using the full power of the selected language, and there is no need to define new functions through the API to interact with the data.

1.1 Beyond the API, using the R language

In our approach, we select the R language [14] due to is power and because is a widely accepted language in the statistical community. R is a free software environment for statistical computing and graphics; see [16] [17] or the web site http://r-project.org. R software can be executed on a wide variety of UNIX platforms, on Windows, on Linux and on MacOS.

This approach is opposite to the approach followed by API development. In this approach, the system allows an authorized user, or program, to access the data and obtain, using R syntax, all the data and information desired. The concern is related not with the implementation of new APIs or protocols to allow access to specific statistical information or data but with limiting the amount of information that can be obtained over the web. This implies limiting the R operations that can be implemented on the server. Fortunately, this configuration can be accomplished through the RServe package [18], which allows the user to define what instructions can be used over the web.

The power of R does not rely only on strong statistical and graphical facilities but also on versatility. Any element of the research community can improve the system by adding new modules to perform statistical operations. One of the packages we need for our approach is RServe. R usually works in standalone applications, and to connect the different services to R, the R-Serve package must be used. R-Serve can be executed from a command. RServe is a TCP/IP server that allows other programs to use the R facilities from various languages without the need to initialize R or link to the R library [19]. Each

connection has a separate workspace and working directory, which is an essential feature for this project.

The sequences to start using the service are (i) start the R console, (ii) on the console, load the RServe library, and (iii) start the RServe server.

For most users, the default configuration is satisfactory; however, for this project, RServe must be configured to coordinate the different elements that comprise the system. RServe usually works with several default parameters that can be modified in the *config* file. The configuration file is located at /etc/Rserv.conf (on a Linux server, this location can be changed during compilation, specifying the option -DCONFIG_FILE=<new path>). New configuration files can be added with the command --RS-conf (this is an argument in the command line). The complete documentation of the package can be found in [18].

1.1.1 Using R on the statistical study lifecycle

Three main areas must be covered: the management of a questionnaire (starting a new study), the management of the matrixes related to the study, and the management of the operations that can be applied to the public matrixes of the study. In each one of these three areas, we propose to use R language as a basic element to simplify the interaction. This leads to a simplification in the maintainability and further expansion of the system.

To prepare a new questionnaire, first and foremost, the questions must be defined. This is not an easy task because of the diversity of questions that can appear in a single questionnaire and also because the various surveys must consistently be related to each other to make it possible to obtain accurate conclusions over time. Various alternatives exists to prepare surveys, e.g., [12], or [13]. Using these alternatives, the questions can be defined, and they can be sorted on questionnaires that the respondents must answer. Often, these alternatives can export the data to various formats for posterior analysis (such as Triple-S). In our proposal, the relations between the various questions that compose the questionnaires must also be defined; this information (which can be stored in the database for its posterior use) helps us in the review of the complete history of the questions. The answers to the various questionnaires and the history of changes are also available. For example, if we include a question such as, "What party would you vote for in the next election?" and in a new version of a questionnaire, it changes to "If elections were to be held tomorrow, what party or coalition would you vote for?" we must keep the relation between both questions, indicating that they represent the same underlying concept. This simplifies the statistical use in the operations tool, merging the information to construct, for example, a time series.

In that sense, the present approach simplifies the ulterior data management; however, this implies that the uploading process is not easy because it is necessary to create the relationships of the questions, surveys and answers in the database. Additionally, the matrix files can be large and represented in various formats. In our approach, all the information is transformed to a specific XML file that always has the same structure. This enables the user to work with surveys that have the answers in several formats, such as Excel, SPSS, Minitab or R, among many others.

Thanks to the use of an XML base representation for the uploading and management of the data matrixes, it is possible to incorporate tools that access the questions. These questions can be presented to the user in various ways, i.e., *editions*. All of the editions of a question can be related, simplifying the operation of merging surveys. The users can build a new questionnaire, and after the questionnaires are defined in the system, they can be related in a matrix that contains the data obtained from the respondents. The key element of our proposed approach is to always retain the relation between the questions, the questionnaires and the answers.

Finally, and because we propose to use the R language, the users can execute the operations written in R (from a subset of the allowed operations) with the data loaded on the system.In this approach, the relation between all of the various questions is preserved. Additionally, the R language will be used as an API to obtain information from the system instead of defining an API.

III. THE UPCEO APPLICATION

Three institutions are involved in this real project, the *Centre d'Estudis d'Opinió* (CEO), the InLab FIB and the Centre de *Telecomunicacions i Tecnologies de la Informació* (CTTI). The CEO is the official survey institute of the Generalitat de Catalunya. It handles the government's political surveys, barometers, election studies, and other public opinion polls in Catalonia. As defined in their institutional functions, "It is a tool (the CEO) of the Catalan government aimed at providing a rigorous and quality service to those institutions and individuals interested in the evolution of Catalan public opinion." One of its commitments is to make the information readily accessible to the public.

InLab FIB is an innovation and research lab based in the Barcelona School of Informatics, Universitat Politècnica de Catalunya - Barcelona Tech (UPC) that integrates academic personnel from various UPC departments and its own technical staff to provide solutions to a wide range of demands that involve several areas of expertise. InLab FIB, formerly LCFIB, has more than three decades of experience in developing applications using the latest ICT technologies, collaborating in various research and innovation projects and creating customized solutions for public administrations, industry, large companies and SMEs using agile methodologies.

The *Centre de Telecomunicacions i Tecnologies de la Informació* (CTTI) [20] is an infrastructure that can host all of the services that the various organizations that belong to the *Generalitat de Catalunya* requires. This infrastructure is maintained by a licensed private enterprise (now T-Systems). This is convenient for the project because, when the CEO publishes a new study, the quantity of resources required to supply the punctual demand can be bigger than the resources required in a usual day. Additionally, because CTTI ensures that the system is working 24/7, it can be convenient for the

daily work to provide the infrastructure for the CEO database to store all of the information regarding the studies. The CEO primarily manages surveys related to political public opinion. The studies derived from these surveys are published on the CEO website to ensure that the public has knowledge about the studies.

We implement a system to simplify the management and use of statistical information over a web. The specific implementation is represented in Figure 1. The system is composed of different layers, each one of which is related to the various services that the system must provide. The web server is based on a WebLogic Oracle® application [21], using Apache Struts [22] [23] and Java as the infrastructure to define the interface of the system and to establish communication with the R system. The main purpose of using R is to implement various operations that deal with data (see 1.1.1). As an example, we use R to obtain the data from the matrix and the surveys that usually are in the original form of Excel spreadsheets, SPSS files or SAS files; here, R is used as the bridge between all of the various file formats. The R language can be used by users and other applications as an API to communicate with the system to obtain statistical data. In Figure 1, the structure of the system is shown. The entire system is on the CTTI cloud infrastructure. The various files related to the application are stored on an NAS system. The studies are stored in an Oracle database to manage the various files of the system. The R application is installed on the system with the RServe package, defining a set of operations (as an API) and publishing them on the internet using the WebLogic platform.

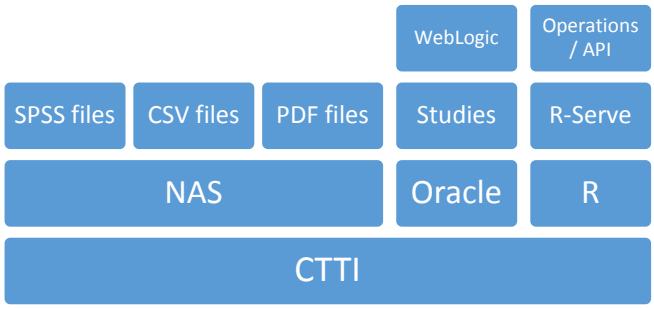


Figure 1. System structure.

From an operations point of view, when a user requests a specific study, he obtains its related documents, mainly *.pdf* files and links to other data related to the survey. With these data, the user can perform various operations (with R), obtaining new data and information. These results can then be exported in CSV file format that can be analyzed in more detail using any statistical package. As shown in Figure 1, the matrix is stored in its original form on the NAS, implying that various formats must be stored in the system. This way, the information generation process can be reproduced exactly as it was by the analyst.

The main file formats that can be used by the CEO analyst are Excel spreadsheets, SPSS *.sav* files and *.csv* files. R is a key element to manage this diversity of formats. Because the application uses R, the information can be read and operated. R can also store or export the new matrix of data in a new format that can be stored again in the database or managed by an external user.

The various functionalities in the system are:

Questionnaire manager manages the questions related to each one of the different questionnaires of the system; see Figure 2 and Figure 3. In our approach, all of the questions must be related to allow a temporal analysis of the data stored on the database.

Matrix manager manages the information related to the matrix generated by the surveys; see Figure 4.

Operation shows the information to the users and other applications (websites) through the R language.

The application can be accessed at http://ceo.gencat.cat/ceop/AppJava/pages. The website is in the Catalan language, and the option that gives access to the operations is "*Banc de dades del BOP*," located at the bottom of the page. This option leads users to the page where a specific study, http://ceo.gencat.cat/ceoa/AppJava/OperacionsExtern.do, is found. This initial listing shows the latest studies performed by the CEO analysts.

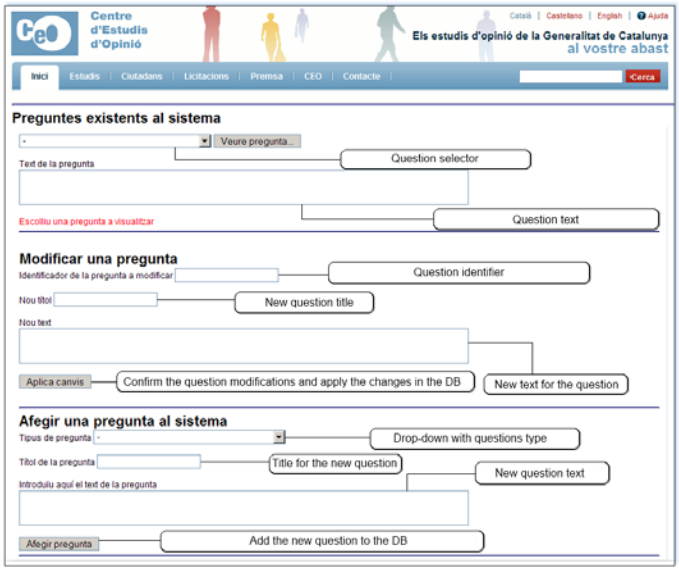


Figure 2. The process of creating a new question is integrated into the application, simplifying the process of reuse and relating the questions of all the questionnaires that exist in the system, as is proposed by our approach.

ISBN: 978-1-61804-292-7

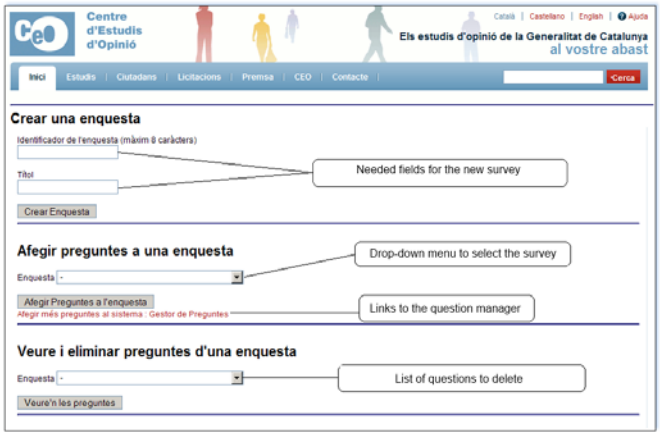


Figure 3. The process of defining a new survey can be performed entirely in the application, simplifying the survey management, as well as its posterior use.

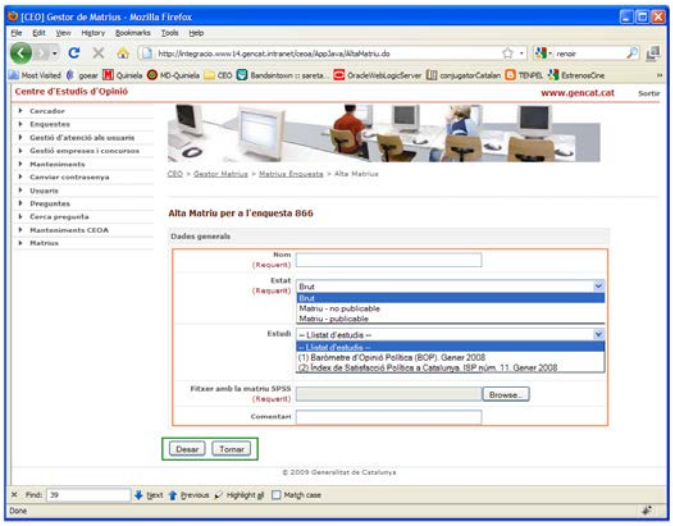


Figure 4. Uploading a new matrix containing the data of a survey to the system.

IV. UPCEO Implementation and Calibration

The entire application resides as a cloud solution supported by the *Generalitat de Catalunya*, hosted by the *Centre de Telecomunicacions i Tecnologies de la Informació* (CTTI). In this cloud solution, the options to work and to modify the upload code are limited, as is explained in section A. Because of the complexity of the structure and the required security concerns, a test infrastructure was implemented to test and implement the R operations. The test infrastructure is composed of a server and a client. On the server side, a machine acts as a Web server (using IBM WebLogic), hosting the MySQL database, storing the data on the NAS (Network Attached Storage) and executing R-Serve. On the client side, a java program (implemented on NetBeans and named JGUIforR; see Figure 5) is used to define the GUI and the R code needed to execute the operations and manage the matrixes.

The client application must first be connected with the server side. The IP of the R server instance we want to use is defined. In this case, the application is connecting with a server that is executed on the same machine as the JGUIforR.

Once this is completed, the connection with the server is established using the File menu. Two options are available. **RComand** implies that the user is working with a local instance of R. In that case, it is not necessary to define the IP. **RComandTCP** implies that the user is working with a remote instance of R; in that case, the IP of the remote server must be defined.

If the connection is established without error, a message appears in the **R Comands** window showing the version of the R engine used on the server side.

To start working, a dataset must be selected, in this case, an SPSS® dataset. Opening a new dataset is as easy as going to the File menu and selecting a new **Matrix** of data.

Once the matrix is loaded, a message is shown to the user in the **R Comands** area, as shown in Figure 5. At this point, all the operations are active, and the user can start working with the matrix.

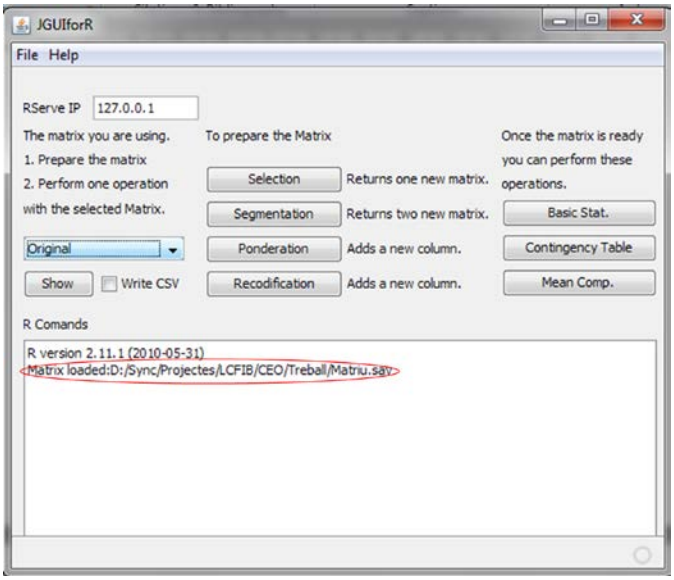


Figure 5. Matrix successfully loaded. All of the options are now activated, and the user can start working with the matrix. The source code of JGUIforR can be downloaded for free at https://svn.java.net/svn/jguiforr~jguifor/.

CEO analysts use this software to understand the operations that the system publishes and to understand the behavior desired in the final implementation of the client, using, in that case, Apache struts [22] to build the website.

As shown in Figure 5, the operations are divided into two main groups. The first includes the preparation of the matrix, selection of a portion of the data of the entire matrix, segmentation of the matrix, weighting of some of the columns of the matrix and recodification. The other operations that can be executed operate over this matrix (calculating the mean, the max, the min values, compiling a contingency table or performing a mean comparison between two variables, etc.).

A. Deploying the system

After the operations perform as expected on the Java platform, the system can be deployed on the CTTI infrastructure. This project represents the first deployment of

ISBN: 978-1-61804-292-7

RServe on the CTTI infrastructure, which implies the need to define roles and protocols to ensure 24/7 support. The system also has high security concerns. First, the application is deployed on the working server, a machine accessible only to the computers located at the InLab FIB laboratory. Once the application passes the tests on this machine, it is deployed at the integration level of the CTTI infrastructure. Here, the application is tested in an environment that is not equal to the production environment but has similar security levels and the same software. After the application performs well there, it can be deployed to a preproduction level. Here, the application runs on an exact replica of the final infrastructure, on the same hardware and executing the same software that the application will find in the production environment. At this level, a set of tests are performed, and the application must pass all of them to be deployed to the production level.

At the production level, the application is available for public use. This is the last step of the deployment, and the current state of the case presented here. Once the system is deployed, the operations performed by the user must never modify the information stored in the server. The system must also be able to store information regarding the various activities that each of the users performs.

When an operation is selected, the R syntax is stored in the database. This syntax is not executed immediately on the system; it is only executed when the user requests results (for example, executes the operations of basic statistics, a contingency table or a mean comparison). This is because the time required to perform an operation bottlenecks at the transference of the data and establishing the connections between the client and RServe. After the connections are established (less than a minute), R performs well and returns the new data very fast.

V. Concluding remarks

This study develops a novel approach to present statistical information over the web following the open-data philosophy. In this approach, the R statistical package is a key element to manage and display the information, allowing the user to perform a number of statistical operations with the data.

From the point of view of data management, the structure of the surveys, the structure that relates the questionnaires and the questions and the related matrix that contains the data, often follow different formats in a real environment. This is true even if a single team manages the information because technology changes and the tools used can be diverse, depending on the objectives of the specific work. This ecosystem of data formats often makes working with the data more difficult. Thus, mechanisms are necessary to translate the information from one format to another. Often, these mechanisms are prone to errors and require the use of tools that are often not well-known by all of the members of the team. In this approach, R is the bridge between the various formats that are stored in the database and is also the language used to recover and work with the information contained in the system. Thus, the CEO analysts store the information in the system using the format they use and understand, and the system is able, using R, to work with the data and to formulate new matrixes of data that can be used again by the experts using their common statistical tools.

Because the system must be able to work at all times, a cloud solution must be implemented to simplify the management of the infrastructure. The amount of access of the external users depends on several factors, e.g., when a new study is offered to the public. This implies that, at times, the traffic to the site is heavy, an aspect that can become a problem for the servers and site management. The cloud solution proposed stores all the information obtained from the CEO studies, allowing 24/7 access to all the information by all the users, and allowing, depending on the user role, the manipulation of the data and the creation of new information and matrixes. Working with the data is accomplished using R as a statistical engine; a user can execute queries and obtain new information regarding the matrixes of data related to a survey. Additionally, because all the operations implemented use R syntax, adding new operations is easy and only requires the addition of a new R code and the definition of a new interface. Thus, the systems implemented based on this approach are extremely scalable and expandable.

Since all of the access to the statistical information is based on the R language, new websites or applications (such as JGUIforR) can be developed that access the data through the use of R statements. This implies that the application goes further than the definition of an API because it uses a statistical language. The power and extensibility of R ensures that we can obtain all the information needed, and the user must only define the subset (if it is needed) of the R instructions that an external user (application or website) can execute. Currently, researchers from various Catalonian institutions are building their own mash-ups using the application. In the future, more capabilities will be added to the application by adding new R language instructions open to public use. There is an additional goal of open access to the institutions, allowing them to access all the information from the CEO servers and define the queries they need for each application (in the broad sense that an application can be a simple query that can reside in a spreadsheet, or a complete web application with various mash-ups).

Last but not least, a set of operations can be defined as an R script. This definition implies that repetitive operations can be performed with fewer errors and in less time.

VI. Acknowledgements

This paper is the result of hard work done by InLab FIB and CEO for three years. We wish to thank the different personnel that are involved with different stages of the project, especially Marta Cuatrecases, Joan Giralt Duran, Sara Royuela Alcazar, José Francisco Crespo Sanjusto, Albert Carrera Mateu and Xavier Canal Masjuan, as members of the InLab FIB that actively developed the application, and Rosa Maria Capo as a member of the CEO that helped us in the development of the tool.

VII. References

[1] World Data Center, "World Data System of International Council for Science," 2010. [Online]. Available: http://www.icsu-wds.org/. [Accessed 11 11 2011].

[2] Organisation For Economic Co-Operation And Development, "OEOECD Principles and Guidelines for Access to Research Data from Public Funding," 2007.

[3] Open Knowledge Foundation, "Legal tools for Open Data," 2011. [Online]. Available: http://opendatacommons.org/. [Accessed 11 11 2011].

[4] open3, "DataMaps.eu," 2011. [Online]. Available: http://www.datamaps.eu/. [Accessed 11 11 2011].

[5] Socrata, Inc, "Socrata, The Open Data Company," 2011. [Online]. Available: http://www.socrata.com/. [Accessed 11 11 2011].

[6] Federal Government, "Data.gov Empowering People," 2011. [Online]. Available: http://www.data.gov/. [Accessed 11 11 2011].

[7] Code for America Labs, Inc , "Code for America," 2011. [Online]. Available: http://codeforamerica.org/. [Accessed 14 11 2011].

[8] Leipziger Agenda 21, "API.LEIPZIG," 2011. [Online]. Available: http://www.apileipzig.de/. [Accessed 14 11 2011].

[9] B. Sundgren, "Making Statistical Data More Available," in *Workshop on R&D Opportunities in Federal Information Services.*, Virginia, USA., 1997.

[10] T. Assini, "NESSTAR: A Semantic Web Application for Statistical Data and Metadata.," in *WWW2002 Conference.*, Hawai, 2002.

[11] New York State Senate, "NYSenate.gov Application Protocol Interface (API)," 2011. [Online]. Available: http://www.nysenate.gov/developers/api. [Accessed 14 11 2011].

[12] Snap Surveys Ltd, "Online surveys," 2012. [Online]. Available: http://www.snapsurveys.com/. [Accessed 20 10 2012].

[13] University of Ottawa, "Snap Surveys," 2012. [Online]. Available: http://www.ccs.uottawa.ca/webmaster/survey/. [Accessed 20 10 2012].

[14] J. Adler, R in a Nutshell: A Desktop Quick Reference, O'Reilly Media, 2009.

[15] P. Teetor, R Cookbook, O'Reilly Media, Inc., 2011.

[16] F. Murtagh, Correspondence analysis and data coding with Java and R, C. S. a. D. A. Chapman and Hall, Ed., 2008.

[17] M. W. Trosset, An introduction to statistical inference and its applications with R, vol. 81, Chapman and Hall., 2010.

[18] Rforge.net, "Rserve - Binary R server," 2011. [Online]. Available: http://www.rforge.net/Rserve/doc.html. [Accessed 14 11 2011].

[19] S. Urbanek, "RServe," 2010. [Online]. Available: http://www.rforge.net/Rserve/. [Accessed 05 July 2010].

[20] Generalitat de Catalunya, "DOGC núm. 5359 - 15/04/2009," 2009. [Online]. Available: http://www.gencat.cat/diari/5359/09082146.htm. [Accessed 9 9 2010].

[21] Oracle, "Oracle Weblogic Server," 2010. [Online]. Available: http://www.oracle.com/technetwork/middleware/weblogic/overview/index.html. [Accessed 11 11 2010].

[22] Apache Software Foundation, "Apache Struts," 2010. [Online]. Available: http://struts.apache.org/. [Accessed 11 11 2010].

[23] C. Cavaness, Programming Jakarta Struts,, O'Reilly Media, 2004.

[24] D. Moore, "SQL Loader," 2003. [Online]. Available: http://www.oracleutilities.com/OSUtil/sqlldr.html. [Accessed 22 10 2012].

[25] A. Billington, "external tables in oracle 9i," 6 2007. [Online]. Available: http://www.oracle-developer.net/display.php?id=204. [Accessed 20 10 2012].

[26] ORACLE-BASE.com, "XMLType Datatype In Oracle9i," 2012. [Online]. Available: http://www.oracle-base.com/articles/9i/xmltype-datatype.php. [Accessed 20 10 2012].

Pau Fonseca i Casas is an associate professor of the Department of Statistics and Operational research of the Technical University of Catalonia, teaching in Statistics and Simulation areas. He owns a Ph.D. in Computer Science on from Technical University of Catalonia.
He works in the InLab FIB (http://inlab.fib.upc.edu/) as a head of the Environmental Simulation area, developing Simulation projects since 1998. He has been involved in more than 20 competitive projects and has published more than 80 papers on journals, conferences and books. He is also a lecturer on Universitat Politècnica de Catalunya – BarcelonaTech, and collaborates with the Universitat Oberta de Catalunya, teaching in Simulation and Statistics area at degree and master levels. His research interests are discrete simulation applied to industrial, environmental and social models, and the formal representation of such models. His website is http://www-eio.upc.es/~pau/.

Raül Tormos is senior survey researcher at the Centre d'Estudis d'Opinió, the official institute for public opinion studies of the Government of Catalonia (Spain). He is also lecturer at the Autonomous University of Barcelona, the University of Barcelona, and the School of Public Administration of Catalonia. He teaches quantitative methods, comparative analysis, official statistics and survey methodology, both at graduate and undergraduate levels. He obtained his PhD (European Doctor) in political science at the Universitat Autònoma de Barcelona. Earlier, he was awarded a full-year stipend by the European Commission (under the TMR funding scheme) as pre-doctoral research fellow at the Mannheim Center for European Social Research (MZES), University of Mannheim. His research interests involve the study of values, attitudes and political behavior, age-period-cohort effects, quantitative research methods and survey methodology. He has done specialized training at the University of Essex, Universidad de Salamanca, Research and Expertise Centre for Survey Methodology, University of Oslo, and University of California at Berkeley. His research has been published in journals such as European Political Science Review or Revista Española de Investigaciones Sociológicas.

Josep Casanovas is a full professor in Operations Research, specializing in Simulation Systems. He is one of the founders of the Barcelona School of Informatics (FIB), of which he was Dean from 1998 to 2004. He is also the director of inLab FIB, a research lab that has been particularly active in technology transfer to business. Among his recent projects is the cooperation in the creation of simulation environments for people and vehicle flow in the new Barcelona airport terminal. He has led several EU-funded projects in the area of simulation and operations research and is a strong advocate of the knowledge and technology transfer function between university and society.

Multisoliton solutions to a generalized AKNS equations with variable coefficients

Sheng Zhang and Xu-Dong Gao

***Abstract*—**In this paper, Hirota's bilinear method is extended to the new and more general AKNS equations with variable coefficients. As a result, one-soliton solutions and two-solition solutions are obtained, from which the uniform formulae of *n*-soliton solutions are derived. It is shown that the Hirota's bilinear method can also be used for constructing multisoliton solutions of some other nonlinear partial differential equations with variable coefficients.

***Keywords*—**Multisoliton solution, Hirota's bilinear method, AKNS equations with variable coefficients.

I. Introduction

It is well known that nonlinear physical phenomena are often related to nonlinear partial differential equations (PDEs), which are involved in many fields from physics to biology, chemistry, mechanics, etc. As mathematical models of the phenomena, the investigation of exact solutions of nonlinear PDEs will help to understand these phenomena better. With the development of soliton theory, finding multisoliton solutions of nonlinear PDEs has gradually developed into a significant direction in nonlinear science and some effective methods have been proposed and developed, such as the inverse scattering transformation [1], Hirota's bilinear method [2], Bäcklund transformation [3], Painlevé expansion [4], homogeneous balance method [5], and the function expansion methods and some others [6]-[14]. Among these methods, Hirota's bilinear method [2] is a purely algebraic method used for constructing multisoliton solutions of nonlinear PDEs, the process of which is fairly simple and convenient for computer operation. Not only is it applicable to KdV equation, mKdV equation and sine-Gordon (sG) equation, but also can be used for nonlinear differential-difference equations (DDEs) [15]-[22]. More and more studies show that the Hirota's bilinear method is a more extensively applicable approach to solve nonlinear PDEs. Recently, the study of variable-coefficient PDEs has attracted much attention because most of real nonlinear physical equations possess variable coefficients.

In this article, we will extend Hirota's bilinear method to construct multisoliton solutions of the following new and more general AKNS equations

$$q_t = a_3(t)(q_{xxx} - 6qrq_x) + a_2(t)(-q_{xx} + 2q^2 r) + a_1(t)q_x - a_0(t)q, \tag{1.1a}$$

$$r_t = a_3(t)(r_{xxx} - 6qrr_x) + a_2(t)(r_{xx} - 2r^2 q) + a_1(t)r_x + a_0(t)r. \tag{1.1b}$$

which is a special case at $m=3$ of the generalized AKNS hierarchy with variable coefficients

$$\begin{pmatrix} q \\ r \end{pmatrix}_t = \sum_{i=0}^{m} a_i(t) L^i \begin{pmatrix} -q \\ r \end{pmatrix}, \ (m = 1, 2, \cdots), \tag{1.2}$$

where the recursion operator is employed as

$$L = \sigma\partial + 2\begin{pmatrix} q \\ -r \end{pmatrix}\partial^{-1}(r, q), \ \sigma = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \ \partial = \frac{\partial}{\partial x},$$

$$\partial^{-1} = \frac{1}{2}\left(\int_{-\infty}^{x} \mathrm{d}x - \int_{x}^{\infty} \mathrm{d}x\right).$$

If setting $a_m(t) = 1$, $a_{m-1}(t) = \cdots = a_1(t) = a_0(t) = 0$, then from (1.2) we can obtain the following known constant-coefficient AKNS hierarchy [20]:

$$\begin{pmatrix} q \\ r \end{pmatrix}_t = L^m \begin{pmatrix} -q \\ r \end{pmatrix}, \ (m = 1, 2, \cdots). \tag{1.3}$$

It should be noted that (1.1) includes the KdV equation, the MKdV-NLS equation as special cases as long as selecting $a_3(t) = -1$, $a_2(t) = a_1(t) = a_0(t) = 0$, $r = -1$ and $a_3(t) = -1$, $a_2(t) = i$, $a_1(t) = a_0(t) = 0$, $r = -q$, respectively.

The rest of this paper is organized as follows. In Section 2, we extend Hirota's bilinear method for constructing multisoliton solutions of the variable-coefficient AKNS equations (1.1). In Section 3, we conclude this paper.

This work was supported by the Natural Science Foundation of Liaoning Province (L2012404) of China, the PhD Start-up Funds of Bohai University (bsqd2013025) and Liaoning Province of China (20141137), the Liaoning BaiQianWan Talents Program (2013921055) and the Natural Science Foundation of China (11371071).

S. Zhang is with the School of Mathematics and Physics, Bohai University, Jinzhou 121013, PR China (corresponding author to provide phone: 086-416-3400149; e-mail: szhangchina@ 126.com).

X. D. Gao is with the School of Mathematics and Physics, Bohai University, Jinzhou 121013, PR China (e-mail: 986242791@qq.com)

II. MULTISOLITON SOLUTIONS

Firstly, we take the following transformation:

$$q=\frac{g(x,t)}{f(x,t)},\ r=\frac{h(x,t)}{f(x,t)}, \tag{2.1}$$

and then obtain the bilinear forms of (1.1) as follows:

$$\begin{aligned} D_t g\cdot f &= a_3(t)[D_x^3 g\cdot f-\frac{3}{f^2}(D_x^2 f\cdot f+2gh)D_x^1 g\cdot f] \\ &+a_2(t)[-D_x^2 g\cdot f+\frac{g}{f}(D_x^2 f\cdot f+2gh)] \\ &+a_1(t)D_x^1 g\cdot f-a_0(t)D_x^0 g\cdot f, \end{aligned} \tag{2.2a}$$

$$\begin{aligned} D_t g\cdot f &= a_3(t)[D_x^3 g\cdot f-\frac{3}{f^2}(D_x^2 f\cdot f+2gh)D_x^1 g\cdot f] \\ &+a_2(t)[D_x^2 h\cdot f-\frac{g}{f}(D_x^2 f\cdot f+2gh)] \\ &+a_1(t)D_x^1 h\cdot f+a_0(t)D_x^0 h\cdot f. \end{aligned} \tag{2.2b}$$

Further supposing that

$$D_x^2 f\cdot f+2gh=0, \tag{2.3}$$

then (2.2) are reduced to:

$$\begin{aligned} &[D_t-a_3(t)D_x^3+a_2(t)D_x^2-a_1(t)D_x^1 \\ &\quad +a_0(t)D_x^0]g\cdot f=0, \end{aligned} \tag{2.4a}$$

$$\begin{aligned} &[D_t-a_3(t)D_x^3-a_2(t)D_x^2-a_1(t)D_x^1, \\ &\quad -a_0(t)D_x^0]h\cdot f=0, \end{aligned} \tag{2.4b}$$

here the following Hirota's bilinear operator is employed:

$$\begin{aligned} &D_t^m D_x^n g(x,t)\cdot f(x,t). \\ &=(\partial_t-\partial_{t^*})^m(\partial_t-\partial_{x^*})^n g(x,t)f(x^*,t^*)\big|_{x^*=x,t^*=t}. \end{aligned} \tag{2.5}$$

Especially, (2.5) gives $D_x^0 g\cdot f=D_t^0 g\cdot f=0$.

We next construct multisoliton solutions of (1.1) by means of (2.3) and (2.4). For the one-soliton solution, we suppose

$$f=1+\sum_{i=1}^{\infty}\varepsilon^{2i}f^{(2i)},\ g=\sum_{i=1}^{\infty}\varepsilon^{2i-1}g^{(2i-1)}, \tag{2.6a}$$

$$h=\sum_{i=1}^{\infty}\varepsilon^{2i-1}h^{(2i-1)}, \tag{2.6b}$$

and substitute them into (2.3) and (2.4) and then collect the coefficients of the same order of ε yields a system of differential equations

$$\begin{aligned} &g_t^{(1)}-a_3(t)g_{xxx}^{(1)}+a_2(t)g_{xx}^{(1)}-a_1(t)g_x^{(1)} \\ &\quad +a_0(t)g^{(1)}=0, \end{aligned} \tag{2.7a}$$

$$\begin{aligned} &h_t^{(1)}-a_3(t)h_{xxx}^{(1)}-a_2(t)h_{xx}^{(1)}-a_1(t)h_x^{(1)} \\ &\quad -a_0(t)h^{(1)}=0, \end{aligned} \tag{2.7b}$$

$$f_{xx}^{(2)}=-g^{(1)}h^{(1)}, \tag{2.7c}$$

$$\begin{aligned} &[D_t-a_3(t)D_x^3+a_2(t)D_x^2-a_1(t)D_x^1+a_0(t)D_x^0] \\ &\quad (g^{(3)}\cdot 1+g^{(1)}\cdot f^{(2)})=0, \end{aligned} \tag{2.7d}$$

$$\begin{aligned} &[D_t-a_3(t)D_x^3-a_2(t)D_x^2-a_1(t)D_x^1-a_0(t)D_x^0] \\ &\quad (h^{(3)}\cdot 1+h^{(1)}\cdot f^{(2)})=0, \end{aligned} \tag{2.7e}$$

and so forth. If letting

$$g^{(1)}=\mathrm{e}^{\xi_1},\ \xi_1=k_1x-\sum_{j=0}^{3}(-1)^j k_1^j\int a_j(t)\mathrm{d}t+\xi_1^0, \tag{2.8a}$$

$$h^{(1)}=\mathrm{e}^{\eta_1},\ \eta_1=l_1x+\sum_{j=0}^{3}l_1^j\int a_j(t)\mathrm{d}t+\eta_1^0, \tag{2.8b}$$

be two solutions of (2.7a) and (2.7b), from (2.7c) we obtain

$$f^{(2)}=\mathrm{e}^{\xi_1+\eta_1+\theta_{13}},\mathrm{e}^{\theta_{13}}=-\frac{1}{(k_1+l_1)^2}. \tag{2.9}$$

Substituting (2.9) into (2.7d)-(2.7e) and those behind, we can verify that if

$$g^{(3)}=h^{(3)}=f^{(4)}=\cdots=0, \tag{2.10}$$

then (2.7d)-(2.7e) and those behind all hold. In this case, we write

$$f_1=1+\mathrm{e}^{\xi_1+\eta_1+\theta_{13}},\ g_1=\mathrm{e}^{\xi_1},\ h_1=\mathrm{e}^{\eta_1}, \tag{2.11}$$

and hence obtain the following one-soliton solutions of (1.1):

$$q=\frac{\mathrm{e}^{\xi_1}}{1+\mathrm{e}^{\xi_1+\eta_1+\theta_{13}}},\ r=\frac{\mathrm{e}^{\eta_1}}{1+\mathrm{e}^{\xi_1+\eta_1+\theta_{13}}}. \tag{2.12}$$

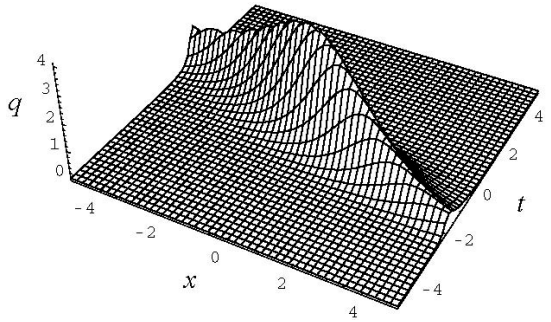


Fig. 1. Spatial structure of one-soliton solution q of (2.12).

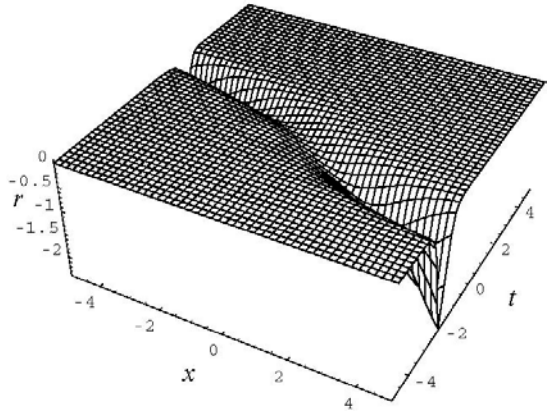


Fig. 2. Spatial structure of one-soliton solution r of (2.12).

In Figs. 1 and 2, the spatial structures of one-soliton solutions (2.12) are shown, where the parameters are selected as $k_1=1.2$, $l_1=1$, $a_0(t)=0.5\mathrm{sech}t$, $a_1(t)=\mathrm{e}^t$, $a_2(t)=\sin t$, $a_3(t)=1+0.5t^2$, $\xi_1^0=0$, $\eta_1^0=i\pi$.

If selecting

$$g^{(1)} = \mathrm{e}^{\xi_1} + \mathrm{e}^{\xi_2},$$

$$\xi_i = k_i x - \sum_{j=0}^{3} (-1)^j k_i^j \int a_j(t)\mathrm{d}t + \xi_i^0,\ i = 1,2 \qquad (2.13a)$$

$$h^{(1)} = \mathrm{e}^{\eta_1} + \mathrm{e}^{\eta_2},$$

$$\eta_i = l_i x + \sum_{j=0}^{3} l_i^j \int a_j(t)\mathrm{d}t + \eta_i^0,\ i = 1,2 \qquad (2.13b)$$

then a direct computation gives

$$f^{(2)} = \mathrm{e}^{\xi_1+\eta_1+\theta_{13}} + \mathrm{e}^{\xi_1+\eta_2+\theta_{14}} + \mathrm{e}^{\xi_2+\eta_1+\theta_{23}} + \mathrm{e}^{\xi_2+\eta_2+\theta_{24}}, \qquad (2.14a)$$

$$g^{(3)} = \mathrm{e}^{\xi_1+\xi_2+\eta_1+\theta_{12}+\theta_{13}+\theta_{23}} + \mathrm{e}^{\xi_1+\xi_2+\eta_2+\theta_{12}+\theta_{14}+\theta_{24}}, \qquad (2.14b)$$

$$h^{(3)} = \mathrm{e}^{\xi_1+\eta_1+\eta_2+\theta_{13}+\theta_{14}+\theta_{34}} + \mathrm{e}^{\xi_2+\eta_1+\eta_2+\theta_{23}+\theta_{24}+\theta_{34}}, \qquad (2.14c)$$

$$f^{(4)} = \mathrm{e}^{\xi_1+\xi_2+\eta_1+\eta_2+\theta_{12}+\theta_{13}+\theta_{14}+\theta_{23}+\theta_{24}+\theta_{34}}, \qquad (2.14d)$$

where

$$\mathrm{e}^{\theta_{12}} = -(k_1 - k_2)^2,\ \mathrm{e}^{\theta_{34}} = -(l_1 - l_2)^2,$$

$$\mathrm{e}^{\theta_{i(j+2)}} = -\frac{1}{(k_i + l_j)^2},\ i, j = 1,2. \qquad (2.15)$$

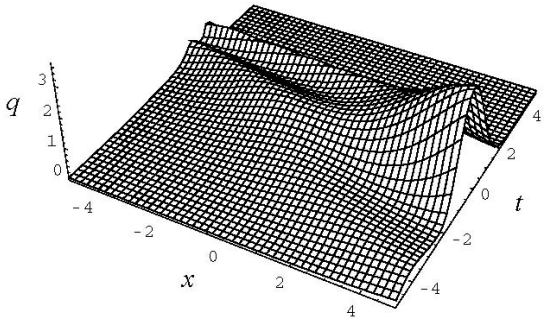


Fig. 3. Spatial structure of one-soliton solution q of (2.18).

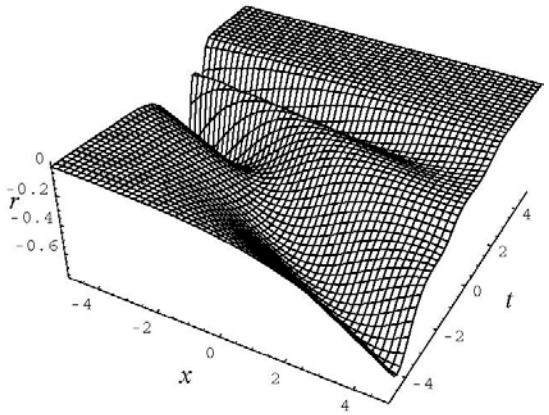


Fig. 4. Spatial structure of one-soliton solution r of (2.18).

Substituting (2.14) into (2.7c)-(2.7e) and those behind, we can verify that if

$$g^{(5)} = h^{(5)} = f^{(6)} = \cdots = 0, \qquad (2.16)$$

then (2.7c)-(2.7e) and those behind all hold. In this case, we write

$$f_2 = 1 + \mathrm{e}^{\xi_1+\eta_1+\theta_{13}} + \mathrm{e}^{\xi_1+\eta_2+\theta_{14}} + \mathrm{e}^{\xi_2+\eta_1+\theta_{23}} + \mathrm{e}^{\xi_2+\eta_2+\theta_{24}} + \mathrm{e}^{\xi_1+\xi_2+\eta_1+\eta_2+\theta_{12}+\theta_{13}+\theta_{14}+\theta_{23}+\theta_{24}+\theta_{34}}, \qquad (2.17a)$$

$$g_2 = \mathrm{e}^{\xi_1} + \mathrm{e}^{\xi_2} + \mathrm{e}^{\xi_1+\xi_2+\eta_1+\theta_{12}+\theta_{13}+\theta_{23}} + \mathrm{e}^{\xi_1+\xi_2+\eta_2+\theta_{12}+\theta_{14}+\theta_{24}}, \qquad (2.17b)$$

$$h_2 = \mathrm{e}^{\eta_1} + \mathrm{e}^{\eta_2} + \mathrm{e}^{\xi_1+\eta_1+\eta_2+\theta_{13}+\theta_{14}+\theta_{34}} + \mathrm{e}^{\xi_2+\eta_1+\eta_2+\theta_{23}+\theta_{24}+\theta_{34}}, \qquad (2.17c)$$

and hence determine the following two-soliton solutions of (1.1):

$$q = \frac{g_2}{f_2},\ r = \frac{h_2}{f_2}. \qquad (2.18)$$

In Figs. 3 and 4, the spatial structures of two-soliton solutions (2.18) are shown, where we select the parameters as $k_1 = 0.6$, $k_2 = 1.1$, $l_1 = 0.7$, $l_2 = 0.4$, $a_0(t) = 0.5\mathrm{sech}t$, $a_1(t) = \mathrm{e}^t$, $a_2(t) = \sin t$, $a_3(t) = 1 + 0.5t^2$, $\xi_1^0 = 0$, $\xi_2^0 = 0$, $\eta_1^0 = i\pi$, $\eta_2^0 = i\pi$.

Generally, if taking

$$g^{(1)} = \sum_{i=0}^{n} \mathrm{e}^{\xi_i},\ \xi_i = k_i x - \sum_{j=0}^{3} (-1)^j k_i^j \int a_j(t)\mathrm{d}t + \xi_i^0, \qquad (2.19a)$$

$$h^{(1)} = \sum_{i=0}^{n} \mathrm{e}^{\eta_i},\ \eta_i = l_i x + \sum_{j=0}^{3} l_i^j \int a_j(t)\mathrm{d}t + \eta_i^0, \qquad (2.19b)$$

then the solutions of (2.7a)-(2.7e) and those equations behind can be expressed by

$$f_n = \sum_{\mu=0,1} A_1(\mu) \mathrm{e}^{\sum_{i=1}^{2n} \mu_i \xi_i + \sum_{1 \le i < j}^{2n} \mu_i \mu_j \theta_{ij}}, \qquad (2.20a)$$

$$g_n = \sum_{\mu=0,1} A_2(\mu) \mathrm{e}^{\sum_{i=1}^{2n} \mu_i \xi_i + \sum_{1 \le i < j}^{2n} \mu_i \mu_j \theta_{ij}}, \qquad 2.20b)$$

$$h_n = \sum_{\mu=0,1} A_3(\mu) \mathrm{e}^{\sum_{i=1}^{2n} \mu_i \xi_i + \sum_{1 \le i < j}^{2n} \mu_i \mu_j \theta_{ij}}, \qquad (2.20c)$$

$$\xi_i = k_i x + \sum_{j=0}^{3} (-1)^j k_i^j \int a_j(t)\mathrm{d}t + \xi_i^0, \qquad (2.20d)$$

$$\eta_i = l_i x + \sum_{j=0}^{3} l_i^j \int a_j(t)\mathrm{d}t + \eta_i^0, \qquad (2.20e)$$

$$\mathrm{e}^{\theta_{ij}} = -(k_i - k_j)^2,\ \mathrm{e}^{\theta_{(i+n)(j+n)}} = -(l_i - l_j)^2, \qquad (2.20f)$$

$$\mathrm{e}^{\theta_{i(j+n)}} = -\frac{1}{(k_i + l_j)^2},\ i, j = 1,2,\cdots,n, \qquad (2.20g)$$

we can obtain the following uniform formula of the *n*-soliton solutions of (1.1):

$$q = \frac{g_n}{f_n},\ r = \frac{h_n}{f_n}. \qquad (2.21)$$

where the summation $\Sigma_{\mu=0,1}$ refers to all possible combinations of each $\mu_i = 0,1$ for $i = 1,2,\cdots,n$, $A_1(\mu)$, $A_2(\mu)$ and $A_3(\mu)$ denote that when we select all the possible combinations $\mu_j = 0,1$ $(j = 1,2,\cdots,2n)$ the following conditions hold, respectively

$$\sum_{j=1}^{n}\mu_j=\sum_{j=1}^{n}\mu_{n+j}\,,\ \sum_{j=1}^{n}\mu_j=\sum_{j=1}^{n}\mu_{n+j}+1,$$

$$\sum_{j=1}^{n}\mu_j+1=\sum_{j=1}^{n}\mu_{n+j}\,.$$

To the best of our knowledge, the obtained one-soliton solutions (2.12), two-soliton solutions (2.18), and *n*-soliton solutions (2.21) are new, they have not been reported in literature.

III. Conclusions

We have successfully obtained one-soliton solutions, two-soliton solutions and the uniform formulae of *n*-soliton solutions of a new AKNS equations with variable coefficients through Hirota's bilinear method. It is easy to see that the obtained one-soliton solutions (2.12), two-soliton solutions (2.18) and *n*-soliton solutions (2.21) include integrable functions $\alpha_1(t)$, $\alpha_2(t)$ and $\alpha_3(t)$, which provide enough freedom for us to describe enrich structures of these obtained soliton solutions. In the procedure of extending Hirota's bilinear method to the variable-coefficient AKNS equations (1.1), one of the key steps is to reduce (1.1) to the bilinear forms (2.3) and (2.4) by the transformation (2.1) introduced in this work. This paper shows that Hirota's bilinear method may provide us with an effective mathematical tool for constructing multi-solition solutions of some other nonlinear PDEs with variable coefficients. This is our task in future.

References

[1] C. S. Garder, J. M. Greene, M. D. Kruskal, and R. M. Miura, "Method for solving the Korteweg-de Vries equation," *Phys. Rev. Lett.* vol. 19, no. 12, pp. 1095–1097, Nov. 1967.

[2] R. Hirota, "Exact solution of the Korteweg-de Vries equation for multiple collisions of solitons," *Phys. Rev. Lett.* vol. 27, no. 18, pp. 1192–1194, Nov. 1971.

[3] M. R. Miurs, *Bäcklund Transformation*. Berlin, Springer, 1978.

[4] J. Weiss, M. Tabor, and G. Carnevale, "The Painlevé property for partial differential equations," *J. Math. Phys.* vol. 24, no. 3, pp. 522–526, Mar. 1983.

[5] M. L. Wang, "Solitary wave solutions for a variant Boussinesq equations," *Phys. Lett. A* vol. 199, no. 3-4, pp. 169–172, Mar. 1995.

[6] E. G. Fan, "Travelling wave solutions in terms of special functions for nonlinear coupled evolution systems," *Phys. Lett. A* vol. 300, no. 2-3, pp. 243–249, Jul. 2002.

[7] E. G. Fan and H. H. Dai, "A direct approach with computerized symbolic computation for finding a series of traveling waves to nonlinear equations," *Comput. Phys. Commun.* Vol. 153, no.1, pp. 17–30, Jun. 2003.

[8] J. H. He and X. H. Wu, "Exp-function method for nonlinear wave equations," Chaos Soliton. Fract. vol. 30, no. 3, pp. 700–708, Nov. 2006.

[9] S. Zhang and T. C. Xia, "A generalized auxiliary equation method and its application to (2+1)-dimensional asymmetric Nizhnik-Novikov -Vesselov equations," *J. Phys. A: Math. Theor.* vol. 40, no. 2, pp. 227–248, Jan. 2007.

[10] E. G. Fan, K. W. Chow, and J. H. Li, "On doubly periodic standing wave solutions of the coupled higgs field equation," *Stud. Appl. Math.* vol. 128, no. 1, pp. 86–105, Jan. 2012.

[11] W. X. Ma and J. H. Lee, "A transformed rational function method and exact solutions to 3+1 dimensional Jimbo-Miwa equation," Chaos Soliton. Fract. vol. 42, no. 3, pp. 1356–1363, Mar. 2009.

[12] Z. Y. Yan, "Localized analytical solutions and parameters analysis in the nonlinear dispersive Gross–Pitaevskii mean-field GP (m,n) model with space-modulated nonlinearity and potential," *Stud. Appl. Math.* vol. 132, no. 3, pp. 266–284, Apr. 2014.

[13] C. Q. Dai, Y. Y. Wang, Q. Tian, and J. F. Zhang, "The management and containment of self-similar rogue waves in the inhomogeneous nonlinear Schrödinger equation," *Ann. Phys.* vol. 327, no. 2, pp. 512–521, Feb. 2012.

[14] C. Q. Dai, X. G. Wang, and G. Q. Zhou, "Stable light-bullet solutions in the harmonic and parity-time-symmetric potentials," *Phys. Rev. A* vol. 89, no. 1, 013834(7pp.), Jan. 2014

[15] R. Hirota, *The Direct Method in Soliton Theory*. New York, Cambridge University Press, 2004.

[16] M. V. Balashov, "A property of the ansatz of Hirota's method for quasilinear parabolic equations," *Math. Notes* vol. 71, no. 3-4, pp. 339–354, Mar. 2002.

[17] R. Hirota, "Exact solution of the modified Korteweg-de Vries equation for multiple collisions of solitons," *J. Phys. Soc. Japan.* vol.33, no. 5, pp. 1456–1458, May 1972.

[18] I. Mcarthur and C. M. Yung, "Hirota bilinear form for the super-kdV hierarchy," *Mod. Phys. Lett. A* vol. 8, no. 18, pp. 1739–1745, Jun. 1993.

[19] Q. P. Liu, X. B. Hu, and M. X. Zhang, "Supersymmetric modified Korteweg-de Vries equation: Bilinear approach," *Nonlinearity* vol.18, no. 4, pp. 1597–1603, Jul. 2005.

[20] D. Y. Chen, *Introduction to Soliton.* Beijing, Science press, 2006.

[21] A. M. Wazwaz, "The Hirota's bilinear method and the tanh-coth method for multiple-soliton solutions of the Sawada-Kotera-Kadomtsev -Petviashvili equation," *Appl. Math. Comput.* vol. 200, no., pp. 160–166, 2008.

[22] S. Zhang and D. Liu, "Multisoliton solutions of a (2+1)-dimensional variable-coefficient Toda lattice equation via Hirota's bilinear method," *Can. J. Phys.* vol. 92, no. 3, pp. 184–190, Jan. 2014.

Generalizing Certain Properties of Decomposable Systems

Cristina Serbanescu, Ioan Bacalu

Abstract—In this paper we try to generalize certain results of the spectral theory for a single S-decomposable (S-spectral) operator to S-decomposable (S-spectral) systems of operators. We investigate the behaviour of S-decomposable (S-spectral) systems with respect to direct sums, by proving that the direct sum of two operator systems is an S-decomposable (S-spectral) system if and only if each system is S-decomposable (S-spectral); then one can prove several remarks concerning projections, separate parts of the Taylor spectrum, etc. Some of the previous general properties can be easily obtained as corollaries of certain theorems from homology and cohomology theory and from exterior product theory, but we also do some natural changes and basic calculations.

The spectral decompositions are related to differential equations and to systems of differential equations ([20]) and can have various applications in quantum mechanics, in bifurcation and fractal theories ([3]).

Keywords— S-decomposable (S-spectral) system; S-spectral capacity (S-spectral measure); exterior form; homology and cohomology modules.

I. Preliminaries

Let document E^n be the exterior algebra generated by n-tuple of indeterminates $\sigma=(s_1,s_2,...,s_n)$ over the field of complex numbers $\mathbb{C}$ ([21]). E^n is the complex algebra with identity e satisfying the relations $s_i\wedge s_j=-s_j\wedge s_i$, where by $(s_i,s_j)\to s_i\wedge s_j$ we denote multiplication in E^n. The algebra E^n is graded and $E^n=\sum_{p=0}^{\infty}\oplus E_p^n$, where E_p^n is generated by the elements of the form $s_{j_1}\wedge s_{j_2}\wedge...\wedge s_{j_p}$, with $1\le j_1\le j_2\le...\le j_p\le n$, for $p>0$, $E_p^n\wedge E_q^n\subset E_{p+q}^n$. We take $E_0^n\approx\mathbb{C}$, where the elements of E_0^n represent multiplies of the identity. Also, $E_n^n\approx\mathbb{C}$ has the single basis element $s_1\wedge s_2\wedge...\wedge s_n$ and $E_p^n=(0)$, for $p>n$.

Let X be a Banach space and let $a=(a_1,a_2,...,a_n)\subset\mathbf{B}(X)$ be a system of commuting operators. Let A be a complex algebra of operators whose centre containing the operators $a_1,a_2,...,a_n$. We denote by $\Lambda^p[\sigma,X]=E_p^n(X)=X\otimes_{\mathbb{C}}E_p^n$ ([16], [21]) the space of all exterior forms of degree p in s, having coefficients in X. The space $\Lambda^p[\sigma,X]$ can be viewed as a module over any operator algebra A, having the above property. By writing xs for $x\otimes s$, $x\in X$, $s\in E^n$, we note that $\Lambda^p[\sigma,X]$ is composed by elements written as:

$$\psi=\sum_{1\le j_1<j_2<...<j_p\le n}x_{j_1j_2...j_p}s_{j_1}\wedge s_{j_2}\wedge...\wedge s_{j_p},$$

$x_{j_1j_2...j_p}\in X$, for $p>0$.

We have the equalities $\Lambda^0[\sigma,X]=\Lambda^n[\sigma,X]=X$ and also we put $\Lambda^p[\sigma,X]=0$ for $p<0$ or $p>n$.

Through the spectrum of a system $a=(a_1,a_2,...,a_n)\subset\mathbf{B}(X)$ one comprehends, broadly speaking, the complement in $\mathbb{C}^n$ of the set of all $z=(z_1,z_2,...,z_n)\in\mathbb{C}^n$ having the property of nonsingularity for the system $z-a=(z_1-a_1,z_2-a_2,...,z_n-a_n)$. Using the sense given to the notion of nonsingularity, one can obtain several notions of the spectrum. We are interested in the Taylor spectrum ([21]), because it seems to have more advantages over the classical ones.

According to J.L. Taylor, the nonsingularity for the system $z-a$ means the exactness of a certain sequence

(determined by using the space and the operators). This sequence is a variant of the elementary sequence

$$0 \to X \xrightarrow{z-a} X \to 0$$

that – in the case of a single operator – shows the property of $z-a$ being both one to one and onto.

There are two types of sequences used to define the nonsingularity of a system of operators: the Koszul chain complex ([17]) or the cochain complex very much similar with a complex of differential forms. Both of them can be described in terms of exterior algebra; a natural duality between the two complexes makes them simultaneously exact and therefore it define the same notion of nonsingularity.

The common space basis of the two complexes is represented by the space $\Lambda^p[\sigma, X]$ and both complexes are different only through the link operators (boundary and coboundary operators, respectively).

If $1 \le p \le n$, we denote by

$$\delta_p = \delta_p(a) : \Lambda^p[\sigma, X] \to \Lambda^{p-1}[\sigma, X]$$

the operator defined by

$$\delta_p\left(xs_{j_1} \wedge ... \wedge s_{j_p}\right) = \sum_{i=1}^{p} (-1)^{i-1} a_i x s_{j_1} \wedge ... \wedge \hat{s}_{j_i} \wedge ... \wedge s_{j_p}$$

and

$$\delta_p\left(\sum_{1 \le j_1 < ... < j_p \le n} x_{j_1 ... j_p} s_{j_1} \wedge ... \wedge s_{j_p}\right) = \sum_{1 \le j_1 < ... < j_p \le n} \delta_p\left(x_{j_1 ... j_p} s_{j_1} \wedge ... \wedge s_{j_p}\right)$$

where the circumflex accent placed over an element marks its absence; for $p \le 0$ or $p > n$, we put $\delta_p = 0$.

We also denote by

$$\delta^p = \delta^p(a) : \Lambda^p[\sigma, X] \to \Lambda^{p+1}[\sigma, X]$$

the homomorphism that acts on an exterior form $\psi \in \Lambda^p[\sigma, X]$, defined by the left exterior multiplication of this form by $\alpha = a_1 s_1 + ... + a_n s_n$ (i.e. $\psi \to \alpha \wedge \psi$); for $p \le 0$ or $p > n$, we put $\delta^p = 0$). Using the commutativity of the system $a = (a_1, a_2, ..., a_n)$, the relations $\delta_p \delta_{p+1} = 0$ and $\delta^{p+1} \delta^p = 0,\ p \in \mathbb{Z}$ are verified. The chain complex composed by the modules $\Lambda^p[\sigma, X]$ and the boundary operators $\delta_p,\ p \in \mathbb{Z}$ is called the Koszul complex associated with the system $a = (a_1, a_2, ..., a_n) \subset \mathbf{B}(X)$ and it is denoted by $E(X, a)$. The chain complex represented by the modules $\Lambda^p[\sigma, X]$ and the coboundary operators $\delta^p,\ p \in \mathbb{Z}$ is denoted by $F(X, a)$. Hence we have

$$E(X,a): 0 \to X = \Lambda^n[\sigma, X] \xrightarrow{\delta_n} \Lambda^{n-1}[\sigma, X] \xrightarrow{\delta_{n-1}} ... \xrightarrow{\delta_3} \Lambda^2[\sigma, X] \xrightarrow{\delta_2} \Lambda^1[\sigma, X] \xrightarrow{\delta_1} \Lambda^0[\sigma, X] = X \to 0$$

and

$$F(X,a): 0 \to X = \Lambda^0[\sigma, X] \xrightarrow{\delta^0} \Lambda^1[\sigma, X] \xrightarrow{\delta^1} \Lambda^2[\sigma, X] \xrightarrow{\delta^2} ... \xrightarrow{\delta^{n-2}} \Lambda^{n-1}[\sigma, X] \xrightarrow{\delta^{n-1}} \Lambda^n[\sigma, X] = X \to 0$$

.

Broadly speaking, the above sequences are not exact. The homology modules of the Koszul complex $E(X, a)$ are the sequences of quotients A-modules:

$$H_p(X,a) = \operatorname{Ker}\left(\delta_{p+1} : \Lambda^{p+1} \to \Lambda^p\right) / \operatorname{Im}\left(\delta_p : \Lambda^p \to \Lambda^{p-1}\right),\ p \in \mathbb{Z}$$

and the cohomology modules of chain complex $F(X, a)$ are denoted by

$$H^p(X,a) = \operatorname{Ker}\left(\delta^p : \Lambda^p \to \Lambda^{p+1}\right) / \operatorname{Im}\left(\delta^{p-1} : \Lambda^{p-1} \to \Lambda^p\right),\ p \in \mathbb{Z}$$

.

One can easily verify that the two complexes $E(X, a)$ and $F(X, a)$ are equivalent with respect to the notion of exactness ([16]).

Definition 1.1. ([16]) The system $a = (a_1, a_2, ..., a_n) \subset \mathbf{B}(X)$ is said to be *nonsingular* on X if the Koszul complex $E(X, a)$ is exact or, equivalently, the complex $F(X, a)$ is exact. The set of those $z = (z_1, z_2, ..., z_n) \in \mathbb{C}^n$ for which the system $z - a = (z_1 - a_1, ..., z_n - a_n)$ is nonsingular on X is called the *resolvent set of* a *on* X and is denoted by $r(a, X)$. The complement in $\mathbb{C}^n$ of this set,

$\mathbb{C}^n \setminus r(a, X)$, is said to be *the spectrum of* a *on* X and is denoted by $\sigma(a, X)$.

We shall use the following spaces of X-valued functions defined on an open set $U \subset \mathbb{C}^n$: $\mathbf{B}(U, X)$ – the space of all continuous functions admitting (in the distribution sense) continuous partial derivatives with respect to $\overline{z}_1, \overline{z}_2, ..., \overline{z}_n$ ([17]); $\mathbf{B}_0(U, X)$ – the subspace of $\mathbf{B}(U, X)$ consisting of all functions with compact support; $C^\infty(U, X)$ – the space of all continuous functions admitting partial derivatives of any order; $C_0^\infty(U, X)$ – the subspace of $C^\infty(U, X)$ consisting of all functions with compact support; $\mathbf{U}(U, X)$ – the space of analytic functions on U. In addition, we permanently use the fact that $\mathbf{B}(U, X) = C^\infty(U, X)$ ([24]).

If $U \subset \mathbb{C}^n$ is open, $\mathbf{F}$ is one of the function spaces described above and $\sigma = (s_1, s_2, ..., s_n)$ is a system of indeterminates, then we denote by α the operator that acts on an exterior form $\psi \in \Lambda^p[\sigma, \mathbf{F}]$ in indeterminates $\sigma = (s_1, s_2, ..., s_n)$ with coefficients in $\mathbf{F}$ ([22]), in the following way:

$$(\alpha\psi)(z) = \left[(z_1 - a_1)s_1 + (z_2 - a_2)s_2 + ... + (z_n - a_n)\right] \wedge \psi(z),\ z \in U$$

and we also denote by $\alpha \oplus \overline{\partial}$ the operator that acts similarly on the exterior forms $\psi \in \Lambda^p[\sigma \cup d\overline{z}, \mathbf{F}]$ in indeterminates σ and $d\overline{z} = (d\overline{z}_1, d\overline{z}_2, ..., d\overline{z}_n)$ with coefficients in $\mathbf{F}$ ([22]):

$$\left((\alpha \oplus \overline{\partial})\psi\right)(z) = \left[(z_1 - a_1)s_1 + (z_2 - a_2)s_2 + ... + (z_n - a_n)s_n + \frac{\partial}{\partial \overline{z}_1} d\overline{z}_1 + ... + \frac{\partial}{\partial \overline{z}_n} d\overline{z}_n\right] \wedge \psi(z).$$

Definition 1.2. ([16]) *The analytic resolvent set* of x *with respect to* $a = (a_1, a_2, ..., a_n) \subset \mathbf{B}(X)$ is the set of all $z = (z_1, z_2, ..., z_n) \in \mathbb{C}^n$ such that there are an open neighborhood V of z and n X-valued analytic functions $f_1, f_2, ..., f_n$ on V, satisfying the identity

$$x = (\zeta_1 - a_1) f_1(\zeta) + ... + (\zeta_n - a_n) f_n(\zeta),\ \zeta \in V.$$

The complement of this set in $\mathbb{C}^n$ is said to be *the analytic spectrum of* x *with respect to* a. We shall denote them by $\rho(a, x)$, respectively $\sigma(a, x)$.

Definition 1.3. ([16]) *The resolvent set* of x *with respect to* $a = (a_1, a_2, ..., a_n) \subset \mathbf{B}(X)$, denoted by $r(a, x)$, is the union of all open sets V in $\mathbb{C}^n$ having the property that there is an exterior form $\psi \in \Lambda^{n-1}\left[\sigma \cup d\overline{z}, C^\infty(V, X)\right]$ satisfying the equality

$$x(s_1 \wedge s_2 \wedge ... \wedge s_n) = \left[(z_1 - a_1)s_1 + ... + (z_n - a_n)s_n + \frac{\partial}{\partial \overline{z}_1} d\overline{z}_1 + ... + \frac{\partial}{\partial \overline{z}_n} d\overline{z}_n\right] \wedge \psi(z)$$

.

The complement in $\mathbb{C}^n$ of this set is called *the spectrum* of x *with respect to* a and is denoted by $sp(a, x) = \mathbb{C}^n \setminus r(a, x)$.

In order to obtain a global solution ψ for the equation $sx = (\alpha \oplus \overline{\partial})\psi$, it is necessary that the system $a = (a_1, a_2, ..., a_n) \subset \mathbf{B}(X)$ to satisfy a similar condition to the single-valued extension property from the case of a single operator. This condition is expressed by the following cohomology property:

Definition 1.4. ([16]) We say that the system $a = (a_1, a_2, ..., a_n) \subset \mathbf{B}(X)$ verifies *the cohomology property* (L) or a has *the single-valued extension property* if

$$H^{n-1}\left(C^\infty(G, X), \alpha \oplus \overline{\partial}\right) = 0$$

for any open set $G \subset \mathbb{C}^n$.

For the commuting operator system $a = (a_1, a_2, ..., a_n) \subset \mathbf{B}(X)$, let S_a be the compact minimal set with the property that $H^{n-1}\left(C^\infty(G, X), \alpha \oplus \overline{\partial}\right) = 0$, for any open set $G \subset \mathbb{C}^n$ such that $G \cap S_a = \varnothing$ (minimal means that S_a is the intersection of all compact sets having the property above).

S_a is called *the spectral analytic residuum* of the system a ([25]).

Obviously, $S_a \subset \sigma(a, X)$ and $H^i\left(C^\infty\left(\mathbb{C}^n \setminus S_a, X\right), \alpha \oplus \overline{\partial}\right) = 0,\ 0 \le i \le n-1;$ if $S_a = \varnothing$, then a verifies condition (L) (i.e. a has the single-valued extension property) ([25]).

For the system a with the spectral analytic residuum S_a, if H is an arbitrary subset of $\mathbb{C}^n$ such that $H \supset S_a$, we denote

$$X_{[a]}(H)=\{x;\ x\in X,\ sp(a,x)\subset H\} \text{ and}$$

$$X_a(H)=\{x;\ x\in X,\ \sigma(a,x)\subset H\};$$

$X_{[a]}(H)$ and $X_a(H)$ are linear subspaces of X and $X_a(H)\subset X_{[a]}(H)$.

II. Procedure for direct sums of exterior forms. Direct sums of homology and cohomology modules.

In this chapter we remind some results published in [26]

Lemma 2.1. *If $\Lambda^p[\sigma,X]$ and $\Lambda^p[\sigma,Y]$ are the spaces of all exterior forms of grade p in s ($\sigma=(s_1,s_2,...,s_n)$) with coefficients in X, respectively Y, then*

$$\Lambda^p[\sigma,X]\oplus\Lambda^p[\sigma,Y]=\Lambda^p[\sigma,X\oplus Y].$$

Remark 2.2. If in the previous lemma we replace the system of indeterminates σ with the system $\sigma\cup d\overline{z}=(s_1,s_2,...,s_n,d\overline{z}_1,d\overline{z}_2,...,d\overline{z}_n)$ and the spaces X and Y with $C^\infty(G,X)$, respectively $C^\infty(G,Y)$ $\left(G\subset\mathbb{C}^n \text{ open set}\right)$, using moreover the obvious equality $C^\infty(G,X)\oplus C^\infty(G,Y)=$ $=C^\infty(G,X\oplus Y)$, we obtain

$$\Lambda^p\left[\sigma\cup d\overline{z},C^\infty(G,X)\right]\oplus\Lambda^p\left[\sigma\cup d\overline{z},C^\infty(G,Y)\right]$$
$$=\Lambda^p\left[\sigma\cup d\overline{z},C^\infty(G,X\oplus Y)\right]$$

.

Lemma 2.3. *Let A, A', B, B' be modules over an algebra such that $A'\subset A, B'\subset B$ and let h,k be two arbitrary maps between arbitrary given sets. Then we have*

$$A/A'\oplus B/B' = A\oplus B/A'\oplus B',$$
$$\operatorname{Ker}h\oplus\operatorname{Ker}k = \operatorname{Ker}(h\oplus k),$$
$$\operatorname{Im}h\oplus\operatorname{Im}k = \operatorname{Im}(h\oplus k).$$

In the same manner, for a number $n>2$ of modules and applications we have:

$$A_1/A_1'\oplus A_2/A_2'\oplus...\oplus A_n/A_n'$$
$$=(A_1\oplus A_2\oplus...\oplus A_n)/A_1'\oplus A_2'\oplus...\oplus A_n',$$
$$\operatorname{Ker}h_1\oplus\operatorname{Ker}h_2\oplus...\oplus\operatorname{Ker}h_n$$
$$=\operatorname{Ker}(h_1\oplus h_2\oplus...\oplus h_n),$$
$$\operatorname{Im}h_1\oplus\operatorname{Im}h_2\oplus...\oplus\operatorname{Im}h_n$$
$$=\operatorname{Im}(h_1\oplus h_2\oplus...\oplus h_n).$$

Proof. One can easily prove by direct verification, and for $n>2$, eventually, by mathematical induction.

Proposition 2.4. *If $a=(a_1,a_2,...,a_n)\subset\mathbf{B}(X)$, $b=(b_1,b_2,...,b_n)\subset\mathbf{B}(Y)$ are two operator systems and H^p are the homology modules, then we have*

$$H^p(X,z-a)\oplus H^p(Y,z-b)=H^p\left(X\oplus Y,z-(a\oplus b)\right),$$
$$H^p\left(C^\infty(G,X),\alpha\oplus\overline{\partial}\right)\oplus H^p\left(C^\infty(G,Y),\beta\oplus\overline{\partial}\right)=$$
$$=H^p\left(C^\infty(G,X\oplus Y),(\alpha\oplus\beta)\oplus(\overline{\partial}\oplus\overline{\partial})\right)$$

for any $z\in\mathbb{C}^n$ and $G\subset\mathbb{C}^n$ open set.

Lemma 2.5. *Let $a=(a_1,a_2,...,a_n)\subset\mathbf{B}(X)$, $b=(b_1,b_2,...,b_n)\subset\mathbf{B}(Y)$ be two commuting systems of operators. Then the system $a\oplus b=(a_1\oplus b_1,a_2\oplus b_2,...,a_n\oplus b_n)\subset\mathbf{B}(X\oplus Y)$ has the property that the corresponding Taylor spectrums verifies the equality*

$$\sigma(a\oplus b,X\oplus Y)=\sigma(a,X)\bigcup\sigma(b,Y).$$

Proposition 2.6. *Let $a=(a_1,a_2,...,a_n)\subset\mathbf{B}(X)$, $b=(b_1,b_2,...,b_n)\subset\mathbf{B}(Y)$. The systems a and b verify condition (L) if and only if the system $a\oplus b=(a_1\oplus b_1,a_2\oplus b_2,...,a_n\oplus b_n)\subset\mathbf{B}(X\oplus Y)$ verifies condition (L).*

Proposition 2.7. *If $a=(a_1,a_2,...,a_n)\subset\mathbf{B}(X)$ and $b=(b_1,b_2,...,b_n)\subset\mathbf{B}(Y)$ are two systems of operators that verify condition (L), then we have the equalities*:

1. $\rho(a\oplus b,x\oplus y)=\rho(a,x)\bigcap\rho(b,y)$;

2. $\sigma(a\oplus b,x\oplus y)=\sigma(a,x)\bigcup\sigma(b,y)$;
3. $r(a\oplus b,x\oplus y)=r(a,x)\bigcap r(b,y)$;
4. $sp(a\oplus b,x\oplus y)=sp(a,x)\bigcup sp(b,y)$;
5. $(X\oplus Y)_{[a\oplus b]}(F)=X_{[a]}(F)\oplus Y_{[b]}(F)$;
6. $(X\oplus Y)_{a\oplus b}(F)=X_a(F)\oplus Y_b(F)$

where $x\in X$, $y\in Y$ *and* $F\subset\mathbb{C}^n$ closed.

Moreover, for n *systems of operators* ($n>2$) *we have*:

1. $\rho(a_1\oplus a_2\oplus...\oplus a_n,x_1\oplus x_2\oplus...\oplus x_n)=\rho(a_1,x_1)\bigcap\rho(a_2,x_2)\bigcap...\bigcap\rho(a_n,x_n)$;
2. $\sigma(a_1\oplus a_2\oplus...\oplus a_n,x_1\oplus x_2\oplus...\oplus x_n)=\sigma(a_1,x_1)\bigcup\sigma(a_2,x_2)\bigcup...\bigcup\sigma(a_n,x_n)$;
3. $r(a_1\oplus a_2\oplus...\oplus a_n,x_1\oplus x_2\oplus...\oplus x_n)=r(a_1,x_1)\bigcap r(a_2,x_2)\bigcap...\bigcap r(a_n,x_n)$;
4. $sp(a_1\oplus a_2\oplus...\oplus a_n,x_1\oplus x_2\oplus...\oplus x_n)=sp(a_1,x_1)\bigcup sp(a_2,x_2)\bigcup...\bigcup sp(a_n,x_n)$;
5. $(X_1\oplus X_2\oplus...\oplus X_n)_{[a_1\oplus a_2\oplus...\oplus a_n]}(F)=X_{1[a_1]}(F)\oplus X_{2[a_2]}(F)\oplus...\oplus X_{n[a_n]}(F)$
6. $(X_1\oplus X_2\oplus...\oplus X_n)_{a_1\oplus a_2\oplus...\oplus a_n}(F)=X_{1a_1}(F)\oplus X_{2a_2}(F)\oplus...\oplus X_{na_n}(F)$.

Proposition 2.8. *Let* $a=(a_1,a_2,...,a_n)\subset\mathbf{B}(X)$ *be a commuting system of operators verifying condition* (L). *Then* $X_{[a]}(F)$ *is a linear manifold ultrainvariant to* a, *in other words it is invariant to all operators* $b\in\mathbf{B}(X)$ *commuting with every* a_i $(i=1,2,...,n)$. *If* $X_{[a]}(F)$ *is closed, for* $F\subset\mathbb{C}^n$ *closed and* $\sigma\left(a,X_{[a]}(F)\right)\subset F$, *then* $X_{[a]}(F)$ *is a spectral maximal space of* a, *i.e. for any subspace* Y *invariant to* a *with* $\sigma(a,Y)\subset F$ *we have* $Y\subset X_{[a]}(F)$.

III. Several properties of S-decomposable systems

Definition 3.1. For the Banach space X, let $\mathcal{S}(X)$ be the family of all linear closed subspaces of X, let $S\subset\mathbb{C}^n$ be a compact set and let $\mathcal{F}_S^n$ be the family of all closed subsets $F\subset\mathbb{C}^n$ which have the property: either $F\bigcap S=\varnothing$ or $F\supset S$.

The application $\mathcal{E}_S:\mathcal{F}_S^n\to\mathcal{S}(X)$ is called S-*spectral capacity* if it verifies the conditions:

(1) $\mathcal{E}_S(\varnothing)=\{0\}$, $\mathcal{E}_S\left(\mathbb{C}^n\right)=X$;

(2) $\mathcal{E}_S\left(\bigcap_{i=1}^{\infty}F_i\right)=\bigcap_{i=1}^{\infty}\mathcal{E}_S(F_i)$, for any series $\{F_i\}_{i\in\mathbb{N}}\subset\mathcal{F}_S^n$;

(3) for any open finite S-covering $\{G_S\}\bigcup\{G_j\}_{j=1}^m$ of $\mathbb{C}^n$ we have

$$X=\mathcal{E}_S(\bar{G}_S)+\sum_{j=1}^{m}\mathcal{E}_S(\bar{G}_j).$$

A commuting operator system $a=(a_1,a_2,...,a_n)\subset\mathbf{B}(X)$ is said to be S-*decomposable* if there is an S-spectral capacity $\mathcal{E}_S$ such that

(4) $a_j\mathcal{E}_S(F)\subset\mathcal{E}_S(F)$, for any $F\in\mathcal{F}_S^n$ and for any $1\le j\le n$;

(5) $\sigma\left(a,\mathcal{E}_S(F)\right)\subset F$, for any $F\in\mathcal{F}_S^n$.

In case $S=\varnothing$, $\mathcal{F}_{\varnothing}^n=\mathcal{F}\left(\mathbb{C}^n\right)$ is the family of all closed subsets of $\mathbb{C}^n$, the S-spectral capacity is said to be spectral capacity and the system is decomposable ([16]). We must notice that for operator systems having $n\ge2$, we do not know whether the definition of S-decomposability (and of the decomposability) given for an operator ([10]) is equivalent with Definition 3.1 or not.

Definition 3.2. Let $a=(a_1,a_2,...,a_n)\subset\mathbf{B}(X)$ be a commuting operator system and let $S\subset\mathbb{C}^n$ be a fixed compact subset. We say that a verifies *the cohomology property* (L_S) if

$$H^{n-1}\left(C^{\infty}(G,X),\alpha\oplus\bar{\partial}\right)=0$$

for any open set $G\subset\mathbb{C}^n$, with $G\bigcap S=\varnothing$.

Theorem 3.3. *Let* $a=(a_1,a_2,...,a_n)\subset \mathbf{B}(X)$ *be an* S*-decomposable system and let* $\mathcal{E}_S$ *be an* S*-spectral capacity for a. Then* $\mathcal{E}_S(F)$ *is a spectral maximal space of* a*, where* $F\subset\mathbb{C}^n$ *is closed.*

Proof. Let Y be a closed subspace of X invariant to a with $\sigma(a,Y)\subset F$, for a certain closed set $F\subset\mathbb{C}^n$. Let us choose $F\supset S$. Therefore there is an open S-covering $\{G_S, G\}$ of $\mathbb{C}^n$ such that $G_S\supset F$ and $\bar{G}\bigcap F=\varnothing$ and

$$X=\mathcal{E}_S(\bar{G}_S)+\mathcal{E}_S(\bar{G}).$$

From an isomorphism theorem, the quotient space $X/\mathcal{E}_S(\bar{G}_S)$ is isomorphic to

$$\mathcal{E}_S(\bar{G})/\mathcal{E}_S(\bar{G}_S)\bigcap\mathcal{E}_S(\bar{G})=\mathcal{E}_S(\bar{G})/\mathcal{E}_S(\bar{G}_S\bigcap\bar{G})$$

.

According to Taylor's Theorem concerning the inclusion of the spectra ([21]) we obtain

$\sigma(a,\mathcal{E}_S(\bar{G})/\mathcal{E}_S(\bar{G}_S\cap\bar{G}))\subset\sigma(a,\mathcal{E}_S(\bar{G}_S\cap\bar{G}))\cup\sigma(a,\mathcal{E}_S(\bar{G}))\subset(\bar{G}_S\cap\bar{G})\cup\bar{G}=\bar{G}$

,

meaning

$$\sigma(a,X/\mathcal{E}_S(\bar{G}_S))\subset\bar{G}.$$

Let us make the following notations: φ is the canonical map of X on $Z=X/\mathcal{E}_S(\bar{G}_S)$, b_i is the restriction of a_i to Y, c_i is the operator induced by a_i in $Z=X/\mathcal{E}_S(\bar{G}_S)$ and τ is the restriction of φ to Y. Putting $b=(b_1,b_2,...,b_n)$, $c=(c_1,c_2,...,c_n)$ we have found

$$\sigma(b,Y)\bigcap\sigma(c,Z)\subset F\bigcap G=\varnothing.$$

If f is the germ of the analytic function which is equal to 1 on $\sigma(b,Y)$, respectively to 0 on $\sigma(c,Y)$, then $f(b)=I_Y$ and $f(c)=0$. According to Proposition 3.2.1, [8], one can obtain $\varphi\cdot I_Y=0$, therefore $Y\subset\mathcal{E}_S(\bar{G}_S)$. Since G_S is arbitrary under the property $G_S\supset F$, we can deduce that

$$Y\subset\bigcap\{\mathcal{E}_S(\bar{G}_S);\ G_S\supset F\}=\mathcal{E}_S(F).$$

The case $F\in\mathcal{F}_S^n$, $F\bigcap S=\varnothing$ can be obtained in a similar way.

Corollary 3.4. *If* $a=(a_1,a_2,...,a_n)\subset\mathbf{B}(X)$ *is an* S*-decomposable system, then* a *admits a unique* S*-spectral capacity* $\mathcal{E}_S$.

Proof. Let $\mathcal{E}_{1S}$ and $\mathcal{E}_{2S}$ be two S-spectral capacities for a; then, according to the previous theorem, $\mathcal{E}_{1S}(F)$ and $\mathcal{E}_{2S}(F)$ are spectral maximal spaces of a and from the inclusions

$$\sigma(a,\mathcal{E}_{1S}(F))\subset F,\ \sigma(a,\mathcal{E}_{2S}(F))\subset F$$

it follows that

$$\mathcal{E}_{1S}(F)\subset\mathcal{E}_{2S}(F),\ \mathcal{E}_{2S}(F)\subset\mathcal{E}_{1S}(F),$$

consequently the two S-spectral capacities coincide.

Theorem 3.5. *If* $a=(a_1,a_2,...,a_n)\subset\mathbf{B}(X)$ *is* S*-decomposable, then* $S_a\subset S$.

Proof. With minor modifications, the proof is similar with the proof for the case of decomposable systems (when $S=\varnothing$) (Proposition 2.1.4, [16]).

We intend to show that for any open polydisc $U\subset\mathbb{C}^n$ such that $U\bigcap S=\varnothing$ we have

$$H^i(\mathbf{U}(U,X),\alpha)=0\ \ (0\le i\le n-1).$$

One can contend this by mathematical induction on i. Let us begin with the initial step $i=0$. Let $f\in\mathbf{U}(U,X)$ such that $\alpha f=0$; according to Proposition 3.5.8, [8], we have $f=0$ on any polydisc D' with $\overline{D}'\subset U$ and hence $f=0$ on U. We assume that for any open polydisc $D\subset\mathbb{C}^n$ with $D\bigcap S=\varnothing$ we have

$$H^{i-1}(\mathbf{U}(D,X),\alpha)=0$$

for fixed i, $0<i\le n-1$ and we must prove that $H^i(\mathbf{U}(U,X),\alpha)=0$.

Let $\{D_\nu\}$ be a sequence of open polydiscs, $D_\nu\bigcap S=\varnothing$ such that $\overline{D}_\nu\subset D_{\nu+1}$, for any ν with $\bigcup_{\nu=1}^{\infty}D_\nu=U$ and let $\psi\in\Lambda^i[\sigma,\mathbf{U}(U,X)]$ such that $\alpha\psi=0$. Applying Proposition 3.5.8, [8] for D_2, we deduce that there is an exterior form $\varphi_1\in\Lambda^{i-1}[\sigma,\mathbf{U}(D_2,X)]$ such that $\psi=\alpha\varphi_1$ on D_2; analogously, we may find a form φ_2' on D_3 with

$\psi = \alpha\,\varphi_2'$ on D_3. Consequently, one can obtain $\alpha\left(\varphi_1 - \varphi_2'\right) = 0$ on D_2, whence, by applying the inductive hypothesis, we can deduce there is a form $\chi \in \Lambda^{i-2}\left[\sigma, \mathbf{U}\left(D_2, X\right)\right]$ such that $\varphi_1 - \varphi_2' = \alpha\,\chi$. We preserve a sufficient number of terms from the Taylor decomposition of χ on D_2 such that χ' (the preserved part) verifies the inequality $\left\|\alpha\,\chi - \alpha\,\chi'\right\| < \frac{1}{2}$ on $\bar{D}_1$. If we replace φ_2' with $\varphi_2 = \varphi_2' + \alpha\,\chi'$, we find a form on D_3 such that $\psi = \alpha\,\varphi_2$ on D_3 and

$$\left\|\varphi_1 - \varphi_2\right\| = \left\|\varphi_1 - \varphi_2' - \alpha\,\chi'\right\| = \left\|\alpha\,\chi - \alpha\,\chi'\right\| < \frac{1}{2} \text{ on } \bar{D}_1.$$

Analogously, we can define a sequence of forms $\left\{\varphi_\nu\right\}$, $\varphi_\nu \in \Lambda^{i-1}\left[\sigma, \mathbf{U}\left(D_{\nu+1}, X\right)\right]$ under the conditions: $\psi = \alpha\,\varphi_\nu$ on $D_{\nu+1}$ and $\left\|\varphi_{\nu+1} - \varphi_\nu\right\| < \frac{1}{2^{\nu+1}}$ on $\bar{D}_\nu$.

The sequence $\left\{\varphi_\nu\right\}$ obviously converges to a form $\varphi \in \Lambda^{i-1}\left[\sigma, \mathbf{U}\left(U, X\right)\right]$ having analytic coefficients on U and satisfying the equation $\psi = \alpha\,\varphi$ on U, hence the inductive proof is ended.

Finally, we observe that $H^i\left(\mathbf{U}\left(U, X\right), \alpha\right) = 0$ $\left(0 \le i \le n-1\right)$ implies

$$H^i\left(C^\infty\left(G, X\right), \alpha \oplus \bar{\partial}\right) = 0 \ \left(0 \le i \le n-1\right)$$

where U is any open polydisc of $\mathbb{C}^n$ and G is any open subset of $\mathbb{C}^n$ such that $U \cap S = \varnothing$ and $G \cap S = \varnothing$; the proof of this observation is given in [16], Theorem 1.5.16, for any $U, G \subset \mathbb{C}^n$. Applying the definition of S_a, it results that $S_a \subset S$.

Remark 3.6. If $a = \left(a_1, a_2, ..., a_n\right) \subset \mathbf{B}\left(X\right)$ is an S-decomposable system, on account of the proof of Theorem 3.5 we have

$$H^{n-1}\left(C^\infty\left(G, X\right), \alpha \oplus \bar{\partial}\right) = 0$$

for any open set $G \subset \mathbb{C}^n$ with $G \cap S = \varnothing$, hence a verifies the cohomology property $\left(L_S\right)$.

Remark 3.7. If $a = \left(a_1, a_2, ..., a_n\right) \subset \mathbf{B}\left(X\right)$ is an S-decomposable system and $\mathcal{E}_S$ is its S-spectral capacity, then we have the following properties:

1) $\mathcal{E}_S\left(F_1 \cup F_2\right) = \mathcal{E}_S\left(F_1\right) \oplus \mathcal{E}_S\left(F_2\right)$, for $F_1,\ F_2 \in \mathcal{F}_S^n$, $F_1 \cap F_2 = \varnothing$.

2) $\mathcal{E}_S\left(F\right)$ is a spectral maximal space of a, for any $F \in \mathcal{F}_S^n$.

3) If $S_a = \varnothing$, then $\mathcal{E}_S\left(F\right) = X_a\left(F\right)$, for any $F \in \mathcal{F}_S^n$ and $\sigma\left(a, X_a\left(F\right)\right) \subset F$.

4) If $S_a \neq \varnothing$, then

$\mathcal{E}_S\left(F\right) = X_a\left(F\right)$, for any $F \in \mathcal{F}_S^n$, $F \supset S$ and $\sigma\left(a, X_a\left(F\right)\right) \subset F$ and

$\mathcal{E}_S\left(F\right) = Y_F$, for any $F \in \mathcal{F}_S^n$, $F \cap S = \varnothing$, where Y_F is the spectral maximal space of a defined by the equality

$$\mathcal{E}_S\left(F \cup S\right) = \mathcal{E}_S\left(F\right) \oplus \mathcal{E}_S\left(S\right) = X_a\left(F \cup S\right) = Y_F \oplus X_a\left(S\right)$$

and

$\sigma\left(a, Y_F\right) \subset F$.

Lemma 3.8. *If the system* $a = \left(a_1, a_2, ..., a_n\right) \subset \mathbf{B}\left(X\right)$ *is decomposable or* S*-decomposable and* $X_a\left(F\right)$ *are the spectral subspaces associated with* a*, then*

$$X_a\left(F_1 \cap F_2\right) = X_a\left(F_1\right) \cap X_a\left(F_2\right)$$

for $F_1, F_2 \subset \mathbb{C}^n$ *closed when* a *is decomposable, for* $F_1, F_2 \in \mathcal{F}_S^n$ *when* a *is* S*-decomposable with* $S_a = \varnothing$*, for* $F_1, F_2 \in \mathcal{F}_S^n$, $F_1 \supset S$, $F_2 \supset S$, *when* a *is* S*-decomposable with* $S_a \neq \varnothing$.

Moreover, for $F_1, F_2, ..., F_i, ...$ *closed subsets of* $\mathbb{C}^n$ *or* $F_i \in \mathcal{F}_S^n$, $i \in I$ *we have*

$$X_a\left(\bigcap_{i \in I} F_i\right) = \bigcap_{i \in I} X_a\left(F_i\right).$$

Proof. Let us suppose that a is decomposable and let $x \in X_a\left(F_1 \cap F_2\right)$; therefore $\sigma\left(a, x\right) \subset \ \subset F_1 \cap F_2$, hence $\sigma\left(a, x\right) \subset F_1$, $\sigma\left(a, x\right) \subset F_2$, consequently

$x \in X_a(F_1)$, $x \in X_a(F_2)$, whence $x \in X_a(F_1) \bigcap X_a(F_2)$. Conversely, let $x \in X_a(F_1) \bigcap X_a(F_2)$; then we have $\sigma(a,x) \subset F_1$, $\sigma(a,x) \subset F_2$, hence $\sigma(a,x) \subset F_1 \bigcap F_2$, from where $x \in X_a(F_1 \bigcap F_2)$.

The second equality is proof in the same manner as above or by mathematical induction using the first equality.

Proposition 3.9. (Criterion of S-decomposability for operator systems)

A commuting operator system $a = (a_1, a_2, ..., a_n) \subset \mathbf{B}(X)$ *is* S*-decomposable if and only if the following conditions are verified*:

(I) a *verifies condition* (L_S), *the space* $X_a(F)$ *is closed and* $\sigma(a, X_a(F)) \subset F$, *for any* $F \in \mathcal{F}_S^n$, $F \supset S$;

(II) *for any open* S*-covering* $\{G_S\} \bigcup \{G_j\}_{j=1}^m$ *of* $\mathbb{C}^n$ *and for any* $x \in X$ *we have*

$x = x_S + x_1 + x_2 + ... + x_m$, *with* $\sigma(a, x_S) \subset G_S$, $\sigma(a, x_j) \subset G_j$, $j = 1, 2, ..., m$.

Proof. Let us suppose that a is S-decomposable and let $\mathcal{E}_S$ be its S-spectral capacity. According to Theorem 3.5, $S_a \subset S$ and from Remark 3.6, a verifies condition (L_S) (the system a verifies condition (L_S) means that a verifies condition (L) on $\mathbb{C}^n \setminus S$, in other words the system a has the single-valued extension property (modulo S) on $\mathbb{C}^n \setminus S$).

Let $F \in \mathcal{F}_S^n$, $F \supset S$; then $S_a \subset F$ and $X_a(F)$ makes sense. According to Theorem 3.3 and Remark 3.7, $\mathcal{E}_S(F)$, for $F \in \mathcal{F}_S^n$, is spectral maximal space of a, $\mathcal{E}_S(F) = X_a(F)$, for $F \in \mathcal{F}_S^n$, $F \supset S$ and $\sigma(a, \mathcal{E}_S(F)) \subset F$. Therefore the space $X_a(F)$ is a spectral maximal space of a, hence it is closed.

The second assertion of hypothesis is obviously verified, from condition (3) of the definition of S-decomposability.

Conversely, let us suppose that the conditions (I) and (II) are satisfied. Let us first consider the case when the system a has the single-valued extension property, i.e. $S_a = \varnothing$.

When $S_a = \varnothing$, it is known that $X_a(F)$ makes sense, for any $F \in \mathcal{F}_S^n$.

Let us denote by $\mathcal{F}_S'^n$ the family of all subsets $F \in \mathcal{F}_S^n$ with the property $F \supset S$ and by $\mathcal{F}_S''^n$ the family of all subsets $F \in \mathcal{F}_S^n$ having the property $F \bigcap S = \varnothing$. It is easy to observe that if we take $F' \in \mathcal{F}_S'^n$ and $F'' \in \mathcal{F}_S''^n$, then we have $F' \bigcap F'' \in \mathcal{F}_S''^n$, hence the intersections of sets of both families $\mathcal{F}_S'^n$ and $\mathcal{F}_S''^n$ are in fact intersections only with sets from $\mathcal{F}_S''^n$.

The space $X_a(F)$ being closed, for any $F \in \mathcal{F}_S^n$, we denote by $\mathcal{E}_S$ the application defined by the identity

$\mathcal{E}_S(F) = X_a(F)$, for $F \in \mathcal{F}_S^n$

and it is easy to show that $\mathcal{E}_S$ is an S-spectral capacity for a. Indeed, from Corollary 1.5.10, [16], $\sigma(a,x) = \varnothing$ implies $x = 0$, hence

$\mathcal{E}_S(\varnothing) = X_a(\varnothing) = \{0\}$,

$\mathcal{E}_S(\mathbb{C}^n) = X_a(\mathbb{C}^n) = X_a(\sigma(a, X)) = X$.

Let us verify the equality:

$\mathcal{E}_S\left(\bigcap_{i \in I} F_i\right) = \bigcap_{i \in I} \mathcal{E}_S(F_i)$, $F_i \in \mathcal{F}_S^n$, when $S_a = \varnothing$.

Let $x \in \mathcal{E}_S\left(\bigcap_{i \in I} F_i\right) = X_a\left(\bigcap_{i \in I} F_i\right)$;

then $\sigma(a,x) \subset F_i$, for all i, hence $x \in \bigcap_{i \in I} X_a(F_i) = \; = \bigcap_{i \in I} \mathcal{E}_S(F_i)$. Conversely, let $x \in \bigcap_{i \in I} \mathcal{E}_S(F_i) = \bigcap_{i \in I} X_a(F_i)$, hence $x \in X_a(F_i)$, for all i, whence it results that $\sigma(a,x) \subset F_i$, for all i. Then $\sigma(a,x) \subset \bigcap_{i \in I} F_i$,

consequently $x \in X_a\left(\bigcap_{i\in I} F_i\right) = = \mathcal{E}_S\left(\bigcap_{i\in I} F_i\right)$.

Let $\{G_S\} \cup \{G_j\}_{j=1}^m$ be an S-covering of $\mathbb{C}^n$. From condition (II), it follows that

$$X = X_a(\bar{G}_S) + \sum_{j=1}^m X_a(\bar{G}_j) = \mathcal{E}_S(\bar{G}_S) + \sum_{j=1}^m \mathcal{E}_S(\bar{G}_j).$$

On account of the relations above, it results that $\mathcal{E}_S$ is an S-spectral capacity for a. Since $X_a(F)$ is ultrainvariant to a, for any $F \in \mathcal{F}_S^n$ (Proposition 2.8), by using (I) we deduce

$$a_j \mathcal{E}_S(F) = a_j X_a(F) \subset X_a(F) = \mathcal{E}_S(F),$$

$1 \le j \le n$ and

$$\sigma(a, \mathcal{E}_S(F)) = \sigma(a, X_a(F)) \subset F$$

therefore a is S-decomposable.

Let us now consider the case $S_a \neq \varnothing$; then $X_a(F)$ does not make sense, for any $F \in \mathcal{F}_S^n$. Then $X_a(S \cup F)$ is a spectral maximal space of a, for any $F \in \mathcal{F}_S^n$ (see Remark 3.7) and we can write

$$X_a(S \cup F) = X_a(S) \oplus Y_F, \text{ for } F \in \mathcal{F}_S''^n,$$

where Y_F is also a spectral maximal space of a and $\sigma(a, Y_F) \subset F$.

If we define $\mathcal{E}_S$ by the following equalities

$$\mathcal{E}_S(F) = X_a(F), \text{ for } F \in \mathcal{F}_S'^n \ (F \supset S)$$

$$\mathcal{E}_S(F) = Y_F, \text{ for } F \in \mathcal{F}_S''^n \ (F \cap S = \varnothing)$$

where the space Y_F is defined above, we can show that $\mathcal{E}_S$ is an S-spectral capacity for a.

From the equalities

$$X_a(S \cup \varnothing) = X_a(S) \oplus Y_\varnothing = X_a(S)$$

$$X_a(S \cap \varnothing) = X_a(S) \cap Y_\varnothing = Y_\varnothing$$

it follows that $\mathcal{E}_S(\varnothing) = Y_\varnothing = \{0\}$. We obviously have that $\mathcal{E}_S(\mathbb{C}^n) = X$, because $\mathcal{E}_S(\mathbb{C}^n) = .$

$= X_a(\mathbb{C}^n) = X_a(\sigma(a, X)) = X$. Let us verify condition (2) from Definition 3.1:

$$\mathcal{E}_S\left(\bigcap_{i\in I} F_i\right) = \bigcap_{i\in I} \mathcal{E}_S(F_i), \qquad F_i \in \mathcal{F}_S^n, \qquad \text{when } S_a \neq \varnothing.$$

It is easily seen that the inclusion $F_1 \subset F_2$, for $F_1, F_2 \in \mathcal{F}_S^n$ implies $\mathcal{E}_S(F_1) \subset \mathcal{E}_S(F_2)$. Indeed, for $F_1, F_2 \in \mathcal{F}_S^n$ with $F_1 \supset S$ and $F_2 \supset S$, it is obviously that $X_a(F_1) \subset X_a(F_2)$, hence $\mathcal{E}_S(F_1) \subset \mathcal{E}_S(F_2)$; for $F_1, F_2 \in \mathcal{F}_S^n$ with $F_1 \cap S = \varnothing$ and $F_2 \cap S = \varnothing$, from the inclusion

$$Y_{F_1} \oplus X_a(S) = X_a(F_1 \cup S) \subset X_a(F_2 \cup S)$$

$$= Y_{F_2} \oplus X_a(S)$$

it follows that $Y_{F_1} \subset Y_{F_2}$.

Let us take $F_i \in \mathcal{F}_S^n$, with $F_i \supset S, i \in I$; we have $\bigcap_{i\in I} F_i \supset S$, $X_a\left(\bigcap_{i\in I} F_i\right)$ makes sense and using Lemma 3.8 we can write

$$\mathcal{E}_S\left(\bigcap_{i\in I} F_i\right) = X_a\left(\bigcap_{i\in I} F_i\right)$$

$$= \bigcap_{i\in I} X_a(F_i) = \bigcap_{i\in I} \mathcal{E}_S(F_i).$$

When $F_i \in \mathcal{F}_S^n$, with $F_i \cap S = \varnothing, i \in I$, then $\bigcap_{i\in I} F_i \subset F_j$, $j \in I$ implies $Y_{\bigcap_{i\in I} F_i} \subset Y_{F_j}$, $j \in I$, hence $Y_{\bigcap_{i\in I} F_i} \subset \bigcap_{i\in I} Y_{F_i}$; but $Y = \bigcap_{i\in I} Y_{F_i}$ is a spectral maximal space of a and $\sigma(a, Y) \subset \bigcap_{i\in I} F_i$, hence

$$Y \subset X_a\left(\left(\bigcap_{i\in I} F_i\right) \cup S\right) = Y_{\bigcap_{i\in I} F_i} \oplus X_a(S),$$

whence $Y = \bigcap_{i\in I} Y_{F_i} \subset Y_{\bigcap_{i\in I} F_i}$, consequently

$$\mathcal{E}_S\left(\bigcap_{i\in I} F_i\right) = \bigcap_{i\in I} \mathcal{E}_S(F_i).$$

If we consider two sets $F_1 \in \mathcal{F}_S'^n$ and $F_2 \in \mathcal{F}_S''^n$, then

$\mathcal{E}_S(F_1 \cap F_2) = \mathcal{E}_S(F_1) \cap \mathcal{E}_S(F_2)$. Indeed, it is obviously that $\mathcal{E}_S(F_1 \cap F_2) = Y_{F_1 \cap F_2} \subset Y_{F_1} \cap Y_{F_2} = X_a(F_1) \cap Y_{F_2} = \mathcal{E}_S(F_1) \cap \mathcal{E}_S(F_2)$. From the inclusion $\sigma\left(a, Y_{F_1} \cap Y_{F_2}\right) \subset F_1 \cap F_2$, it follows that $Y_{F_1} \cap Y_{F_2} \subset X_a\left(\left(F_1 \cap F_2\right) \cup S\right) = Y_{F_1 \cap F_2} \oplus X_a(S)$, hence $Y_{F_1} \cap Y_{F_2} \subset Y_{F_1 \cap F_2}$.

Finally, if $F_i \in \mathcal{F}_S^n$, $i \in I$ are arbitrary, then by putting $F_i' = F_i$, for $F_i \supset S$ and $F_i'' = F_i$, for $F_i \cap S = \varnothing$, we obtain

$$\mathcal{E}_S\left(\bigcap_{i\in I} F_i\right) = \mathcal{E}_S\left(\left(\bigcap_{i\in I} F_i'\right) \cap \left(\bigcap_{i\in I} F_i''\right)\right)$$

$$= \mathcal{E}_S\left(\bigcap_{i\in I} F_i'\right) \cap \mathcal{E}_S\left(\bigcap_{i\in I} F_i''\right) =$$

$$= \left(\bigcap_{i\in I} \mathcal{E}_S(F_i')\right) \cap \left(\bigcap_{i\in I} \mathcal{E}_S(F_i'')\right) = \bigcap_{i\in I} \mathcal{E}_S(F_i).$$

The conditions (3), (4), (5) from Definition 3.1 can be easily verified, by using conditions (I) and (II) of the hypothesis and therefore a is S-decomposable.

REFERENCES

[1] ALBRECHT, E.J., ESCHMEIER, J., *Analytic functional models and local spectral theory*, Proc. London Math. Soc., **75**, 323-345, 1997.

[2] ALBRECHT, E.J., VASILESCU, F.H., *On spectral capacities*, Rev. Roum. Math. Pures et Appl., **18**, 701-705, 1974.

[3] ALBRECHT, E.J., RICKER, W.J., *Local spectral theory for operators with thin spectrum*, Spectral Analysis and Its Applications: Ion Colojoară Anniversary Volume. Theta Series in Advanced Mathematics, 2003.

[4] APOSTOL, C., *Spectral decompositions and functional calculus*, Rev. Roum. Math. Pures et Appl., **13**, 1481-1528, 1968.

[5] BACALU, I. *On restrictions and quotients of decomposable operatos*, Rev. Roum. Math. Pures et Appl., **18**, 809-813, 1973.

[6] BACALU, I., *Măsuri spectrale reziduale*, St. Cerc. Mat. **27**, 377-379, 1975.

[7] BACALU, I., *Some properties of decomposable operators*, Rev. Roum. Math. Pures et Appl., **21**, 177-194, 1976.

[8] BACALU, I., *Descompuneri spectrale reziduale* I (1980), II (1980), III (1981), St. Cerc. Mat., **32**.

[9] COLOJOARĂ, I., FOIAŞ, C., *Quasi-nilpotent equivalence of not necessarily commuting operators*, J. Math. Mech., **15**, 521-540, 1966.

[10] COLOJOARĂ, I., FOIAŞ, C., *Theory of generalized spectral operators*, Gordon Breach, Science Publ., New York-London-Paris, 1968.

[11] COLOJOARĂ, I., *Generalized spectral operators*, Rev. Roum. Math. Pures et Appl., **7**, 459-465, 1962.

[12] COLOJOARĂ, I., *Spectral theory elements*, Pacific. J. Math., **4**, 321-354, 1954.

[13] DOWSON, H.R., *Restrictions of spectral operators*, Proc. London Math. Soc., **15**, 437-457, 1965.

[14] DUNFORD, N., *Spectral Operators*, Pacific J. Math., **4**, 321-354, 1954.

[15] ESCHMEIER, J., *On two notions of the local spectrum for several commuting operators*, Michigan Math. J., **30**, 245-248, 1983.

[16] FRUNZĂ, ŞT., *O teorie axiomatică a descompunerilor spectrale pentru sisteme de* operatori I *şi* II, St. Cerc. Mat., I, **27**, 655-711, 1975 and II, **29**, 329-376, 1977.

[17] KOSZUL, J.L., *Homologie et Cohomologie des Algèbres de Lie*, Soc. Math France, **78**, 65-127, 1950.

[18] MAC LANE, S., *Homology*, Springer-Verlag, New York/Berlin, 1963.

[19] MAC LANE, S.,BIRKOFF, G., *Algebra*, Mac Millan, 1967.

[20] NAGY, B., *Differential operators and spectral decomposition*, Colloquia Mathematica Societatis Jonas Bolyai, **35**, Functions, Series, Operators, Budapest (Hungary), 1980.

[21] TAYLOR, J.L., *A joint spectrum for several commuting operators*, J. Func. Anal., **6**, 172-191, 1970.

[22] TAYLOR, J.L., *The analytic functional calculus for several commuting operators*, Acta. Math., **125**, 1-38, 1970.

[23] VASILESCU, F.H., *Residually decomposable operators in Banach spaces*, Tôhoku Mat. Journ., **21**, 509-522, 1969.

[24] VASILESCU, F.H., *Calcul funcţional analitic multidimensional*, Ed. Academiei, 1979.

[25] VASILESCU, F.H., *Analytic Functional Calculus and Spectral Decompositions*, Ed. Academiei and D. Reidel Publishing Company, Bucharest and Dordrecht, 1982.

[26] ZAMFIR,M.,BACALU,I.,*Direct sums of decomposable systems*,Mathematical Modelling in Civil Engineering,vol 7,2011

First A. Author (M'76–SM'81–F'87) and the other authors may include biographies at the end of regular papers. Biographies are often not included in conference-related papers. This author became a Member (M) of **INASE** in 1976, a Senior Member (SM) in 1981, and a Fellow (F) in 1987. The first paragraph may contain a place and/or date of birth (list place, then date). Next, the author's educational background is listed. The degrees should be listed with type of degree in what field, which institution, city, state or country, and year degree was earned. The author's major field of study should be lower-cased.

The second paragraph uses the pronoun of the person (he or she) and not the author's last name. It lists military and work experience, including summer and fellowship jobs. Job titles are capitalized. The current job must have a location; previous positions may be listed without one. Information concerning previous publications may be included. Try not to list more than three books or published articles. The format for listing publishers of a book within the biography is: title of book (city, state: publisher name, year) similar to a reference. Current and previous research interests ends the paragraph.

The third paragraph begins with the author's title and last name (e.g., Dr. Smith, Prof. Jones, Mr. Kajor, Ms. Hunter). List any memberships in professional societies other than the **INASE**. Finally, list any awards and work for **INASE** committees and publications. If a photograph is provided, the biography will be indented around it. The photograph is placed at the top left of the biography. Personal hobbies will be deleted from the biography.

ISBN: 978-1-61804-292-7

An exact method for solving Multi-Objective Stochastic Integer Linear Programming

Salima Amrouche
USTHB, Department of Operational Research
Faculty of Mathematics
BP 32 El-Alia Bab-Ezzouar 16111
Algiers, Algeria
Email: samrouche@usthb.dz

Mustapha Moulaï
USTHB, LaROMaD Laboratory,
BP 32 El-Alia Bab-Ezzouar 16111
Algiers, Algeria
Email: mmoulai@usthb.dz

***Abstract*—Multi-objective stochastic integer programming is an optimization technique in which the objective functions and some constraints of an optimization problem contains integer variables and random data which follow discrete probability distribution. Once a problem requires a stochastic formulation, a first step consist in transforming the problem into an equivalent deterministic formulation. In the second case, it is necessary to transforming the multi-objective problem into a mono-objective problem. An algorithm combined the integer L-shaped method and branch and bound method with efficient cuts concept for the search of integer efficient solutions.This approach has the advantage to give the decision maker the efficient solution and their corresponding optimal cost values of the random constraint violations. A numerical example is included for illustration.**

I. Problem Formulation

We consider multi-objective integer linear programming problems involving random variable coefficients in both objective functions and some constraints of the following model (see [1], [6]):

$$(P)\begin{cases} min\ C_i(\xi)\,x, \quad i=1,...,k \\ T(\xi)\,x = h(\xi) \\ x \in S,\ x\ integer \end{cases}$$

where $k \geq 2$; C_i, T and h are random matrices defined on some probability space (Ω, E, P); $S = \{x \in R^n | Ax = b, x \geq 0\}$. The vector $b \in R^m$ and A the $m \times n$ real matrix are given and x is to be determined.

A. The transformed deterministic model

Assume that we have a joint finite discrete probability distribution of the random data : $\{(\xi^r,\ p_r),\ r=1,...,R\}$. R is the number of realizations (scenarios).

For each realization ξ^r we associate a criterion $C_i(\xi^r)\,x$, a matrix $T(\xi^r)$, a vector $h(\xi^r)$ and a recourse matrix $W(\xi^r) = W$. In this article the recourse matrix W does not change, this is called fixed recourse [3]. We assume that the decision maker is able to satisfactorily specify the penalties q^r of the constraint violations z_r. The above problem is equivalent to the so-called Deterministic Equivalent (DE) Multi-Objective Integer Linear Programs $MOILP$ [2].

$$(DE)\begin{cases} min\ Z_i = \acute{Z}_i + Q(x),\ i=1,...,k \\ x \in S,\ x\ integer \end{cases} \tag{1}$$

where $\acute{Z}_i = E_\xi[C_i(\xi)\,x]$, $Q(x) = E_\xi[Q(x,\xi)]$ and

$$Q(x,\xi^r) = \min\{(q^r)^t z \mid T(\xi^r)x + W(\xi^r)z = h(\xi^r), z \geq 0\}. \tag{2}$$

B. The transformed mono-objective model

We transform the DE problems in the following way

$$(DE_\lambda)\begin{cases} min\ \lambda^T \acute{Z} + Q(x) \\ x \in S,\ x\ integer \end{cases}$$

where $\lambda^T = (\lambda_1, ..., \lambda_k) \geq 0$, with at least one component strict inequality.

II. Solution method

$$(P_l)\begin{cases} \min \quad \lambda^T E\left[C_i^T(\xi)\,x\right] + \theta \\ D_l x \geq d_l, \qquad l=1,...,N \\ E_l x + \theta \geq e_l, \qquad l=1,...,M \\ x \in S_l \end{cases} \tag{3}$$

(P_l) obtained at node l in a structured tree.
N_l is the indices set of non-basic variables of efficient solution x^l;
$H_l = \left\{ j \in N_l \mid \widehat{c}^i_{jl} < 0 \right\}$, where $\widehat{c}^i_{jl}$ is the component j of the reduced cost vector of the objective function $\acute{Z}_i$.
$S_{l+1} = \left\{ x \in S_l \ \setminus \sum_{j \in H_l} x_j \geq 1 \right\}$.
N and M indicate the number of feasibility cuts and optimality cuts, respectively, added until step l.

Starting with $\theta = -\infty$ and without feasibility cuts, optimality cuts and efficiency cuts. The objective $\lambda^T E\left[C^T(\xi)\,x\right]$ is minimized under the deterministic constraints. If for some realizations the second stage problems yielded by the solution x are not feasible, a feasibility cut is introduced. Then, the problem is optimized again to obtain another feasible point x. If x is not integer,

ISBN: 978-1-61804-292-7

create two new branches at fractional component of x; and get the ideal solution of each node then append the new nodes to the list of dangling node if these nodes is not dominated by ideal solution. Using x^l, we solve the recours problem for all realizations $\xi^r,\ r \in \{1,...,R\}$, and compute $Q\left(x^l\right)$. If $\theta^l < Q\left(x^l\right)$ then a new optimality cut $E_l x \geq e_l$ is added to the current problem(P_l). In presence of integer optimal solution, An efficient cut $\sum_{j \in H_l} x_j \geq 1$ is then added for deleting integer solutions that are not efficient and the new program is solved. The method terminates when all the created nodes are fathomed.

III. Numerical Illustration

Let us consider the following example with a similar structure to that of problem $(P), k = 3, n_0 = 4, m_0 = m = n = 2,\ R = 2.$

Deterministic constraints : $\begin{cases} -4x_1 + 2x_2 \geq -8 \\ x_1 + x_2 \leq 5 \end{cases}$

Objective function

$$C\left(\xi^1\right) = \begin{pmatrix} -9 & 4 \\ 3 & -5 \\ 8 & -11 \end{pmatrix}, C\left(\xi^2\right) = \begin{pmatrix} 3 & -2 \\ 7 & 1 \\ -4 & 9 \end{pmatrix};$$

$$T\left(\xi^1\right) = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}, T\left(\xi^2\right) = \begin{pmatrix} 1 & 0 \\ 3 & 4 \end{pmatrix};$$

$$h\left(\xi^1\right) = \begin{pmatrix} 3 \\ 5 \end{pmatrix}, h\left(\xi^2\right) = \begin{pmatrix} 6 \\ 1 \end{pmatrix};$$

$q^1 = q\left(\xi^1\right) = (1,0,6,2)^t$, $q^2 = q\left(\xi^2\right) = (5,3,2,1)^t$; $p\left(\xi^1\right) = \frac{1}{2}, p\left(\xi^2\right) = \frac{1}{2}$;

$$W(\xi) = W = \begin{pmatrix} -2 & -1 & 2 & 1 \\ 3 & 2 & -5 & -6 \end{pmatrix};$$

$\acute{Z}_1 = (-3,1)$, $\acute{Z}_2 = (5,-2), \acute{Z}_3 = (2,-1)$.

$\lambda^T = (1,\ 1,\ 1), \lambda^T E\left[C^i(\xi)\right] = (4,-2)$

- **First iteration**

$\Xi_0 = \{(P_0)\}$. The resolution of the problem (P_0)

$$(P_0) \begin{cases} \min & 4x_1 - 2x_2 \\ s.t. & 4x_1 - 2x_2 + x_3 = 8 \\ & x_1 + x_2 + x_4 = 5 \\ & x_1,\ x_2,\ x_3,\ x_4 \geq 0 \end{cases}$$

gives the following table I:

, minimum is at $x^0 = (0,5)$.

- Test of feasibility:

$$h\left(\xi^1\right) - T\left(\xi^1\right)x^0 = \begin{pmatrix} -7 \\ 0 \end{pmatrix}$$

$$h\left(\xi^2\right) - T\left(\xi^2\right)x^0 = \begin{pmatrix} 6 \\ -19 \end{pmatrix}$$

TABLE I
Stochastic feasible solution

B	Rhs	x_1	x_4
x_3	18	6	2
x_2	5	1	1
$\lambda^T \acute{Z}$	10	6	2
$\acute{Z}_1$	-5	-4	-1
$\acute{Z}_2$	10	7	2
$\acute{Z}_3$	5	3	1

TABLE II
Optimality cut

B	Rhs	x_1	x_4
x_3	18	6	2
x_2	5	1	1
θ	$\frac{26}{4}$	$\frac{7}{4}$	$\frac{5}{4}$
$\lambda^T \acute{Z}$	10	6	2
$\acute{Z}_1$	-5	-4	-1
$\acute{Z}_2$	10	7	2
$\acute{Z}_3$	5	3	1

$$\begin{cases} \max & -7\sigma_1^1 + 0\sigma_1^2 \\ s.t. & \sigma_1^T W \leq 0 \\ & \sigma_1^1 + \sigma_1^2 \leq 1 \end{cases}$$

, maximum is at : $\sigma_1^T = (0,0)$

$$\begin{cases} \max & 6\sigma_2^1 - 19\sigma_2^2 \\ s.t. & \sigma_2^T W \leq 0 \\ & \sigma_2^1 + \sigma_2^2 \leq 1 \end{cases}$$

, maximum is at : $\sigma_2^T = (0,0)$

$d_1 - D_1 x^0 = 0, d_2 - D_2 x^0 = 0$. This means that x^0 is feasible for the second stage.

Optimality Test:

$$\begin{cases} \max & -7\pi_1^1 + 0\pi_1^2 \\ s.t. & \pi_1^T W \leq (1,0,6,2)^T \end{cases}$$

,maximum is at : $\pi_1^T = \left(-1, \frac{-1}{2}\right)$

$$\begin{cases} \max 6\pi_2^1 - 19\pi_2^2 \\ \pi_2^T W \leq (5,3,2,1)^T \end{cases}$$

, maximum is at : $\pi_2^T = (1,0)$.

$e_1 = \sum_{r=1}^{2} p_r \pi_r^t h\left(\xi^r\right) = \frac{1}{4}$, $E_1 = \sum_{r=1}^{2} p_r \pi_r^t T\left(\xi^2\right) = \left(\frac{1}{2},\ \frac{-5}{4}\right)$, $Q\left(x^0\right) = e_1 - E_1 x^0 = \frac{13}{2}$, $\theta^0 = -\infty < \frac{13}{2} = Q\left(x^0\right)$,

$E_0 x + \theta^0 \geq e_0 \Leftrightarrow \frac{-7}{4}x_1 - \frac{5}{4}x_4 + s_1 - \theta = \frac{-26}{4}$. Add this optimality cut to the last obtained table I to obtain table II.

The minimum is reached to $x^0 = (0,5)$, $Z^0 = \left(\frac{23}{2}, \frac{-7}{2}, \frac{3}{2}\right)$.

$\bar{Z} = (+\infty, ..., +\infty)$ is dominated by Z^0 then the set of non dominated solutions is updated $ND = ND \cup \left\{Z^0\right\}$

ISBN: 978-1-61804-292-7

TABLE III
STOCHASTIC EFFICIENT SOLUTIONS

Iteration	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$
x^i	$(2,3)$	$(1,3)$	$(1,4)$	$(0,4)$	$(0,5)$
θ^i	3	$\frac{7}{2}$	$\frac{19}{4}$	$\frac{21}{4}$	$\frac{13}{2}$
$\acute{Z}^i$	$(-3,4,1)$	$(0,-1,-1)$	$\left(1,\frac{-7}{2},-2\right)$	$(4,-8,-4)$	$(5,-10,-5)$
$\tilde{Z}^i=\acute{Z}^i+\theta^i$	$(0,7,4)$	$\left(\frac{7}{2},\frac{5}{2},\frac{5}{2}\right)$	$\left(\frac{23}{4},\frac{5}{4},\frac{11}{4}\right)$	$\left(\frac{37}{4},\frac{-11}{4},\frac{5}{4}\right)$	$\left(\frac{23}{2},\frac{-7}{2},\frac{3}{2}\right)$

Proceeding in this manner, the algorithm terminates when all dangling nodes are fathomed, then the set of the whole integer efficient solutions is given by table III:

IV. CONCLUSION

In this paper, an exact method combining integer L-Shaped method [7] with efficient cuts is described for generating all efficient solutions for multiple objective stochastic integer linear programming problems [4], [5], [8] is presented.

REFERENCES

[1] M. ABBAS, F. BELLAHCENE "Cutting plane method for multiple objective stochastic integer linear programming", European journal of Operational Research 168 (2006) 967-984.

[2] M. Abbas; M.E. Chergui; M. Ait Mehdi "Efficient cuts for generating the non-dominated vectors for Multiple Objective Integer Linear Programming", International Journal of Mathematics in Operational Research. Volume 2008, DOI: 10.1504/IJMOR.2012.046690, pages:302-316.

[3] S. Amrouche, M. Moulaï "Multi-objective stochastic integer linear programming with fixed recourse", International Journal of Multicriteria Decision Making (IJMCDM), 2012 Vol. 2 No. 4, pages:355-378.

[4] F. Ben Abdelaziz, M. Masmoudi,November 2014, "A multiple objective stochastic portfolio selection problem with random Beta",International Transactions in Operational Research, Vol 21, Issue 6, pp 919933.

[5] H. Bonnel, J. Collonge, August 2014,"Stochastic Optimization over a Pareto Set Associated with a Stochastic Multi-Objective Optimization Problem", Journal of Optimization Theory and Applications,Volume 162 Issue 2,pp 405-427

[6] A. Bustos, L. Herrera, E. Jimnez, August 2014, "Efficient Frontier for Multi-Objective Stochastic Transportation Networks in International Market of Perishable Goods", Journal of Applied Research and Technology. JART, Vol. 12. Nm. 04.

[7] CAROE and TIND, 1998. "L-Shaped Decomposition of Two-Stage Stochastic Programs with Integer Recourse", European Journal of Operations Research, 101, 306-316.

[8] Xiaobing Liu, Zhancheng Li, Li He,September 2014, "A Multi-objective Stochastic Programming Model for Order Quantity Allocation under Supply Uncertainty, International Journal of Supply Chain Management IJSCM, Vol. 3, No. 3,pp 24-32.

Efficient matching for the Iterative Closest Point algorithm by using low cost distance metrics

H. Mora-Mora*, J. Mora-Pascual, P. Martinez-Gonzalez, A. Garcia-Garcia

Abstract—Since its introduction, the Iterative Closest Point algorithm (ICP) has become one of the most popular methods for the geometric alignment of three-dimensional models. Given two point clouds, named model and source, the algorithm iteratively refines a transformation which is applied to the source cloud in order to minimize the difference or distance between both point clouds. Many applications of multiple fields currently use this algorithm to reconstruct 2D or 3D surfaces from different data measurements due to its simplicity and effectiveness. However, one of the main problems of the algorithm is the high computational cost of certain complex phases when dealing with high density point clouds. This fact renders impossible some of the applications of the algorithm. The goal of this work is the improvement of the ICP algorithm so that a broader range of computational resources demanding problems can be addressed. For that, a convergence analysis and validation of point-to-point distance metrics with a lower computational cost than the Euclidean one which is used as a de facto standard in the existing implementations in the literature of the algorithm.

Keywords—Convergence of numerical methods, Error analysis, Computational geometry, Computational efficiency, ICP algorithm.

I. Introduction

Nowadays, range sensors obtain depth information so that we can capture three-dimensional datasets from different points of view, each one of them represented using a particular coordinate system. A lot of applications require a full or partial scene reconstruction from the data provided by the sensors over different points of view. In order to reconstruct the surfaces or shapes of the original scene, we have to combine the different datasets with their own coordinate systems in a process called *shape registration*.

The goal of shape registration is the transformation of different three-dimensional datasets to represent them in one common coordinate system so that those elements which overlap in both sets are properly aligned allowing the reconstruction of the original surfaces. The registration process may be applied to rigid or non-rigid shapes. In the case of rigid shapes, the transformation which aligns both surfaces is rigid too (rotation and translation) so that the solution space is bounded to six degrees of freedom (*6DoF*). On the other hand, non-rigid shape require a non-rigid transformation which takes into account the possibility of deformation so that the solution space is increased considerably. A remarkable study of the different registration techniques both rigid and non-rigid is the work by Tam *et al.* [28] which includes an in depth review of the registration problem; other important surveys are the ones by Van Kaik *et al.* [29] (focused on shape correspondence techniques) and Audette *et al.* [3] (centered on shape registration applied to medical images).

In addition, registration can be classified according to its granularity, distinguishing coarse and fine grained methods. The objective of coarse grained registration is to obtain a quick estimate of the transformation to roughly align both shapes while fine grained techniques use that initial estimate to refine it iteratively in order to find the best alignment in terms of precision under a set of restrictions. Reviews of multiple techniques and methods that can be applied to solve both coarse and fine problems have been carried out by Salvi *et al.* [25] paying special attention to precision, robustness and efficiency assessment; subsequently, Wang *et al.* [15] extended this survey.

In this work we will deal with the ICP algorithm which is a fine grained registration method which is currently the most popular algorithm for 3D rigid registration for the Robotics community; the cause of that popularity is mainly its simplicity and effectiveness as well as the many variants that have been developed, adapting the ICP to different scenarios to improve its efficiency and precision [24].

As we mentioned before, a wide set of applications of different fields make use of this algorithm in order to compute rigid registrations due to its simplicity and effectiveness; however, the algorithm has got a high computational complexity (quadratic with respect to the number of points in its original variant) which renders impossible or at least difficults certain applications which require the processing of high density point sets provided by high precision sensors. Many variants have been proposed in the literature to improve its performance, either by reducing the number of points or by decreasing the needed iterations or even reducing the complexity of its most expensive phase in terms of computing resources: the search of nearest neighbors. Nevertheless, despite reducing its complexity, in many cases those variants tend to have a negative impact on precision or even on the convergence domain, limiting the possible application scenarios.

Therefore, it is a fact that any general improvement of the algorithm which is able to accelerate its execution, without affecting the quality nor reducing its possible application scenarios, is a step forward for all those applications with needs of precision and quality in their rigid registrations. In that sense, this work proposes a general improvement for the algorithm, carried out by an interdisciplinary research based on the fusion of mathematical and geometric concepts, such as distance metrics, with its computational component by taking into account their associated operative cost and their impact on the algorithm's execution time. Our research group has expe-

H. Mora-Mora, J. Pascual-Mora, P. Martinez-Gonzalez and A. Garcia-Garcia are with the Department of Computer Technology and Computation, University of Alicante, Spain, 03690, San Vicente del Raspeig, Alicante, Spain. e-mail: ({hmora, jeronimo, pmartinez, agarcia}@dtic.ua.es).

ISBN: 978-1-61804-292-7

rience and successful examples of this kind of research about accelerating mathematical methods at a low level [13], [16], [26]. The hypothesis of the improvement is the following one: distance metrics whose computational cost is reduced respect to the Euclidean's one, like Manhattan or Chebyshev distances, may replace it and effectively reduce the computational cost while preserving the algorithm's convergence properties as well as the registration quality in terms of convergence domain and final registration error.

The rest of this paper is structured as follows: Section II provides a general view of the most remarkable variants of the ICP algorithm. Section III formulates the rigid registration problem, establishing the notation which will be used during the rest of the paper. Section IV describes or proposal in a detailed manner. Subsequently, Section V declares the comparison methodology. Finally, Section VI presents the results of the experiments, which are discussed along Section VII. At last, Section VIII concludes this work with an overview of the results and the accomplished goals to end the paper putting it in context and enumerating some future work possibilities.

II. Related work

The work of Rusinkiewicz and Levoy [24] is one of the main reviews about variants of the ICP algorithm and efforts directed towards the improvement of the algorithm's efficiency at some aspect. That review states that the ICP algorithm was introduced by parallel research works, one conducted by Besl and McKay [4] and the other performed by Chen and Medioni [5]. The difference between both works lays on the scope of the method: on the one hand, Chen and Medioni consider the specific problem of the alignment of multiple range images while Besl and McKay propose the method using a more general point of view, taking into account three-dimensional shape registration using multiple representation forms, being the point set the most popular one. In addition, this method has a singular advantage: its convergence can be proven; that is why, Besl and McKay's paper has become so popular that their approach is considered the standard ICP algorithm.

The variants arising from the original algorithm are usually classified following the taxonomy created by Rusinkiewicz and Levoy [24]. That classification has six categories which represent each one of the phases in which the algorithm can be divided: point selection, matching, pair weighting, outlier removal, error metric and minimization.

Here we will review the most remarkable variants of the algorithm focusing on matching and error metric phases because our proposal will directly impact them.

A. Matching

The matching phase establishes the correspondences between the points of one point cloud and the closest ones in the other. The original algorithm uses the Euclidean distance as the metric to make correspondences. Some variants have focused their contribution on the inclusion of additional properties in the metric like surface normals [22], color [30] and even some invariant features [9]. Other contributions adapted the algorithm to take into account anisotropic and heterogeneous noise, it is the case of the works carried out by Pennec *et al.* [21] and Hansen *et al.* [12] in which the Euclidean metric was replaced by the Mahalanobis distance, proving the possibility of exchanging the distance metrics. Other variants, outside the scope of this work, changed the matching strategy by using expectation maximization techniques [10], normal shooting [5], reverse calibration [2] or point-ray distances [8].

Our proposal introduces changes in this stage, the Euclidean distance metric is replaced by other metrics like Manhattan or Chebyshev to optimize the computational cost.

B. Error metric

There are two main error metrics which are profoundly tested and widely used: the point-to-point metric introduced by Besl and McKay [4] and the point-to-plane one by Chen and Medioni [5]. The point-to-point distance of Besl and McKay consists of the summation of the quadratic distances between the points of the model and source point clouds: $\sum_{i=1}^{N} \|Rd_i + t - m_i\|^2$. On the other hand, the point-to-plane distance of Chen and Medioni takes into account the distance between the points of the source to the tangent planes in which the model points are: $[(Rd_i + t - m_i) \cdot \vec{n_i}]^2$. These basic error metrics can be modified to take into account other variants of the original algorithm to improve its robustness. In fact, many of the previously mentioned variants apply changes over the error metric.

In our case, we will use the original point-to-point error metric by Besl and McKay.

C. Other remarkable variants

Other variants have focused their contribution on the extension of the ICP algorithm for non-rigid registration [1], [6], [9], [18] or even on the inclusion of *a priori* knowledge on the original algorithm [7].

As we previously noted, one of the main problems of the algorithm is its high complexity, quadratic with respect to the number of points, because of the need of computing the distances of all points of one point cloud to all points of the other in order to obtain the closest point. In that sense, a lot of variants have directed their efforts toward the reduction of that complexity by using kd-trees [27], closest points caching [27] or even parallel implementations on CPU [14] or GPU [17], [20], [31].

III. Problem formulation

Rigid registration can be formulated as an optimization problem with certain restriction whose objective is the alignment of surfaces or three-dimensional data, for the sake of simplicity we will assume that the data will be given in the form of point sets or clouds, although it can be applied to several geometric representation forms. In this problem, two point sets of n dimensions $\mathbf{M} \doteq \{m_i\}_{i=1}^{N_M} \in \mathbb{R}^n$ y $\mathbf{D} \doteq \{d_i\}_{i=1}^{N_D} \in \mathbb{R}^n$, also known in the literature of the algorithm as $model$ y $data$ (being $N_M \in \mathbb{R}$ y $N_D \in \mathbb{R}$ the cardinalities of the sets M y D). The objective is to align the source point cloud with the model one, in other words,

ISBN: 978-1-61804-292-7

obtain the rigid transformation Φ which minimizes the mean square error between the model and the source points once the transformation (a rotation R and a translation T) is applied to the source point set D. In order to simplify even more the explanation, we assume that $N_M = N_D$ and each point d_i to have is corresponding point m_i so the objective function that has to be minimized is:

$$f(R,T) = \frac{1}{N_D}\sum_{i=1}^{N_D}\|m_i - R(d_i) - T\|^2 \quad (1)$$

Taking this formulation into account, if the right correspondences between the model points and the ones of the source point cloud are known, we can find the optimal relative transform to align both point sets in one step. However, the difficult part of this problem when applied to a real life scenario is that the correspondences are unknown and it is also possible that some points of the source set have no corresponding point in the model and vice versa. For that, a need for a method for establishing the correspondences arises.

A. The ICP algorithm

The ICP algorithm, originally described by Besl and McKay [4], is one of the most popular and widely used methods for rigid registration. Its functioning is based on the closest point criteria used for establishing the correspondences, so that the corresponding point for a source point is its closest one in the model. The distances between the points are calculated by using the Euclidean distance metric to define the closest point operator.

B. Finding matches

Given two three-dimensional points p_1 and p_2, the Euclidean distance between both of them $d(p_1, p_2)$ is the length of the segment which connects them $\|p_1 - p_2\|$. Given a point p and a point set A, we define the Euclidean distance of the point to the set $d(p, A)$ as the minimum of the distances of p to each one of the points of the set A, in other words, $d(p, A) = \min_{i \in 1...n} d(p, a_i)$. The function c which obtains the closest point to p in the point set A is the following one:

$$c(p, A) = \underset{a \in A}{argmin} \quad d(p, a) \quad (2)$$

The algorithm sets a correspondence between each point d_i of the source point set D and the closest point in the model which will be named $y_i \in Y$, forming the set of closest points to D. From this statement we deduce that $Y \subseteq M$, $y \in M$ and $N_Y = N_D$.

The closest point operator C which produces the point set $Y = C(D, M)$. This operator obtains the set $Y \doteq \{y_i\}_{i=1}^{N_D}$ in which the point y_i is the closest point in the model to the point d_i to the source point set.

$$C(D, M) = \{y_i = c(d_i, M)\}_{i=1}^{N_D} \quad (3)$$

Assuming this closest point criteria, the algorithm ensures the convergence if the initial position of the source point set is close enough to the model set position.

C. Phases of the algorithm

Given that, in general, the correspondences obtained using the closest point operator are not the right nor the best ones from the beginning, the ICP algorithm performs an iterative refinement process. Each one of the iterations comprises three main phases that can be extended, as we observed in Section II, to improve different aspects of the algorithm. The main phases are:

1) **Correspondences or matching**: In this phase, the closest point operator is applied to obtain the closest points set Y.
2) **Transformation calculation or *minimization*:** In this phase, the algorithm tries to find the rotation R and the translation T which minimize the objective function of rigid registration taking into account the correspondences.

$$f(R,T) = \frac{1}{N_D}\sum_{i=1}^{N_D}\|y_i - (R(d_i) - T)\|^2 \quad (4)$$

3) **Update transformation or apply it:** In this last phase, the transformation is accumulated or the source point cloud is transformed by applying it so that the new position for each point d_i is calculated as follows $d_i = R(d_i) + T$.

These phases are repeated until a certain stop criteria such as a limit for the number of iterations or a threshold for the difference of final registration error of the current iteration and the previous one so that the algorithm stops if the transformation has enough refinement. By using this process, the algorithm's convergence is stated in the following theorem: the ICP algorithm always converges monotonically to a local minimum with respect to the objective mean squared distance error function in equation 4.

IV. Proposed improvement

Our proposal has the goal of reducing the execution time of the algorithm in general by means of a reduction of the computational cost of the matching phase. In this way, other existing variants of the algorithm may be able to include this improvement to increase execution speed.

In order to do that, our proposal replaces the Euclidean distance metric, widely used in the existing implementations of the algorithm to find the closest points during the matching phase with the closest points operator, with other metrics with a lower operative cost and thus reducing the execution time of that phase. In addition, these new metrics must provide similar quality as the Euclidean one, since it would not be useful to reduce the execution time if the algorithm is not able to provide an acceptable result because we are in a context of high precision and lots of data.

In the previous section, the ICP algorithm was formulated in function of a distance metric d and its functioning was analyzed as convergent to a local minimum. In this section we will proceed to define what a distance metric is and then analyze the operative cost of the Euclidean, Chebyshev and Manhattan distance metrics, being these two last ones our candidate metrics for replacing the Euclidean one since we expect an inferior computational cost from them. Next, we will carry out an empirical study of the computational

ISBN: 978-1-61804-292-7

cost of the three metrics in order to estimate the possible performance gain or speedup that we can expect when using a certain metric, providing a starting point for comparing the algorithm's performance with the new metrics for matching.

A. Distance metrics

A distance metric d is defined as a non-negative function in a set X so that $d : X \times X \to R$, being R the set of natural numbers. This function describes the distance between points, for example x, y, z, of the X set. Furthermore, it must meet the following conditions [11]:

1) $d(x, y) + d(y, z) \geq g(x, z)$
2) $d(x, y) = d(y, x)$
3) $d(x, x) = 0$
4) $d(x, y) = 0 \implies x = y$

Once the concept of distance metric has been defined, we can address the formulation of the Euclidean metric, used as a standard in the original algorithm, as well as the formulation of the Chebyshev and Manhattan metrics proposed as candidates for the reduction of the computational cost because of their low operative complexity.

1) Original Euclidean distance: The Euclidean distance between the points x and y is defined as the length of the segment which connects both of them $\overline{xy}$. In Cartesian coordinates, if $x = (x_1, x_2, ..., x_n)$ and $y = (y_1, y_2, ..., y_3)$ are two points in a Euclidean n-space, then the distance from x to y or y to x is determined by the equation 5.

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \tag{5}$$

In terms of operative cost, this metric requires the calculation of n products, $2n-1$ sums/subtractions and 1 square root to obtain the distance between two n-dimensional points.

2) Chebyshev distance: The Chebyshev distance between the points x and y is defined as the maximum of the absolute values of the differences between their coordinates. In this way, if $x = (x_1, x_2, ..., x_n)$ and $y = (y_1, y_2, ..., y_3)$, then the Chebyshev distance from x to y or y to x is described by the equation 6.

$$d(x, y) = \max\left(|x_1 - y_1|, |x_2 - y_2|, ..., |x_n - y_n|\right) \tag{6}$$

In terms of operative cost, the Chebyshev metric requires n subtractions, $n - 1$ comparisons and n absolute values to obtain the distance between two n-dimensional points.

3) Manhattan distance: The Manhattan distance between the points x and y is defined as the sum of the absolute values of the differences of their coordinates. Being $x = (x_1, x_2, ..., x_n)$ and $y = (y_1, y_2, ..., y_3)$, then the Manhattan distance from x to y or y to x is determined by the equation 7.

$$d(x, y) = \sum_{i=1}^{n} |x_i - y_i| \tag{7}$$

The operative cost of this metric is $2n-1$ sums/subtraction and n absolute values to obtain the distance between two n-dimensional points x and y.

B. Computational cost comparison

As we could see before, it is clear that both the Chebyshev and Manhattan metrics have a lower operative complexity than the Euclidean one: the n products are replaced by n absolute values whose cost is significantly inferior, and the square root operation is no longer needed. However, this analysis does not allow us to quantify the performance gain obtained when using one metric or another because of the differences of the processor architectures which include a different set of arithmetic operations which are implemented differently across the families of processors and microarchitectures. In this sense, the only way of estimating the performance gain is by determining the specific architecture operations needed to implement the metric and calculate the number of cycles needed for them for each processor family. In addition, modern processors have reached an extremely high complexity level in which superscalar processing and all kinds of instructions optimization techniques take place so that it is quite difficult to estimate *a priori* the gain or speedup obtained by using a metric or another.

In this sense, before carrying out more extensive set of tests to prove the viability of the use of the candidate metrics and their impact on the results of the algorithm, we have performed a set of empirical tests to determine an estimate of the performance gain or speedup that may be expected by using one metric instead of another. By doing this, we can confirm that it is really worth to use low cost distance metrics before modifying the algorithm. The rest of this section is dedicated to describe the viability study.

The different distance metrics have been codified in C++. For each one of them a test has been performed which consisted of the computing of ten millions of distance operations over two three-dimensional points. We have carried out one hundred executions for each test with only the basic services of the computer in execution to avoid altering the results; the arithmetic mean of those one hundred executions has been obtained discarding those times which were deviated a 20% from the median to smooth the noise introduced by the different processes of the operating system being executed in the background.

For that, we have used a typical platform currently, based on the x86 architecture. The machine used to execute the tests has got the following specifications: the operating system is Debian 7.1, 64 bits based version, the processor used is an Intel Core i5 2410M (2 cores, 4 threads, frequency 2.3 GHz), the motherboard chipset is an Intel HM65 Express, the main memory is composed by a single DDR3 4 GiB stick working at a frequency of 1333 MHz.

To compile the executable we have used the C++ GNU compiler *g++* version *4:4.7.2-1*, carrying out a basic compilation with no optimization flags: *g++ -o main main.cpp*.

The results of the benchmark, in terms of percentage of change of the execution time needed to perform the ten million

ISBN: 978-1-61804-292-7

of distance operations, are shown on Table I together with the speedups obtained over the Euclidean metric.

TABLE I
PERCENTAGE OF CHANGE OF EXECUTION TIMES AND SPEEDUPS OF THE DIFFERENT METRICS AFTER TEN MILLION DISTANCE CALCULATIONS OVER THREE-DIMENSIONAL POINTS.

	Euclidean	**Chebyshev**	**Manhattan**
% of change (execution time)	0%	-20%	-43%
Speedup	1.000	1.247	1.766

As we can see, the Chebyshev distance shows a small improvement in terms of execution time of approximately a 20% over the Euclidean distance metric while the Manhattan one achieves a significant improvement which can be quantified in a 43% over the Euclidean one approximately. These facts confirm our operative cost analysis that we performed in previous subsections; it is a remarkable fact that the Mahanttan distance is clearly better than the Chebyshev one while its operative complexity is quite similar; this happens because of the cost of the call to the *maximum* function which is higher than performing a simple sum or subtraction which requires no additional logic.

In addition, we would like to note that we can't expect to obtain an improvement of a 20% or a 43% in the execution time of the algorithm just by simply applying the Chebyshev or Manhattan distance metrics because this change would only accelerate the matching phase which is just a fraction of the overall computation that is performed by the algorithm. The obtained speedup is bounded by different factors such as the computing fraction P represented by the accelerated phase and the total speedup that we can apply to that fraction S_p according to the Amdahl's law [23] shown in the equation 8.

$$S = \frac{1}{(1-P) + \frac{P}{S_p}} \tag{8}$$

V. VALIDATION OF TOPOLOGICAL SPACES

In order to assess the performance and quality of our proposal, we carried out several experiments on a heterogeneous set of synthetic situations trying to generate a representative sample of the different surfaces and scenarios where three-dimensional rigid registration techniques are often used. These scenarios are described in section V-A

For all the experiments, we have used our own Matlab implementation of the original ICP algorithm from Besl and McKay without any optimization and with the needed modifications to include the different metrics. Execution time data has been obtained with the own Matlab tools for time measurement. Each test has been performed one hundred times so we took the average time of all executions, discarding values with a deviation of 20% from the median in order to avoid noise due to system overloads while executing background processes.

For each test scenario we have considered three different situations: one without any noise, other with noise in both the model and the data, and the last one only applying noise to the data to be registered and not to the model. The goal of these situations is to verify the robustness of the proposal against noise. The applied noise consists on the application of random displacements to all coordinates of all the points of the set. This deviations are bounded in order not to deform excessively the point cloud.

Additionally, each scenario with a particular noise situation has been executed for the three distance metrics already mentioned before: Euclidean, Manhattan and Chebyshev. The results of these experiments are presented in Section VI and their discussion in Section VII.

A. Test scenarios

Scenarios are synthetic situations generated from the Venus shape [19] by applying a set of cuts, and deleting some points with the goal of generating an heterogeneous set of possibilities for the registration. For each test, we will try to register a point set into another one called *model*; for each scenario, the source set has been initially transformed applying a rotation of $R_0(30°, 20°, 15°)$ and a translation of $T_0(0.12m, -0.08m, 0.1m)$ to the surface.

1) Full: In this first scenario we perform a registration where both model and data have exactly the same point set: the complete shape composed by 5688 points. The starting situation of this scenario (once applied the initial transformation explained before) is represented without noise in Figure 1. It is a simple situation for the algorithm and it is not very useful from a practical point of view, although it allows us to evaluate the difference among metrics in a basic scenario.

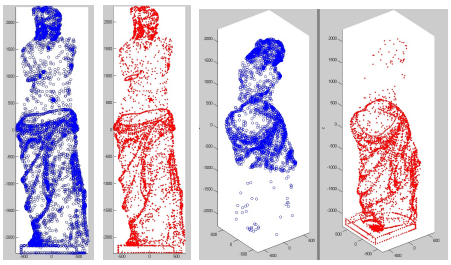


Fig. 1. Left: Model used for the registration and data set for the full scenario. They are the same shape, as we can see. Right: Model used for the registration in the horizontal cut with overlap scenario and data set for the same scenario. We can observe an overlapping area between the two sets.

2) Horizontal cut (top): In this scenario we have performed a horizontal cut to the point set, leaving just the top part of the Venus to be registered with the complete model. The starting situation is shown in Figure 2, without noise. Despite being a simple and unrealistic scenario, this kind of situation is closer to those found in practical applications of the algorithm. Furthermore, this part of the shape has a smaller number of points in a simpler surface.

3) Double horizontal cut: In this situation, which is a combination of the two previous scenarios, we have performed a section of the figure with two horizontal planes, and we will try to register it to the complete model. This situation is common in several practical applications of the algorithm such as tomography image registration. Figure 2 shows the resulting data set to be registered.

ISBN: 978-1-61804-292-7

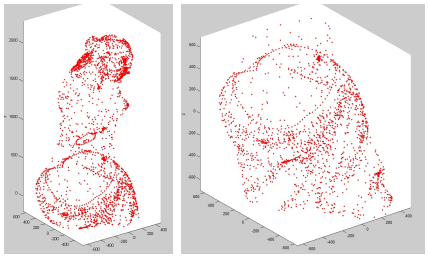


Fig. 2. Data set for top horizontal cut (left) and double horizontal cut (right) scenarios.

4) Horizontal cut (overlap): The last scenario is composed by a combination of both horizontal cuts, top and inferior; in this case a top horizontal cut has been applied to the model (the model consists of the top part of the shape) and an inferior horizontal cut has been done to the data set (the bottom part of the shape), so that an overlapped area exists between the two point sets. This last situation is the one that more resembles a real three dimensional registration application. Figure 1 reflects the starting point for this scenario.

B. Stop criteria

The registration will stop when the difference between the registration error in the current iteration and the one from the previous iteration is less than $0.05mm$. The error is computed as the RMS of the data set and the correspondences established at the iteration k: $e_k = \frac{1}{N_S}\sum_{i=1}^{N_S}\|y_i - s_i\|^2$.

C. Test machine

The machine where we have executed tests has the following specifications: the operating system is Windows 7 Professional (64 bits), the processor is an Intel Core i5 3570K (4 cores, 4 threads, frequency 3.4 Ghz), the motherboard chipset is an Intel Z77, the main memory is composed by two 4 GiB DDR3 modules working with a frequency of 1866 MHz.

The Matlab version that we have used is 8.1 (R2013a).

VI. Experimental results

In this section we will show the obtained results from the experimentation whose methodology was defined in Section V. This result presentation is organized by subsections according to the scenarios described in Section V-A, so that we will show, for each one, the execution time of the registration and its associated error for each one of the metrics proposed in Section IV, and also for each specified noise situation: without noise, with noise applied to both model and data, and with noise only applied to data.

In the same way, we will show a successful registration example for each one of the scenarios, except for the full scenario, whose result is easily imaginable.

A. Full

On Table II we show the experimentation results for the full scenario, there we can observe, regarding execution time, that Chebyshev distance has worse performance than Euclidean one, meanwhile Manhattan distance offers a remarkable improvement in comparison with this last distance. Regarding final registration error, all of them show similar results so neither Chebyshev nor Manhattan get significantly worse quality registrations than the one obtained with the Euclidean distance, except for the case with noise using Chebyshev.

TABLE II
Full scenario results, the execution time is expressed in seconds, and the error (in parentheses) in millimetres.

Noise	Euclid. (s(mm))	Cheby. (s(mm))	Manh. (s(mm))
No	$3.5(6.3 \cdot 10^{-7})$	$9.1(1.1 \cdot 10^{-5})$	$2.9(6.3 \cdot 10^{-7})$
Full	6.8(44.0)	17.0(45.0)	4.8(46.0)
Data	3.4(55.0)	9.8(59.0)	2.9(57.0)

B. Horizontal cut (top)

Table III contains the experimentation results from the top horizontal cut scenario. The information that we can can extract from this data is the same as in the previous scenario: Chebyshev gets worse times meanwhile Manhattan gets an appropriate improvement in this aspect; regarding the registration error, both metrics keep a quality comparable to Euclidean distance except noiseless Chebyshev.

TABLE III
Results of the horizontal cut (top) scenario, the execution time expressed in seconds, and the error in millimetres..

Noise	Euclid. (s(mm))	Cheby. (s(mm))	Manh. (s(mm))
No	$4.6(3.6 \cdot 10^{-12})$	$17.0(8.1 \cdot 10^{-6})$	$4.0(2.9 \cdot 10^{-12})$
Full	6.7(44.0)	23.0(46.0)	5.7(46.0)
Data	4.9(54.0)	17.0(59.0)	4.2(57.0)

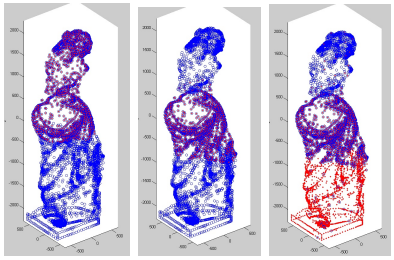


Fig. 3. Left: Successful registration example of top horizontal cut scenario. Middle: Successful registration example of double horizontal cut scenario. Right: Successful registration example of horizontal cut with overlap scenario. (Data set in red, model in blue).

C. Double horizontal cut

In the double horizontal cut scenario, the results shown at Table IV prove that in this case the Euclidean and Manhattan metrics provide a very similar performance and quality, meanwhile Chebyshev one keeps providing significantly worse performance.

ISBN: 978-1-61804-292-7

TABLE IV
RESULTS OF THE DOUBLE HORIZONTAL CUT SCENARIO, THE EXECUTION TIME EXPRESSED IN SECONDS, AND THE ERROR IN MILLIMETRES.

Noise	Euclid. (s(mm))	Cheby. (s(mm))	Manh. (s(mm))
No	$1.6(2.0 \cdot 10^{-5})$	$4.1(2.0 \cdot 10^{-5})$	$1.7(2.0 \cdot 10^{-5})$
Full	2.9(44.0)	6.3(46.0)	2.9(45.0)
Data	2.0(55.0)	3.9(59.0)	1.9(58.0)

D. Horizontal cut (overlap)

At last, data from the horizontal cut with overlap scenario is presented on Table V. It shows again what we have observed in previous scenarios, but in this case the Manhattan metric gets a slightly lower speedup in comparison with the ones obtained in previous scenarios while keeping a very similar quality to the Euclidean metric. Chebyshev metric gets worse performance once again, although registration error remains closer to the original implementation.

TABLE V
RESULTS OF THE HORIZONTAL CUT WITH OVERLAP SCENARIO, THE EXECUTION TIME IN SECONDS, AND THE ERROR IN MILLIMETRES.

Noise	Euclid. (s(mm))	Cheby. (s(mm))	Manh. (s(mm))
No	$2.1(1.1 \cdot 10^{-2})$	$3.9(1.7 \cdot 10^{-2})$	$2.0(1.2 \cdot 10^{-2})$
Full	$2.5(1.3 \cdot 10^{-2})$	$6.0(1.3 \cdot 10^{-2})$	$2.2(1.3 \cdot 10^{-2})$
Data	$2.3(1.3 \cdot 10^{-2})$	$8.0(1.4 \cdot 10^{-2})$	$1.8(1.4 \cdot 10^{-2})$

VII. DISCUSSION

In this secttion we will discuss the results in depth and analyse the conclusions in order to check the possible implications that they may have on our work and the algorithm.

First, it should be noted that the two distance metrics tested showed different results in terms of performance, and quite similar ones regarding the quality of the registration in terms of final error. In those different results, the Chebyshev metric has shown a contradictory behaviour since a performance gain was expected instead of a loss. The Manhattan metric showed the expected behaviour regarding to its performance.

The proposal implemented with the Manhattan metric has shown, in all cases, a better or similar performance than the original implementation with the Euclidean one. Execution time speedups obtained by this metric, calculated with the experimental data of Section VI, are listed by Table VI. In most of the scenarios we can see a performance improvement, a 5% in the worst case and a 29% in the best of them. The average of the computed speedups is about 13%. In some scenarios, such as the double horizontal cut, time results are similar to those obtained by the euclidean metric, so the effect of the metric in them is irrelevant. It is also important to note that greater accelerations have been obtained in those scenarios where a greater number of points on the surfaces were available (full, and top horizontal cut), than those where the figure has more aggressive cuts (double horizontal cut and horizontal cut with overlap).

It is remarkable that in Section IV-B we performed a prediction of a maximum improvement, that was of approximately a -43%, based on empirical data about the execution of C++ implementations of the metrics. In this case we have obtained

TABLE VI
SPEEDUPS COMPUTED FROM TIME DATA PRESENTED ON SECTION VI. THE SPEEDUPS SHOW THE PERFORMANCE GAIN OBTAINED WITH THE MANHATTAN METRIC OVER THE EUCLIDEAN ONE. THE EXECUTION TIME PERCENTAGE OF CHANGE IS SHOWN IN PARENTHESES.

Scenario	Noise	Speedup (Man. over Euc.)
Full	None	1.21 (-17%)
Full	Full	1.42 (-29%)
Full	Data	1.17 (-15%)
Horizontal Cut (Superior)	None	1.15 (-13%)
Horizontal Cut (Superior)	Full	1.18 (-15%)
Horizontal Cut (Superior)	Data	1.17 (-14%)
Double Horizontal Cut	None	0.94 (+6%)
Double Horizontal Cut	Full	1.00 (+0%)
Double Horizontal Cut	Data	1.05 (-5%)
Horizontal Cut (Overlap)	None	1.05 (-5%)
Horizontal Cut (Overlap)	Full	1.14 (-12%)
Horizontal Cut (Overlap)	Data	1.28 (-22%)
Average speedup (Percentage of change)		**1.14 (-13%)**

an improvement of -29% for the noisy top horizontal cut, in the same way that in certain situations such as noiseless horizontal cut with overlap, we have obtained a lower gain than expected with a single -5%; this is due to the same cause as the Chebyshev metric worse performance: changing the distance metric affects the convergence rate because of the quality of the matchings, making it better in some situations where we get a higher gain than the expected due to the lower number of iterations performed, and others where we get a lower convergence speed with more iterations and therefore the obtained performance is lower than the expected.

Despite this fact, it has been experimentally proven that the Manhattan metric offers, in general, better performances in the tested scenarios than the Euclidean one regarding execution time, and at the same time it keeps a similar registration quality, as we show in the final errors listed in Section VI.

VIII. CONCLUSIONS AND FUTURE WORK

In this article, we have presented an improvement proposal for the ICP algorithm which able to reduce its computational cost effectively for high resolution applications. This cost reduction affects the execution time of the algorithm, making it decrease significantly in different scenarios. The performance improvement is due to the replacement of the Euclidean distance metric with another one with less computational cost, such as the Manhattan distance metric. First, we made a brief analysis about the viability of the improvement using the Chebyshev and Manhattan metrics execution times, and comparing them to the Euclidean metric ones; in that analysis we proved, based on the obtained experimental results, that the Chebyshev metric reduced the execution time by a 20%, while Manhattan did it by a 43%.

Since the metric is intensively used during the matching phase to find the closest neighbour for each point, and taking into account that the matching phase represents a significant part of the execution time of the algorithm, a positive impact is caused on performance by using these low cost metrics. The performance improvement obtained by using this modification of the algorithm with the Manhattan distance ranges from 5%

ISBN: 978-1-61804-292-7

to 29% time reduction on tested scenarios, having a mean execution time reduction of a 13%. In some situations with low computational load, the proposal performs roughly at the same level than the original algorithm, with variations which range from -5% (execution time decrease) to 6% (time increase). The speedup is boosted in scenarios with a higher computational load, so that the proposal is quite useful when dealing with high density point sets in high precision applications. Additionally, the registration quality is kept at the same levels as the ones obtained with the Eucliden distance, with variations which, in some scenarios, were slightly better for Manhattan and others for the original implementation. Since fine registration is required in many real-time applications which need a quick and robust response, our proposal has a high impact potential.

Finally, the proposed contribution can be extended or improved in several ways: the same concept of cost reduction could be applied in parallel architectures, both in CPU and GPU to explore both possibilities to increase performance. It is also possible to modify variants where the matching phase has not been altered, in order to apply this performance increase to other contributions that improve other parts of the algorithm. Furthermore, a custom low cost distance metric might be created to test its effect over the algorithm.

REFERENCES

[1] Brian Amberg, Sami Romdhani, and Thomas Vetter. Optimal step nonrigid icp algorithms for surface registration. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE, 2007.

[2] Leopoldo Armesto, J. Minguez, and L. Montesano. A generalization of the metric-based iterative closest point technique for 3d scan matching. In *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, pages 1367–1372, May 2010.

[3] Michel A. Audette, Frank P. Ferrie, and Terry M. Peters. An algorithmic overview of surface registration techniques for medical imaging. *Medical Image Analysis*, 4(3):201 – 217, 2000.

[4] P.J. Besl and Neil D. McKay. A method for registration of 3-d shapes. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 14(2):239–256, Feb 1992.

[5] Y Chen and G Medioni. Object modeling by registration of multiple range images. In *Robotics and Automation, 1991. Proceedings., 1991 IEEE International Conference on*, pages 2724–2729. IEEE, 1991.

[6] Haili Chui and Anand Rangarajan. A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 89(2):114–141, 2003.

[7] Benoît Combès and Sylvain Prima. Prior affinity measures on matches for icp-like nonlinear registration of free-form surfaces. In *Proceedings of the Sixth IEEE International Conference on Symposium on Biomedical Imaging: From Nano to Macro*, ISBI'09, pages 370–373, Piscataway, NJ, USA, 2009. IEEE Press.

[8] Chitra Dorai, Gang Wang, Anil K Jain, and Carolyn Mercer. Registration and integration of multiple object views for 3d model construction. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 20(1):83–89, 1998.

[9] Jacques Feldmar and Nicholas Ayache. Rigid, affine and locally affine registration of free-form surfaces. *International journal of computer vision*, 18(2):99–119, 1996.

[10] Sébastien Granger, Xavier Pennec, and Alexis Roche. Rigid point-surface registration using an em variant of icp for computer guided oral implantology. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2001*, pages 752–761. Springer, 2001.

[11] Alfred Gray, Elsa Abbena, and Simon Salamon. Modern differential geometry of curves and surfaces with mathematica. *Boca Raton, FL*, pages 373–380, 1997.

[12] Mads Fogtmann Hansen, Morten Rufus Blas, and Rasmus Larsen. Mahalanobis distance based iterative closest point. In *Medical Imaging*, pages 65121Y–65121Y. International Society for Optics and Photonics, 2007.

[13] H. Mora Mora JM. García Chamizo, J. Mora Pascual and MT. Signes Pont. Calculation methodology for flexible arithmetic processing. In *IFIP/IEEE International Conference on Very Large Scale Integration*, pages 350–355, 2003.

[14] C. Langis, M. Greenspan, and G. Godin. The parallel iterative closest point algorithm. In *3-D Digital Imaging and Modeling, 2001. Proceedings. Third International Conference on*, pages 195–202, 2001.

[15] Wang Liying and Song WeiDong. A review of range image registration methods with accuracy evaluation. In *Urban Remote Sensing Event, 2009 Joint*, pages 1–8, May 2009.

[16] Higinio Mora-Mora, Jerónimo Mora-Pascual, JuanManuel García-Chamizo, and Antonio Jimeno-Morenilla. Real-time arithmetic unit. *Real-Time Systems*, 34(1):53–79, 2006.

[17] Chad Mourning, Scott Nykl, Huihui Xu, David Chelberg, and Jundong Liu. Gpu acceleration of robust point matching. In *Advances in Visual Computing*, volume 6455 of *Lecture Notes in Computer Science*, pages 417–426. Springer Berlin Heidelberg, 2010.

[18] Daniel Münch, Benoît Combès, and Sylvain Prima. A modified icp algorithm for normal-guided surface registration. In *SPIE Medical Imaging*, pages 76231A–76231A. International Society for Optics and Photonics, 2010.

[19] Department of Computer Science and National University of Taiwan Information Engineering. 3d models. October 2005.

[20] Soon-Yong Park, Sung-In Choi, Jaekyoung Moon, Joon Kim, and Yong-Woon Park. Real-time 3d registration of stereo-vision based range images using gpu. In *Applications of Computer Vision (WACV), 2009 Workshop on*, pages 1–6, Dec 2009.

[21] Xavier Pennec and Jean-Philippe Thirion. A framework for uncertainty and validation of 3-d registration methods based on points and frames. *International Journal of Computer Vision*, 25(3):203–229, 1997.

[22] K. Pulli. Multiview registration for large data sets. In *3-D Digital Imaging and Modeling, 1999. Proceedings. Second International Conference on*, pages 160–168, 1999.

[23] David P. Rodgers. Improvements in multiprocessor system design. *SIGARCH Comput. Archit. News*, 13(3):225–231, June 1985.

[24] S. Rusinkiewicz and M. Levoy. Efficient variants of the icp algorithm. In *3-D Digital Imaging and Modeling, 2001. Proceedings. Third International Conference on*, pages 145–152, 2001.

[25] Joaquim Salvi, Carles Matabosch, David Fofi, and Josep Forest. A review of recent range image registration methods with accuracy evaluation. *Image and Vision Computing*, 25(5):578–596, 2007.

[26] María Teresa Signes, Juan Manuel García, and Higinio Mora. Improvement of the discrete cosine transform calculation by means of a recursive method. *Mathematical and Computer Modelling*, 50(5–6):750 – 764, 2009. Mathematical Models in Medicine and Engineering.

[27] David A. Simon. *Fast and Accurate Shape-based Registration*. PhD thesis, Pittsburgh, PA, USA, 1996. AAI9838226.

[28] G.K.L. Tam, Zhi-Quan Cheng, Yu-Kun Lai, F.C. Langbein, Yonghuai Liu, D. Marshall, R.R. Martin, Xian-Fang Sun, and P.L. Rosin. Registration of 3d point clouds and meshes: A survey from rigid to nonrigid. *Visualization and Computer Graphics, IEEE Transactions on*, 19(7):1199–1217, July 2013.

[29] Oliver Van Kaick, Hao Zhang, Ghassan Hamarneh, and Daniel Cohen-Or. A survey on shape correspondence. In *Computer Graphics Forum*, volume 30, pages 1681–1707. Wiley Online Library, 2011.

[30] S. Weik. Registration of 3-d partial surface models using luminance and depth information. In *3-D Digital Imaging and Modeling, 1997. Proceedings., International Conference on Recent Advances in*, pages 93–100, May 1997.

[31] Aviad Zabatani and Alex M. Bronstein. Parallelized algorithms for rigid surface alignment on gpu. In *Proceedings of the 5th Eurographics Conference on 3D Object Retrieval*, EG 3DOR'12, pages 17–23, Aire-la-Ville, Switzerland, Switzerland, 2012. Eurographics Association.

ISBN: 978-1-61804-292-7

A Parallel Implementation for the Time-Domain Analysis of a Rectangular Reflector Antenna using OpenMP

Ghada M. Sami[1,3] and Khaled Ragab[1,2]
[1]Math. Department, Faculty of Science, Ain Shams University, Cairo, Egypt
[2]College of Computer Sciences and Info. Technology, King Faisal University, Saudi Arabia
[3]Math. Department, College of Science, King Faisal University, Saudi Arabia 1 Department of Electrical

Abstract ─ This paper presents and evaluates a parallel time domain analysis of a rectangular reflector on multicores machine. Rectangular reflector antennas have motivated the time-domain analysis of electromagnetic scattering problems. The asymptotic time domain physical-optics (TDPO) is applied to the analysis of a rectangular reflector illuminated by a Gaussian-impulse. It is a numerical technique used in computational electrodynamics. The effects of time-delayed mutual coupling between points on the surface will be ignored because of utilizing the TDPO method for determining the equivalent surface-current density on the reflector. As a result, the scattered signals at the specular reflection point, at the edges, and at the corners can be clearly distinguished. Furthermore, this paper evaluates and compares the performance of the sequential time-domain analysis against the parallel time-domain analysis on multicores machine.

Index Terms - Parallel computing, Time domain; Rectangular reflector; Electromagnetic scattering.

I. INTRODUCTION

Reflector antennas are intensively applied in the radars, communication, and guidance, etc. Nowadays, the problems of electromagnetic (*EM*) scattering have been widely applied in fields of remote sensing, target identification, radar detection, and so on. The interest in the transient analysis of *EM* phenomena has been growing in recent year. This is due to the advance of Ultra-Wide Band (*UWB*) radars and their associated antennas, various antennas have been proposed for *UWB* application [1], with mobile radio channels by means of their response to pulsed excitation [2]. There are several methods that are used to analysis the *EM* scattering that will be explored in next section. They have inherent difficulties with numerical instability, interpolation errors, and need of extensive computer memory and *CPU* time to solve problems involving large scatterers. It is more efficient do deal with the transient analysis directly in the time domain. The time domain physical-optics (*TDPO*) [5], [6] is an alternative method that requires relatively small amounts of computer memory and CPU time.

Consequently, this paper will focus on implementing the *TDPO* approximation method on parallel computer system. However, this section will discuss several *CEM* numerical methods either implemented in sequential or in parallel as follows.

Physical-optics (PO) approximation is one of these techniques. It has been widely used and considered as a good approximation of the far field electromagnetic scattering [16]. Starting from the *Stratton-Chu* integral equations, the *PO* expressions can be obtained for the *PO* scattered magnetic field in frequency domain [17]. The *PO* approximation is initially applied in the frequency-domain with the inverse Fourier transform [9] and [10]. Those equations are obtained directly from *Maxwell's equations* by applying *Green's theorem* in its vector form [17, 18]. The *PO* requires integration over the illuminated surface of the scatterer. Due to the complex exponential term, the integrand of the *PO* integral is a very oscillatory function, especially at high frequencies. Therefore, it is very expensive to compute these kinds of integrals by simple numerical integration techniques such as Levin's integration method [19]. For large scatterer, the *PO* approximation is an efficient method in the frequency domain [7], [8]. To accelerate the computing of the *PO*, there are some researches that handle *PO* in parallel based shared memory [28] and distributed memory [29, 30].

Moreover, there exist several analytic and numerical techniques for obtaining the response of scattering problems directly in the time domain, which is the most natural approach to be used, such as the finite-difference time-domain method (*FDTD*) [3, 11, 12]. Recently, the *FDTD* method is being used to solve a wide variety of practical problems, because it can be competitive with the *FEM* in terms of versatility and solve time, even on a single PC or laptop computer loaded with a 2 GB memory. However, the main advantage of the *FDTD* becomes increasingly apparent when it is run either on multi-core processors or MPI protocol with low-cost high speed networks, because it

ISBN: 978-1-61804-292-7

can be parallelized more efficiently than the *FEM* [21-25]. There are several works [26, 27] present a hybrid *FDTD* numerical algorithm which has been successfully developed and validated. They employ distributed and shared memory thorough of *MPI* and *OpenMP* [14].

II. THEORY AND FORMULATION

The *TDPO* integral is evaluated over the illuminated with a closed-form expression based on Gaussian-impulse. The formula of the *TDPO* is derived with the inverse Fourier transform. The scattered field of the *TDPO* is obtained as follows [20]:

$$\bar{e}^{TDPO}(\bar{r},t) = -\eta_o \iint_{s`} \frac{1}{4\pi|\bar{r}-\bar{r}`|} \frac{\partial\, \bar{j}_{st}^{PO}(\bar{r}`,\tau(t,\bar{r}`))}{\partial\, t}\, ds` \quad (1)$$

where $\bar{j}_{st}^{PO}$ is given as:

$$\bar{j}_{st}^{PO}(\bar{r}`,\tau(t,\bar{r}`)) =$$

$$\bar{j}_{s}^{PO}(\bar{r}`,\tau(t,\bar{r}`)) - [(\bar{j}_{s}^{PO}(\bar{r}`,\tau(t,\bar{r}`)))\cdot\hat{r}]\cdot\hat{r}\,, \quad (2)$$

$$\bar{j}_{s}^{PO}(\bar{r}`,\tau(t,\bar{r}`)) \;=\; 2\,\hat{n}\times\bar{h}^{inc}(\bar{r}`,\tau(t,\bar{r}`))\,. \quad (3)$$

where the vector ***r`*** locates the integration point on the scatterer surface, ***r*** is the distant observing point, ***c*** is the velocity of the light and is η_o the intrinsic free space impedance,
$\bar{j}_{s}^{PO}(\bar{r}`,\tau(t,\bar{r}`))$ is the surface-current distribution in the time domain and $\bar{h}^{inc}(\bar{r}`,\tau(t,\bar{r}`))$ is the time-domain magnetic field incident on the surface.
The delay time of the propagation is given by:

$$\tau(t,\bar{r}`) = t - \frac{|\bar{r}-\bar{r}`|}{c}. \quad (4)$$

Based on equation (1) the surface-current density does not need to be solved. Consequently minimum computer memory is required and no interpolation evaluation needs to be carried out because the incident fields are known for all positions and times. This benefit makes this approach suitable for limited computer-memory requirement (e.g. personal computer).

Fig. 1 shows the geometry of a rectangular reflector illuminated by an incident wave. We assume that incident wave is bandpass Gaussian-pulse transmit from x-polarized small dipole point source, which has the following form:

$$\bar{h}^{inc}(\bar{r}_k`,t) = \frac{B}{\sqrt{2\pi}\,|\bar{r}_k`|\,\sigma}\,\mathrm{Exp}[-((t-\frac{|\bar{r}_k`|}{c})/\sigma)^2]\;, \quad (5)$$
$$[2\,\mathrm{Cos}[\omega_o\,(t-\frac{|\bar{r}_k`|}{c})]]\,\mathrm{Sin}[\theta_x]\,\hat{\varphi}_x$$

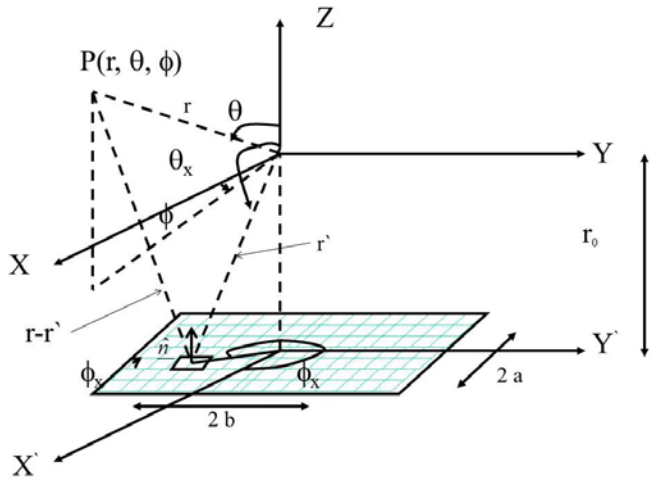


Fig.1. Geometry of a rectangular reflector

σ is the standard deviation of Gaussian envelope, *B* is the magnitude parameter of impulse, and ω_o is the angular frequency.
From Eq.(5) we can form the time – domain representation ***v(t):***

$$\boldsymbol{v(t)} = \mathrm{Exp}[-((t-\frac{|\bar{r}_k`|}{c})/\sigma)^2]\,[\,\mathrm{Cos}\,[\omega_o\,(t-\frac{|\bar{r}_k`|}{c})]]\,, \quad (6)$$

as a real signal, we can write v(t) as:

$$\boldsymbol{v(t)} = \mathrm{Re}\,[\mathrm{Exp}[-((t-\frac{|\bar{r}_k`|}{c})/\sigma)^2]\,[\mathrm{Exp}\,[j\,\omega_o\,(t-\frac{|\bar{r}_k`|}{c})]]]\,,$$

$$= \mathrm{Re}[F\,(t)\,\mathrm{Exp}\,(j\omega_o\,t)]\,, \quad (7)$$

where ***F(t)*** is analytic low pass input signal,

F(t) = I – j Q, where

$$I = \mathrm{Exp}[-((t-\frac{|\bar{r}_k`|}{c})/\sigma)^2]\;\mathrm{Cos}\,[\omega_o\,\frac{|\bar{r}_k`|}{c}],\ \text{and}$$

$$Q = \mathrm{Exp}[-((t-\frac{|\bar{r}_k`|}{c})/\sigma)^2]\;\mathrm{Sin}\,[\omega_o\,\frac{|\bar{r}_k`|}{c}],$$

where ***I*** and ***Q*** are the *In-phase* and *Quadrate* parts. ***F(t)*** corresponds to the complex envelope of ***v(t)*** and useful to know the intensity of the scattered wave in time domain.

The next step is being able to show how such a bandpass system can be given an equivalent baseband representation at the center frequency, as

$$U(t) = \frac{1}{2\pi T} \int_{t-\frac{T}{2}}^{t+\frac{T}{2}} u(\tau)\, \mathrm{Exp}[-j\omega_o \tau]\, d\tau \qquad (8)$$

The baseband output is the sum over each path, of the delayed replicas the baseband input. When we get the U (t), it is possible to draw dB plot, as shown in Fig. 3-b.

III. PARALLEL IMPLEMENTATIONS OF THE TDPO

In this paper, the *OpenMP* [14] programming interface was employed to parallelize the computations of the *EM* based on the *TDPO* method. It was developed on the multicore central processing unit (CPU) in multiple precisions arithmetic. *OpenMP* has been used to parallelize the code and memory-hierarchy-based optimization techniques to reduce the computer time of the code. Using these techniques, the computer time can be reduced in a factor close to the number of cores of the CPU. While acceleration of the computational electromagnetic methods on graphics processing units (GPUs) has recently become a hot topic of investigations, multicore CPU still remains a source of significant computational power comparable to the GPU throughput for specific algorithms [15]. To the best of our knowledge, accurate computation of scattered field of the *TDPO* over rectangular reflector illuminated by a Gaussian-impulse for rectangular require the multiple precision arithmetic, which has not been implemented as a library on GPUs yet. Therefore, it can be anticipated, that the proposed parallel CPU implementation will open the door to the implementation of the *TDPO* method on heterogeneous computing systems simultaneously deploying the computational power of multicore CPUs and GPUs for the tasks best suited for each.

In this paper the authors implemented their computing solution on parallel using *OpenMP* as follows. It begins with a single thread of control, called the *master thread*, which exists for the duration of the program. The set of variables available to any particular thread is called the *thread's execution context*. During execution, the master thread may encounter parallel regions, at which the master thread will fork new threads, each with its own stack and execution context. At the end of the parallel region, the forked threads will terminate, and the master thread continues execution.

The master thread performs the following steps to compute scattered field of the *TDPO*:

- The geometry parameters of the rectangular reflector antennas are input to master thread. In addition, it initializes the constant values, uses equation 3 to calculate surface current density $\bar{j}_s^{PO}$ and then calculates the incident wave $\bar{h}^{inc}(\bar{r}_k^{`}, t)$ that is shown in equation 5,
- Calculate the scattered field of the *TDPO* equation $\bar{e}^{TDPO}(\bar{r}, t)$ by transforming the integral in equation 1 into sum of scatter fields over *M x N* small rectangular reflectors. The master thread creates NTHREADS worker threads where each one calculates the sum over a small rectangle reflector.

The computing of the scattered field of the *TDPO* over rectangular reflector is obtained by equation 1. The rectangular reflector can be divided into *M x N* small rectangular reflectors as shown in fig. 1, where $M=2a/0.1\lambda$, $N=2b/0.1\lambda$, and λ denotes the wavelength. Consequently, the integration in equation 1 can be expressed as the summation of scattered fields over these *M x N* small rectangular reflectors. Therefore, it is time-consuming calculations that need to be performed on parallel. At this time, the code has been parallelized by distributing the *M* vectors of rectangular reflectors into NTHREADS threads. Each thread calculates the sum of the scattered field over *N* rectangular reflectors. The sum total of all the scatter fields by *M x N* small rectangular reflectors constitutes the scattered field by the target as follows.

```
#pragma omp parallel shared () private () {
        #pragma omp for schedule (static)
  for ( i=-M ; i<=M; i++)
     for (j= -N; j<=N; j++)
        sum =sum + Compute_ScatteredField(i, j);
   }
```

The main problem that appears in using *OpenMP* is the order of the loops. If the directions of the observation points are used as outer loop, then each core can compute the scattered field created by all the rectangular in one direction, and at the end, it should store the result in a position of the output vector. But if the index of the rectangular is used as the outer loop, then each core must compute the scatter field over this small rectangular and then use the reduction method to add all the results. Unfortunately, the reduction method is not well implemented for vectors in *OpenMP*, and each core must wait for the others to write their results.

IV. NUMERICAL RESULTS AND DISCUSSION

To explore the effectiveness of the used parallel technique, this paper implemented and carried out sequential and parallel experiments to examine the processing time needed to compute of the scattered field of the *TDPO* over rectangular reflector.

ISBN: 978-1-61804-292-7

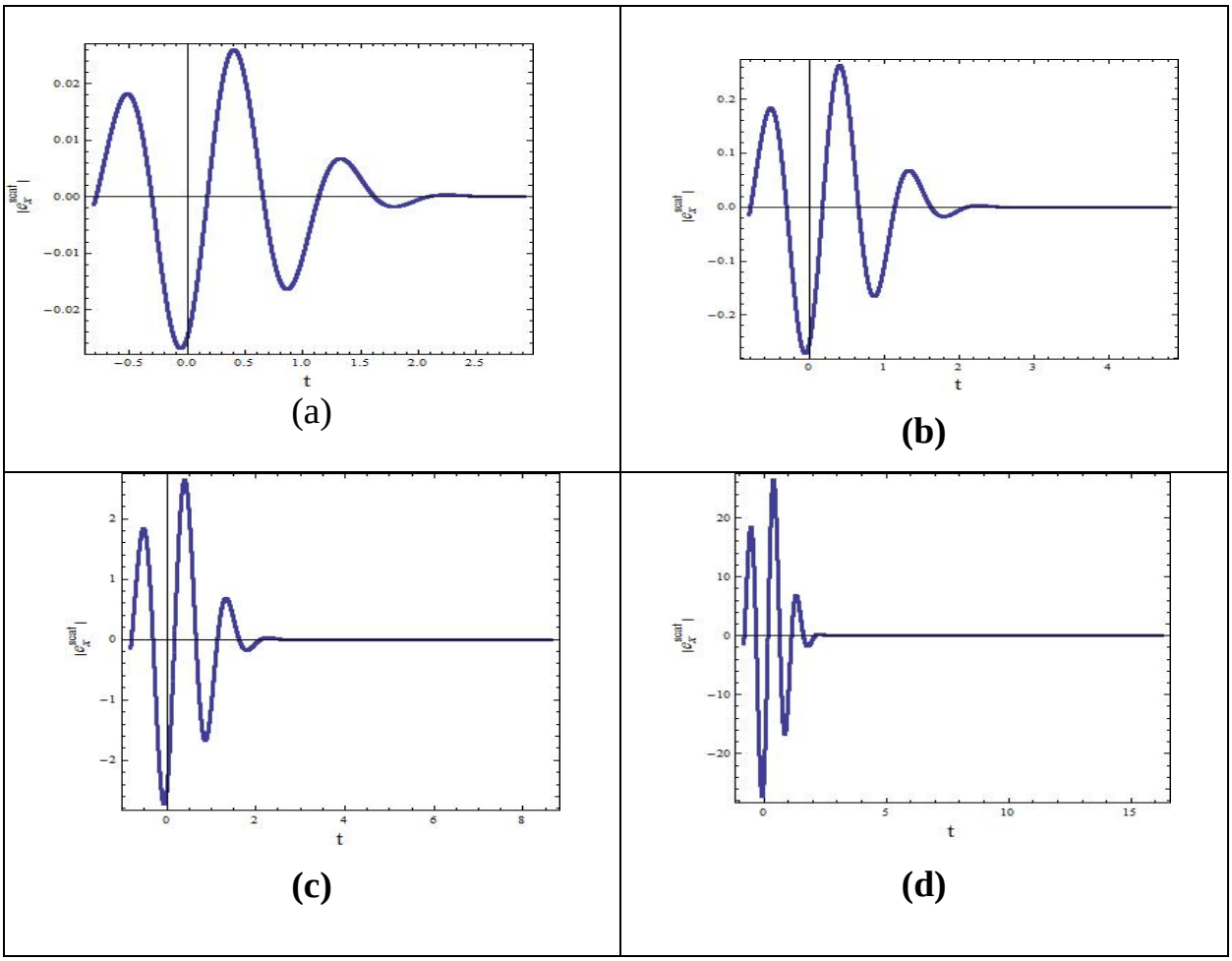


Fig. 2 (a-d). Scattered field rectangular reflector with Gaussian-impulse excitation TDPO with reflector diameters, *d, 2d, 4d* and *8d***,** respectively

A. Setup

Theses computing algorithms were implemented using *Microsoft Visual Studio Professional 2012* on a *HP* server (*ProLiant* ML350p Gen8) with two *Intel Xeon (R)* processors (*E5-2620 @ 2.00 GHz*), each processor has *6* cores and *32* GB RAM. The total number of physical cores is *12*. Hence, it is capable of running *12* threads simultaneously. The multicore *CPU* implementation was performed using the *OpenMP* programming model as in [14, 31].

B. Numerical Results

Numerical results were obtained for a variety of configurations. As a target, we use a PEC rectangular plate as shown in Fig. 1, where λ is the wave length, σ is the standard deviation of Gaussian-impulse and $\tau = \frac{t - t_o}{\sigma}$ and $d = c\sigma$ is the reflector diameter.

Fig. 2 (a-d) shows the scattered field of the *TDPO* of an exact solution. The three scattering components shall be distinct, i.e. specular reflection at the center of rectangle, edge diffraction at the center of the edge, and corner diffraction at the corners shown in Fig. 2 (a-d), respectively. In Fig. 2 (a-d), the reflectors diameters are *d, 2d, 4d* and *8d*, respectively. The scattered field *TDPO* increases with increasing reflectors diameter *d* by factors 2, 4, and 8. The results appear to be more accurate and stable faster than those obtained by frequency domain physical optics [20]. For greater reflector size, the time domain solution requires considerably more computing power consequently we implemented it in parallel.

Fig. 3-a, shows that the observation point is very close to the reflector shadow boundary associated with upper diffraction point, Gaussian-impulse excitation with coordinates r = 100 m, $\theta = 65^o, \phi = 0^o$. Radiation pattern for Gaussian-impulse excitation, based in the peak response at the three scattering components is plotted in Fig. 3-b.

ISBN: 978-1-61804-292-7

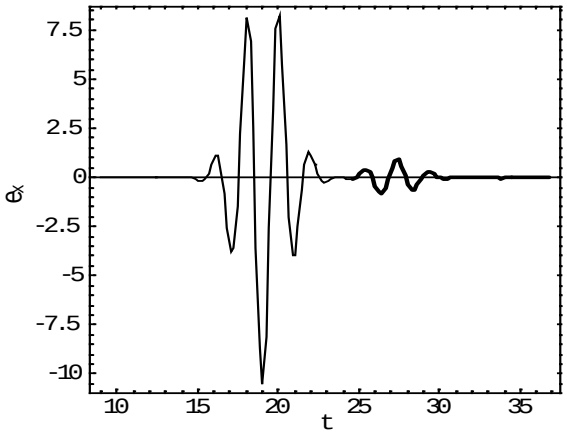


Fig. 3-a Scattered field of a rectangular reflector with Gaussian-impulse excitation at r = 100 m, $\theta = 65^o, \phi = 0^o$

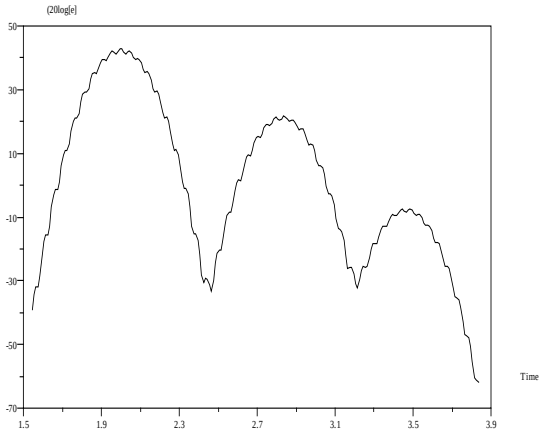


Fig. 3-b. Radiation pattern for Gaussian-impulse excitation.

To compute the scattered field of the *TDPO* over rectangular reflector with diameter *10d* the single-threaded code requires *177.422* seconds. While the multi-threaded code with *12* threads requires *19.73* seconds. The merits of the parallel computing are speedup S_l and efficiency E_l using l parallel threads that can be computed as follows [32], $S_l = T_{sequential} / T_l$ and $E_l = T_{sequntial}/(l \,.\, T_l)$ where $T_{equential}$ is the computing time in sequential, T_l is the computing time using l threads and $1 \leq S_l \leq l$. However, the computing overhead is determined as follows $O(l)=T_l\,(1-E_l)= T_l - (T_{equential}\,/l)$. This experiment shows that with 12 threads the computing is speedup by 8.99x and efficiency is 75%. Fig. 4.a shows the plot showing the speedup as a function of reflector diameters (*d, 2d, 4d, 6d* and *8d* respectively).

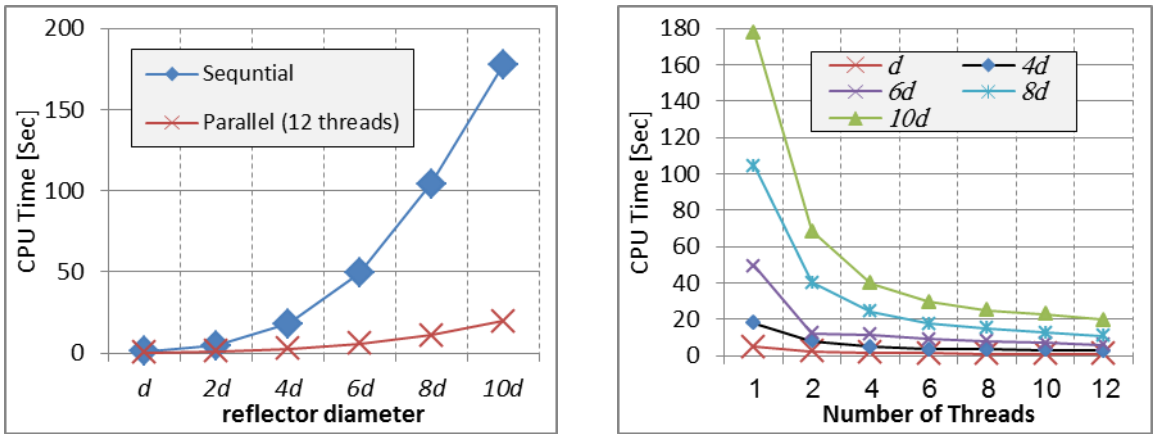
The code is multi-threaded that achieves an excellent speedup when executed on multiple cores. Fig. 4.b demonstrates the required computing time according to different reflector diameters along with increasing the number of parallel threads. This figure shows that for small wavelength the effect of parallel has less significant however it shows significant impact while increasing the reflector's diameter. Moreover, we extend our experiments to reflector diameter 25*d* and we are able to calculate the scatter field in sequential within *12* hours and *14* minutes however it take one hour and *35* minutes and *20* seconds with carrying out *12* threads per *12* cores. This experiment shows that with *12* threads the computing is speedup by 8.02x and efficiency around %.65.

IV. CONCLUSION

To determine the analysis of a rectangular reflector illuminated by a *Gaussian-impulse* considering the *UWB* radar application, this work extends the concept of the frequency-domain physical optics approximation to time-domain. The scattered field of the *TDPO* is obtained by performing the inverse *Fourier transform* over the frequency-domain scattered field that is obtained by calculating the integral over the illuminated surface using the free space *Green's* function. The numerical results show the applicability of *TDPO*, as the scattered signals at the specular reflection point, edge diffraction and corner diffraction. Fig. 2(a-d) shows comparisons of the *TDPO* results with a reference solution based on a frequency domain physical optics. The frequency domain physical optics solution requires considerably more computer time and becomes inherently unstable. Moreover, the *TDPO* can reduce CPU time drastically. The parallel implementation of the *TDPO* is developed over multicores using *OpenMP*. The parallel performance of the parallel *TDPO* program is measured. And the results show that the speed up ratio is approximately equal to *8.99x* with 12 threads.

REFERENCES

[1] C. E. Baum and E. G. Farr, Impulse radiating antennas, Ultra-wideband short-pulse electromagnetic book, edited by H. Bertoni et al., Plenum press, 1993.



(a) (b)

Fig. 4. Speeding up the computing the scattered field of the TDPO over rectangular reflector.

[2] W. Zhang, A wind- band propagation model based on UTD for cellular mobile communication, IEEE Transaction on antennas and propagation, vol. 45, pp.1669-1678, November 1997.

[3] A**.** Tirkas and A. Balanis, Finite-Difference Time-Domain Method for Antenna Radiation, IEEE Trans. On Antenna and Propagation, vol. 40, no. 3, march 1992.

[4] P. R. Rosseau and P. U. Phatak, Time-domain uniform geometrical theory of diffraction for a curved wedge, *IEEE Transactions on Antennas and Propagation***,** vol. 43, pp. 1375-1382, December 1995.

[5] G. Cassis Rego, J. V. Flavio Hasselmann, and Fernando J. S. Moreira, Time-Domain Analysis of a Reflector Anttena Illuminated by a Gaussian Pulse, Journal of Microwave and Optoelectronic, vol. 1, No. 4, Sept. 1999.

[6] L.-X. Yang, D.-B. Ge, and B. Wei, FDTD/TDPO Hybrid Approach for Analysis of the EM Scattering of Combinative Objects, *Progress In Electromagnetics Research*, PIER 76, 275-284, 2007.

[7] W. V. T. Rusch and P. D. Potter, Analysis of Reflector Antenna, New York: Academic, pp. 46-49, 1970.

[8] R. F. Harrington, Time-Harmonic Electromagnetic Fields, New York: McGraw-Hill, 1961, p. 127.

[9] E. M. Kennaugh and R. L. Cosgriff, The Use of Impulse Response in Electromagnetic Scattering Problems, IRE Natl. Conv. Rec., pt. 1, pp. 72 - 77, 1958.

[10] E. M. Kennaugh and D. L. Moffatt, Transient and Impulse Response Approximation, Proc. IEEE, pp. 893-901, Aug., 1965.

[11] K. Vee, Numerical Solution of Initial Boundary Value Problems Involving Maxwell's Equations in Isotropic Media, IEEE Transactions on Antennas and Propagation, Vol. 14, No.5, May 1966, pp. 302-307.

[12] A. Taflove and S. Hagness, Computational Electromagnetics: The Finite-Difference TimeDomain Method, 3rd ed., Artech House, Norwood, MA, 2005.

[13] W. Yu, Raj Mittra, Tao Su, Yongjun Liu, and Xiaoling Yang, *Parallel Finite Difference Time Domain Method*, Artech House, Norwood, MA, 2006.

[14] OpenMP Architecture Review Board, OpenMP application program interface (Version 3.1), Jul. 2011 [Online]. Available: http://www.openmp.org

[15] V. W. Lee *et al.*, Debunking the 100X GPU vs. CPU myth: An evaluation of throughput computing on CPU and GPU, *Proc. 37th Annu. Int. Symp. Comput. Archit.*, 2010, pp. 451–460.

[16] John S. Asvestas, The physical optics method in electromagnetic scattering, Journal of Math. Phys., Vol. 21, No.2, February 1980, pp.0 290 – 299

[17] J. A. Stratton, Electromagnetic Theory, McGraw-Hill, 1941

[18] R. Mittra, Computer Techniques of Electromagnetics, Pergamon Press 1973

[19] **A**. C. Durgun and M. Kuzuoglu , Computation of Physical Optics Integral by LEVIN'S Integration Algorithm, Progress In Electromagnetics Research M, Vol. 6, 59–74, 2009

[20] E.-Y. Sun and W. V. T. Rusch, Time-domain physical-optics, *IEEE Transactions on Antennas and Propagation*, vol. 42, no. 1, pp. 9–15, 1994.

[21] A. Taflove and S. Hagness, *Computational Electromagnetics: The Finite-Difference TimeDomain Method*, 3rd ed., Artech House, Norwood, MA, 2005.

[22] W. Yu, Raj Mittra, Tao Su, Yongjun Liu, and Xiaoling Yang, *Parallel Finite Difference Time Domain Method*, Artech House, Norwood, MA, 2006.

[23] Wenhua Yu, X. Yang, Y. Liu, Lai-Ching Ma, T. Su, N. Huang and R. Mittra, *New Direction in Computational Electromagnetics Solving Large Problems Using the Parallel FDTD on the BlueGene/L Supercomputer Yielding Teraflop-Level Performance*,

IEEE Antennas and Propagation Magazine, , Vol. 50, Apri1 2008, No.23, pp. 20-42.
[24] Wenhua Yu, Raj Mittra, Xiaoling Yang, and Yongjun Liu, Performance Analysis of Parallel FDTD Algorithm on Different Hardware Platforms, Antennas and Propagation Society International Symposium, pp. 1-4, 2009.
[25] Tomasz P. Stefański, Fast Implementation of FDTD-Compatible Green's Function on Multicore Processor, IEEE Antennas and Wireless Propagation Letters, Vol. 11, pp. 81-82, 2012.
[26] LAKSONO ADHIANTO AND BARBARA CHAPMAN, Performance Modeling of Communication and Computation in Hybrid MPI and OpenMP Applications, Proc. of the 12th Int. Conf. on Parallel and Distributed Systems (ICPADS'06) IEE, 2006.
[27] ROBERT ROSENBERG GUY NORTON JORGE C. NOVARINI, Modeling Pulse Propagation and Scattering in a Dispersive Medium: Performance of MPI/OpenMP Hybrid Code, IEEE Conference (SC'06) 2006.
[28] Marcos Arias-Acuña, et. al., Fast Far Field Computation of Single and Dual Reflector Antennas, Journal of Engineering, Hindawi Publishing, Vol. 2013.
[29] Imbriale, W. A. & Cwik, T., A simple physical optics algorithm perfect for parallel computing architecture, In 10th Annual Review of Progress in Appl. Comp. Electromag., pp. 434-441, Monterey, Cal., 1994.
[30] Christian Parrot, et. al., A Distributed Memory Multilevel Fast Physical Optics Algorithm, 3rd European Conference on Antennas and Propagation, Germany, 2009.
[31] Peter Pacheco, An Introduction to Parallel Programming, Morgan Kaufmann, 1st edition, 2011.
[32] Clyde P. Kruskal, et. al., A complexity theory of efficient parallel algorithms, Automata, Languages and Programming Lecture Notes in Computer Science Volume 317, pp. 333-346, 1988.
[33] Rousseau, P.R., et.al., A Time Domain Formulation of the Uniform Geometrical Theory of Diffraction for Scattering From a Smooth Convex Surface, IEEE Transactions on Antennas and Propagation, Vol. 55, No. 6, pp. 1522 – 1534, 2007.

Asset Risk Diversity and Portfolio Optimization with Genetic Algorithm

Jinchuan Ke, Yi Yu, Biyao Yan, and Ying Ren

***Abstract*—**The genetic algorithm has a wide range of search capability, showing a multi-objective optimization problem solving strengths. In this paper, we use the balance of return and risk multi-objective optimization theory based on the Harry M. Markowitz mean variance model, introducing the risk preference as well as VaR constraints in Genetic Algorithm. Through the MATLAB simulation, the stock portfolio is optimized for the investors to make decision.

***Keywords*—**Genetic algorithms, optimization**,** portfolio, stock

I. Introduction

THE most common stock selection and the most important points for us to consider is the degree of return and risk: the return of investment income determines how much risk we determine to bear when we invest. Highly profitable often associated with high risk investments, so investors should not only pursue the high profits and ignore the risk. Both of which should be considered in conjunction with their own situation to be weighed [1-2].

From the theoretical model improvements, the latest research results include the introduction of portfolio theory transaction costs; long-term investment portfolio theory continuous time; VaR portfolio theory; behavioral portfolio theory; non-utility maximization portfolio theory. From the aspect of calculation methods in recent years, there are a lot of approaches different to traditional optimization algorithms, such as genetic algorithms, neural networks, heuristic algorithm, simulated degradation algorithm [3-6].

This paper intends to use the genetic algorithm to measure the portfolio by consideration of the rate of earning and the standard deviation. Pursuant to constitute a multi-objective function of the problem, we use genetic algorithm for optimization of the solution to get the stock of non-dominated solutions.

II. Genetic Algorithm

The basic principle of the genetic algorithm is applied by a gene on the chromosome [7]. Chromosome can solve the problem and find a good one. It requires each chromosome generated by the algorithm to be evaluated and selected based on the fitness of the chromosome, so that the good adaptability chromosome can get more chance to reproduce. In genetic algorithm, the random number is generated by the digital code, namely the chromosome form the initial population. Through the fitness function for each individual with numerical evaluation, the low fitness individuals are eliminated and the high fitness individuals are chosen to participate in genetic manipulation. The function

$$GA = (P(0), N, l, s, g, p, f, t) \quad (1)$$

Where P (0) = ((0), (0), ⋯ , (0)) represents the initial population; N represents the number of individual populations contain; l is the length of the binary string representation; S indicates the selection strategy; g indicates genetic operators, usually including breeding operator: I × I, crossover: I × I, and mutation operators: I × I; p indicates the probability of the child's genetic operator operations, including breeding probability, crossover probability and mutation probability; f is the fitness function; t is the termination criteria. The genetic algorithm is as follows [8]:

(1) Initialization: Since the implementation of population-based genetic algorithm operation, it must be a part of the solution space by a number of initial solution consisting of genetic manipulation to prepare for the initial population. Set evolution generation counter t = 0 and the maximum evolution algebra T. Individuals N is randomly generated as the initial population.

(2) Self-evaluation: Calculate the fitness population P (t) of each individual.

(3) Select the operator: The purpose is to select an excellent number for individuals from the current population, so that they have the opportunity as a parent on behalf of the next generation of breeding.

(4) Cross-operation: The crossover effect in groups. The so-called cross refers to the part of the structure of the two parent individuals to generate new individuals to replace the restructuring operation. The crossover plays a central role in Genetic algorithm.

(5) Mutation operation: A randomly selected string is mutated into a new generation of individuals to provide the opportunity for evolution. Population P (t) after selection, crossover and mutation operation to get the next generation of population P(t+1) thereafter.

This work was supported in part by the Fundamental Research Fund of Beijing Jiaotong University under Grant 2013JBM036.

Jinchuan Ke is with School of Economics & Management, Beijing Jiaotong University, Beijing 100044, P. R. China (e-mail: jchke99@126.com).

Yi Yu, Biyao Yan and Ying Ren are with Beijing Jiaotong University, Beijing 100044, P. R. China

(6) Determine the termination condition: if t = T, places the resulting evolution of the fitness with the largest output as the optimal solution to terminate the calculation.

Cross refers to the part of the structure of the two parent individuals to generate new individuals to replace the restructuring operation. By cross, genetic algorithm can improve search capability to leap. Crossover rate will be based on the population of the two individuals randomly switching certain genes to produce new genetic combinations.

Discrete recombination can exchange the value of variable between individuals. Considering the following individual with three variables:

parent 1: 12 25 5
parent 2: 123 4 34

Each son individual can choose the parent value randomly with equal probability to reorganize a new son individual entity:

Son l; 123 4 5
Son 2: 12 25 34

For the multi-point crossover, m cross location has no repeated random selection. To produce two new generations, continuous exchanges happen in the point of intersection between the variables, except the first variable and the first intersection. For example, the intersection location was 2, 6, 10 as follows.

parent 1: 01110011010 01101111011 Son 1
parent 2: 10101100101 10110000100 Son 2

III. IMPROVED MULTI-OBJECTIVE GENETIC ALGORITHMS

A. Multi-objective Genetic Algorithm

Multi-objective genetic algorithm is to construct a non-dominated set for optimal solution. Evolutionary optimization algorithm can be obtained by a multiple pareto optimal solution, forming a pareto non-dominated set rather than a single solution, it has the advantage of solving multi-objective optimization problem.

SMOCEA has two evolutionary stages: population evolutionary stage and the outstanding individual evolutionary stages. In population evolutionary stage, the various population groups maintain a belief in evolutionary set, and select outstanding individuals to update their belief set from the population evolution. The annexation population reflects the relationship between collaboration and competition among populations. In outstanding individual evolutionary stages, the various groups of individuals maintain the belief set to explore more outstanding individuals in the sparse area neighborhood by way of further evolution.

B. Population culture

Culture algorithm is a double evolutionary mechanism proposed by Robert G · Reynolds in 1994 as a cultural experience in the past. In cultural algorithm, various groups are maintaining a population evolutionary set. The algorithm selects outstanding individuals from the population evolutionary belief set, the set of individual belief is to maintain a certain number of coefficients by extrusion. After the maintenance is completed, the knowledge and cultural formation of population are extracted.

Maintaining the diversity of population distribution is one of the key issues to solve multi-objective evolutionary algorithm. In this paper, we use biodiversity indicators to measure the diversity of belief set.

C. Co-evolution

After the evaluation is completed, the belief set from the finest populations N2 individuals is randomly selected to participate in the next evolution of the population. The number is set to participate:

$$N_2 = \eta \cdot N \cdot \mu \tag{2}$$

Where, η [η ∈ (0,1)] represents outstanding individual utilization; N is the population size; μ [μ ∈ (0,1)] is the proportion of the population.

If μ population is greater, the degree of internal learning (1 -μ) is smaller. Thus, set N '= η • N. Where N' represents the total number of individuals participating in the next generation of evolution, including N1 and N2. The next generation of the population N0 (N0 = N-N1-N2) individuals press the partial order to fill the next population.

D. SMOCEA algorithm flow

After obtaining the data set initialized to n populations and establishing a separate set, the set can be used for the evolution of population evolution based on the objective function. After selecting the outstanding individuals from the population evolutionary belief set into a new set N2, individuals are randomly selected to participate the next evolution. By calculating the population control capabilities, the continuous generation of population control I defined the vulnerable populations.

Termination conditions for each generation sub-populations govern the degree of similarity of sub-optimal capacity and ability to control populations of the weakest among sub-populations of 70%.

IV. EMPIRICAL ANALYSIS

We start from the analysis of the balance between profit ability and risk level, to increase profitability and risk control as the two major goals of investment decision, combining with genetic algorithm for multi-objective optimization and optimal trade-off of stock on the profitability and risk degree. Stock daily net growth rate is identified to measure the earnings, and the net growth rate of standard deviation is used to measure the risk, which constitute a multi-objective function by using genetic algorithm with systematic optimization for a non-dominated solution set, the optimal portfolio.

ISBN: 978-1-61804-292-7

A. *Data processing*

Collect 40 open-end stock 2013 annual data, including cumulative net profit and the risk free rate of return. The average weekly net stock of the net growth rate is calculated for G_w, the stock profit ability A_w and the anti-risk ability R_w are also computed. In addition, based on the original data, the abnormal data is rectified so that each stock can get 45 samples. Then the multi-objective genetic algorithm is used to calculate the optimization results under the balance of profit and risk.

B. *Optimization of multi-objective genetic algorithm*

In accordance with the evolution of evolutionary design process, the 40 stocks initial population is set. For population diversity, this study chose parameters G_w, A_w, R_w with 0.3,0.5,0.7 for testing, and ultimately determine the most appropriate value.

Let the weak dominated population be annexed, retaining the outstanding individuals for population variation and evolution. This paper set the termination criteria with commonly used 200 iterations. In each generation of sub- populations, when the optimal sub-population control ability and the weakest control ability among sub-populations achieve similar degree, then the iteration is terminated. The output of the solution set is the final result, namely the non dominated set.

The solution set is derived from the demand for non-pareto dominating set. By means of Matlab, the investment portfolio results is obtained and shown in Figure 1, the screened six outstanding stocks are number 1, 5, 8, 9, 29, 34 respectively, which is shown in Table 1.

The results show that the number of iteration constraint can not be too strict, otherwise no choice may be excluded; and the iterative constraint can not too loose, otherwise no restraint.

With the increase of generation, the evolution results become better and better. After 200 iterations of population, the optimal solution of the objective function is shown in Figure 2. We run the program 3 times again, and the results are shown in Figure 3, Figure 4, and Figure 5. It can be seen that the stability is very good under 200 iterations, the multi- objective genetic algorithm has achieved a satisfactory result.

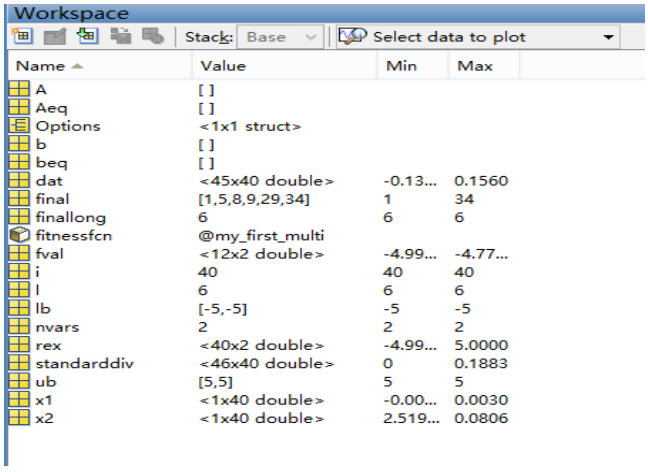


Fig. 1 Genetic algorithm result

Table 1 Multi-objective genetic algorithm result

Multi-objective results	Yield (Gw)	Individual risk (δ)
1 National Agricultural Technology	0.026	0.0017
5 card Electronic	0.030	0.0020
8 Golden Shares	0.029	0.0018
9 letter states pharmacy	0.033	0.0029
29 Santai Electronics	0.029	0.0022
34 Fudan Fuhua	0.030	0.0029

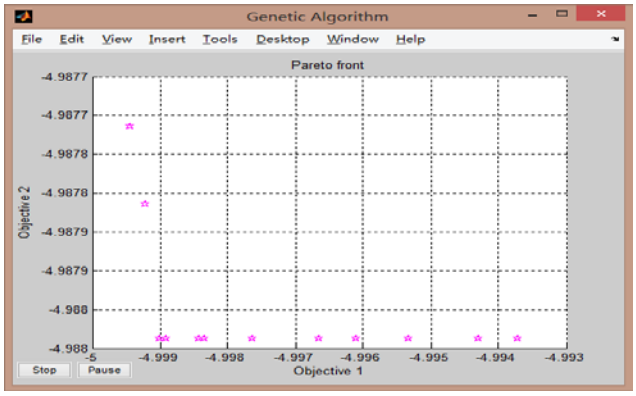


Fig. 2 Optimal solution 1 scatter

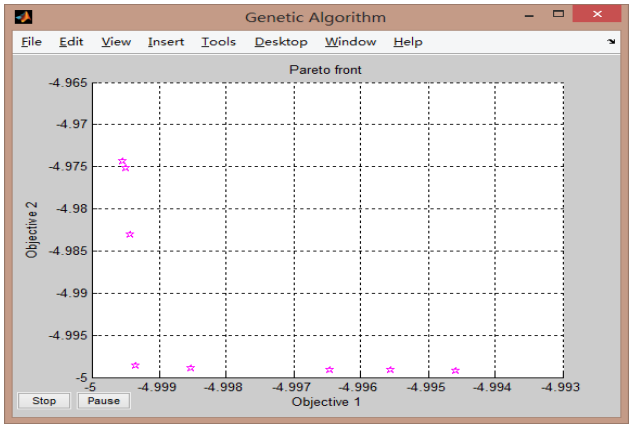


Fig. 3 Optimal solution 2 scatter

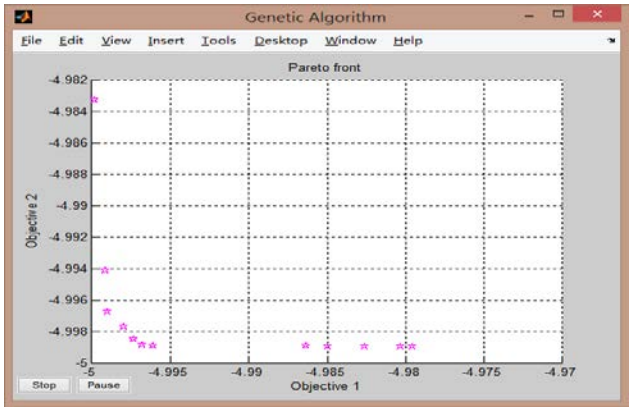


Fig. 4 Optimal solution 3 scatter

ISBN: 978-1-61804-292-7

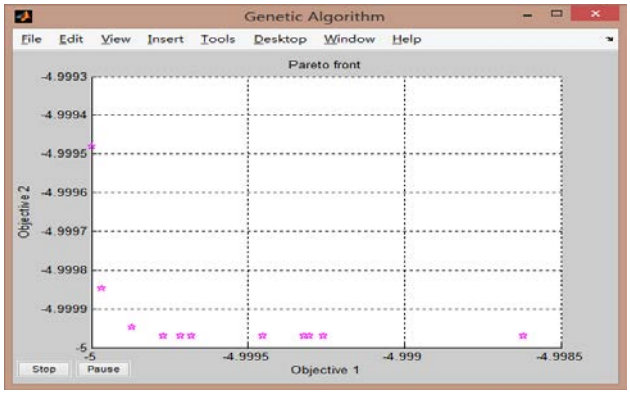


Fig. 5 Optimal solution 4 scatter

C. Comparison with Mean-variance portfolio model

Portfolio is usually expressed as a combination of the expected rate of return. Weighted average expected rate of return can be used to calculate the return of portfolio investment which consist of a variety of securities. Portfolio risk is usually expressed by a combination of standard deviation and the covariance between the weighted average

The return of portfolio is:

$$E(r_p) = \sum_{i=1}^{n} w_i r_i \tag{3}$$

Portfolio risk formula is：

$$\sigma_p = \left[\sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{ij} \right]^{\frac{1}{2}} \tag{4}$$

Based on the data collected from January 1, 2013 to December 30, 2013, we chose the general investor risk preferences to calculate the portfolio by using Markowitz Mean-variance model with the constraint of return and risk have the equal weights, the result shows that the best performance stocks are number 4,10,8,30,29,39 as shown in Table 2.

Table 2 Mean-variance portfolio algorithm result

Artificial screening results	Yield (Gw)	Individual risk (δ)
4 Wright	0.023	0.0052
10 South Building	0.025	0.0073
8 Golden Shares	0.029	0.0018
30 Dongyirisheng	0.021	0.0041
29 Santai Electronics	0.029	0.0022
39 New South Seas	0.013	0.0023

The return and risk calculation by genetic algorithm is obtained such as shown in Figure 2. In accordance with the investment portfolio, the return of six outstanding stock portfolio based on the genetic algorithm is 0.026, and the risk is 0.00341. Manually screened stock portfolio return is 0.020, and the risk is 0.00428, it is concluded that the use of genetic algorithm provides a fair performance for the portfolio.

V. Conclusion

In financial decision, due to the complicated procedures in Markowitz portfolio with many different combinations, the optimal portfolio is usually difficult to obtain. This paper proposes a new model based on genetic algorithms by introducing VaR constraint to solve the problem. In order to meet the requirements of different risk preferences of the portfolio, we also introduced a portfolio's risk preference factor, resulting in a multi-set of optimized portfolio. Through case studies, the genetic algorithm model has got a satisfied result, showing the new model proposed in this paper provides an effective theoretical guidance and decision-making basis for investors.

References

[1] Wang Yan and Hou Couyin, “Determination based on the type of portfolio construction stocks and genetic algorithms”, Forecast, vol. 4, pp.59-66, 2010

[2] Ai Hu, “Column for multi-objective optimization genetic algorithm portfolio model ”, Information Aspect, vol. 7, pp. 67-69, 2009.

[3] Gui-yuan Yang and Xiang Wang, “Portfolio investment based on genetic algorithm, Technology and Industry, vol. 3, pp44-49, 2009.

[4] Ling Yan, “Genetic algorithms and empirical cardinality constraints portfolio optimization analysis”, Journal of Northwest Normal University, vol. 47, no. 2, pp.26-30, 2011.

[5] Xiao-Mei Zhu, Philip Kwok steel, Yuan Yan, “Simulation model portfolio VaR constraint based on genetic algorithm”, Journal of Jianghan University, vol.33, no.4, pp.48-52, 2006.

[6] Yudong Zhang, *Algorithm and Applied Research for Optimization*, World Science Publisher, 2012.

[7] He Yang, Yonggui Du, “Improve search performance improved genetic algorithm selection strategy”, unpublished.

[8] Ying-jie Lei, Shan-wen Zhang, *MATLAB Genetic Algorithm Toolbox and Application*, Xi'an University of Electronic Science and Technology Press, 2005.

Some Properties of The Solution of Beltrami Equation

Melike Aydoğan and Durdane Öztürk

Abstract—Let $f(z) = h(z) + \overline{g(z)}$ be the sense-preserving harmonic mapping, then it satisfies Beltrami differential equation $\overline{f}_{\overline{z}} = \omega(z) f_z$. In the present paper we will investigate the solution of this equation in the open unit disc $\mathbb{D} = \{z \mid |z| < 1\}$.

Keywords—Distortion theorem, Growth theorem, Complex dilatation.

I. Introduction

LET Ω be the family of functions $\phi(z)$ which are analytic in $\mathbb{D}$ and satisfying the conditions $\phi(0) = 0$, $|\phi(z)| < 1$ for all $z \in \mathbb{D}$, and let Ω_a denote the class of functions $\psi(z)$ which are regular in $\mathbb{D}$ and satisfy the condition $|\psi(z)| < a$ for every for every $z \in \mathbb{D}$.

Next, $\mathcal{A}_n$ denote the class of analytic functions of the form $s(z) = z + c_{n+1}z^{n+1} + c_{n+2}z^{n+2} + \ldots$, and let $P_n(a)$ designate the class of functions $p(z) = 1 + p_n z^n + p_{n+1}z^{n+1} + \ldots$ which are analytic in $\mathbb{D}$ and satisfies the condition

$$\left|\frac{1 - p(z)}{1 + p(z)}\right| < a, \tag{1}$$

where $0 < a \leq 1$. Let $s(z)$ be an element of $\mathcal{A}_n$ and satisfies the condition $\left|\dfrac{z\dfrac{s'(z)}{s(z)} - 1}{z\dfrac{s'(z)}{s(z)} + 1}\right| < a$, $(0 < a \leq 1)$ for all $z \in \mathbb{D}$, then $s(z)$ is called starlike of order a. This definition was given by K.S. Padmanabhan [7], the class of such function is denoted by $S_n^*(a)$. Let $s_1(z)$, $s_2(z)$ be elements of $\mathcal{A}_1$, if there exists a function $\phi(z) \in \Omega$ such that $s_1(z) = s_2(\phi(z))$ for every $z \in \mathbb{D}$, then we say that $s_1(z)$ is subordinate to $s_2(z)$ and we write $s_1(z) \prec s_2(z)$. Specially $s_2(z)$ is univalent in $\mathbb{D}$ then $s_1(z) \prec s_2(z)$ if and only if $s(\mathbb{D}) \subset s_2(\mathbb{D})$ and $s_1(0) = s_2(0)$ implies $s_1(\mathbb{D}_r) \subset s_2(\mathbb{D}_r)$, where $\mathbb{D}_r = \{z \mid |z| < r,\ 0 < r < 1\}$. (Subordination and Lindelöf principle [2])

Finally, a planar harmonic mapping in the open unit disc $\mathbb{D}$ is a complex-valued harmonic function f, which maps onto the some planar domain $f(\mathbb{D})$. Since $\mathbb{D}$ is a simply connected domain, the mapping f has a canonical decomposition $f(z) = h(z) + \overline{g(z)}$, where $h(z)$ and $g(z)$ are analytic in $\mathbb{D}$ and have the following power series expansions

$$h(z) = \sum_{n=0}^{\infty} a_n z^n, \quad g(z) = \sum_{n=0}^{\infty} b_n z^n, \tag{2}$$

where $a_n, b_n \in \mathbb{C}$, $n = 0, 1, 2, \ldots$ as usual we call $h(z)$ the analytic part of f and $g(z)$ is co-analytic part of f. An elegant and complete account of the theory of harmonic mapping is given Duren's monograph [1]. Lewy proved in 1936 [2] that the harmonic mapping is locally univalent if and only if its Jacobian $J_f = |h'(z)|^2 - |g'(z)|^2$ is different from zero. In the view of this result, locally univalent harmonic mappings in the unit disc $\mathbb{D}$ are either sense-preserving if $|h'(z)| > |g'(z)|$ or sense-reversing if $|g'(z)| > |h'(z)|$ in $\mathbb{D}$. In this paper, we will restrict ourselves to the study of sense-preserving harmonic mappings. We will also note that $f(z) = h(z) + \overline{g(z)}$ is sense-preserving in $\mathbb{D}$ if and only if $h'(z)$ does not vanish in $\mathbb{D}$, and the second dilatation $\omega(z) = \dfrac{g'(z)}{h'(z)}$ has the property $|\omega(z)| < 1$ for all $z \in \mathbb{D}$. Therefore, the class of all sense-preserving harmonic mappings in the open unit disc with $a_0 = b_0 = 0$ and $a_1 = 1$ will be denoted by S_H, thus S_H contains the standard class S of univalent functions. The family of all mappings S_H with the additional property $g'(0) = 0$, i.e., $b_1 = 0$ is denoted by S_H^0. Hence it is clear that $S \subset S_H^0 \subset S_H$.

In this paper, we consider the class of sense-preserving harmonic mappings which is defined by

$$S_{H(n)} = \Big\{ f(z) = h(z) + \overline{g(z)} \,\Big|\, h(z) = z + a_{n+1}z^{n+1} + a_{n+2}z^{n+2} + \ldots,\ g(z) = b_1 z + b_{n+1}z^{n+1} + b_{n+2}z^{n+2} + \ldots,\ |b_1| < 1 \Big\},$$

and the solution of non-linear elliptic partial differential equation $\overline{f}_{\overline{z}} = \omega(z) f_z$ under the condition $\dfrac{g'(z)}{h'(z)} = b_1 p_n(a)$, $h(z) \in S_n^*(a)$. Therefore, the aim of this paper we will need the following lemmas and theorem.

Lemma 1.1([5]) Let $\phi(z) = \alpha_n z^n + \alpha_{n+1}z^{n+1} + \ldots$ $(\alpha \neq 0, n \geq 1)$ be analytic in $\mathbb{D}$. If the maximum value of $|\phi(z)|$ on the circle $|z| = r < 1$ is attained at $z = z_0$, then we have $z_0\phi'(z_0) = m\phi(z_0)$, $m \geq n$ and every $z \in \mathbb{D}$.

Lemma 1.2([2])

$$p(z) \in P_n(a) \Longleftrightarrow p(z) = \frac{1 - z^n\psi(z)}{1 + z^n\psi(z)}, \quad \psi(z) \in \Omega_a$$

M. Aydoğan and D.Öztürk are with the Department of Mathematics, Işık University, Istanbul, TURKEY e-mail: melike.aydogan@isikun.edu.tr

Manuscript received April 19, 2005; revised January 11, 2007.

Theorem 1.3([2])

$$s(z) \in S_n^*(a) \Longleftrightarrow s(z) = z \exp\left(-2\int_0^a \frac{t^{n-1}\psi(t)}{1+t^n\psi(t)}\right) dt$$

then

$$z\frac{s'(z)}{s(z)} = \frac{1-z^n\psi(z)}{1+z^n\psi(z)},$$

where $\psi(z) \in \Omega_a$.

II. Main Results

Lemma 2.1 Let $h(z)$ be an element of $S_n^*(a)$, then

$$\frac{r}{(1+ar^n)^{2/n}} \leq |h(z)| \leq \frac{r}{(1-ar^n)^{2/n}}, \tag{3}$$

$$\frac{1-ar^n}{(1+ar^n)^{1+2/n}} \leq |h'(z)| \leq \frac{1+ar^n}{(1-ar^n)^{1+2/n}}, \tag{4}$$

and the boundary values of $z\dfrac{h'(z)}{h(z)}$ on $|z| = r$ at $z = z_0$ is

$$z_0\frac{h'(z_0)}{h(z_0)} = \frac{1+2ar^ne^{i\theta}+ar^{2n}}{1-a^2r^{2n}}. \tag{5}$$

These results are sharp because the extremal function is $s(z) = \dfrac{z}{(1-az^n)^{2/n}}$.

Proof. Using Theorem 1.3, then we can write

$$z\frac{h'(z)}{h(z)} = \frac{1-z^n\psi(z)}{1+z^n\psi(z)}, \tag{6}$$

where $\psi(z) \in \Omega_a$. Thus we can write $z\psi(z) = a\phi(z)$, $\phi(z) \in \Omega$. Consequently we have

$$z\frac{h'(z)}{h(z)} = \frac{1-az^{n-1}\phi(z)}{1+az^{n-1}\phi(z)}. \tag{7}$$

On the other hand, the transformation $w(z) = \dfrac{1-az^{n-1}\phi(z)}{1+az^{n-1}\phi(z)}$ maps the circle $|\phi(z)| \leq r$ onto the circle

$$\left|w(z) - \frac{1+a^2r^{2n}}{1-a^2r^{2n}}\right| \leq \frac{2ar^n}{1-a^2r^{2n}}. \tag{8}$$

Using (6), (8) and the subordination principle, then we obtain

$$\left|z\frac{h'(z)}{h(z)} - \frac{1+ar^{2n}}{1-a^2r^{2n}}\right| \leq \frac{ar^n}{1-ar^n}$$

$$\Longrightarrow \frac{1-ar^n}{1+ar^n} \leq Re\left(z\frac{h'(z)}{h(z)}\right) \leq \frac{1+ar^n}{1-ar^n} \tag{9}$$

On the other hand, we have

$$Re\left(z\frac{h'(z)}{h(z)}\right) = r\frac{\partial}{\partial r}\log|h(z)|$$

Therefore the inequality (9) can be written in the following form

$$\frac{1-ar^n}{r(1+ar^n)} \leq \frac{\partial}{\partial r}\log|h(z)| \leq \frac{1+ar^n}{r(1-ar^n)}. \tag{10}$$

Integrating this inequality, we get (3). (4) and (5) are simple consequences of (9).

Remark The proof of this lemma can be found in [6] different way.

Theorem 2.2 The solution of the non-linear elliptic partial differential equation $\overline{f}_{\overline{z}} = \omega(z)f_z$ is

$$\frac{g(z)}{h(z)} = b_1\frac{1-az^{n-1}\phi_1(z)}{1+az^{n-1}\phi_1(z)}, \tag{11}$$

under the condition $\dfrac{g'(z)}{h'(z)} \prec b_1p(z)$, $p(z) \in P_n(a)$, $h(z) \in S_n^*(a)$.

Proof. Since $\dfrac{g(z)}{h(z)} \prec b_1p(z)$, $p(z) \in P_n(a)$, $h(z) \in S_n^*(a)$, then we can write

$$\omega(\mathbb{D}_r) = \left\{\frac{g'(z)}{h'(z)}\right| \left|\frac{g'(z)}{h'(z)} - \frac{b_1(1+a^2r^{2n})}{1-a^2r^{2n}}\right| \leq \frac{2|b_1|ar^n}{1-a^2r^{2n}}, \; 0<r<1\Big\} \tag{12}$$

Now we define the function $\phi(z)$ by

$$\frac{g(z)}{h(z)} = b_1\frac{1-(\phi(z))^n}{1+(\phi(z))^n}, \tag{13}$$

then $\phi(z)$ is analytic and $\phi(0) = 0$. We need to show that $|\phi(z)| < 1$ for all $z \in \mathbb{D}$. Assume to the contrary that there exists a $z_0 \in \partial\mathbb{D}_r$ that $|\phi(z_0)| = 1$. If we take derivative of (13) and after the simple calculations we get

$$\frac{g'(z)}{h'(z)} = b_1\frac{1-(\phi(z))^n}{1+(\phi(z))^n} - 2b_1n\frac{z\phi'(z)(\phi(z))^{n-1}}{(1+\phi(z))^{2n}}\cdot\frac{h(z)}{zh'(z)}$$

Considering (12), Lemma 1.1, Lemma 2.1 and (13) together, then we obtain

$$\omega(z_0) = b_1\frac{1-(\phi(z_0))^n}{1+(\phi(z_0))^n} - 2b_1n\frac{m\phi(z_0)(\phi(z_0))^{n-1}}{(1+\phi(z_0))^{2n}}\cdot\frac{1-a^2r^{2n}}{1+2ar^ne^{i\theta}+ar^{2n}} \notin \omega(\mathbb{D}_r)$$

But this is a contradiction. Therefore $|\phi(z)| < 1$ for all $z \in \mathbb{D}$. On the other hand, using Lemma 1.2 we can write

$$\frac{g(z)}{h(z)} = b_1\frac{1-z^n\psi(z)}{1+z^n\psi(z)}, \quad \psi(z) \in \Omega_a \tag{14}$$

Thus we can write $z\psi(z) = a\phi_1(z)$, $\phi_1(z) \in \Omega$. Consequently

$$\frac{g(z)}{h(z)} = b_1\frac{1-az^{n-1}\phi_1(z)}{1+az^n\phi_1(z)}.$$

This shows that the theorem is true.

Corollary 2.3 Let $f(z) = h(z) + \overline{g(z)}$ be the solution of the non-linear elliptic partial differential equation $\overline{f}_{\overline{z}} = \omega(z)f_z$

ISBN: 978-1-61804-292-7

under the condition $\omega(z) = \dfrac{g'(z)}{h'(z)} = b_1 p(z)$, $p(z) \in P_n(a)$, $h(z) \in S_n^*(a)$. Then

$$\frac{[(1+|b_1|)-(1-|b_1|)ar^n][(1-|b_1|)-(1+|b_1|)ar^n]}{(1-ar^n)^2} \leq$$
$$1-|\omega(z)|^2 \leq$$
$$\frac{[(1+|b_1|)+(1-|b_1|)ar^n][(1-|b_1|)+(1+|b_1|)ar^n]}{(1+ar^n)^2}$$

$$\frac{(1+|b_1|)+(1-|b_1|)ar^n}{1+ar^n} \leq 1+|\omega(z)| \leq \frac{(1+|b_1|)-(1-|b_1|)ar^n}{1-ar^n}$$

$$\frac{(1-|b_1|)-(1+b_1|)ar^n}{1-ar^n} \leq 1-|\omega(z)| \leq \frac{(1-|b_1|)+(1+|b_1|)ar^n}{1+ar^n}$$

Proof. This corollary is a simple consequence of (12).

REFERENCES

[1] Duren, P. , *Univalent Functions*, Grundieren der Mathematischen Wissenschafter Vol. 259, Springer-Verlag, New York, (1983).

[2] Duren, P. , *Harmonic Mappings in the Plane*, Vol. 156 of Cambridge Tracts in Mathematics, Cambridge University Press, Cambridge UK, (2004).

[3] Fukui, S., Sakaguchi, K., *An Extension of S. Ruschweyh*, Ball. Fac. Ed. Wakayama Univ. Nat. Sci., (1980), 1-3.

[4] Goodman A. W. , *Univalent Functions*, Volume I and Volume II, Mariner publishing Company INC, Tampa Florida, (1983).

[5] Jack, I. S. , *Functions Starlike and Convex of order* α, J. London Math. Soc. (2) 3 (1971) 469-474.

[6] Mogra, M.L., *On a Class of Starlike Functions in the Unit Disc*, J. Indian Math. Soc. 40, (1976), 158-165.

[7] Padmanabhan, K. S., *On certain classes of starlike functions in the unit disk*, J. Indian Math. Soc. (N.S.) 32 (1968), 89103. MR-0241626,(39, 2965).

[8] Robinson, R. M . , *Univalent majorants*, Trans. Amer. Math. Soc. 61 (1947), 1-35.

Analysis of Two Masses Sliding along a Cable with Delay

Tea Rukavina, Ivica Kožar

***Abstract*—**In this paper a model of two masses sliding along an elastic cable is presented. There is a delay between two masses, so the problem has been divided into two phases. In phase one there is only one mass and the solution at the end gives the initial conditions for phase two. In phase two the second mass is added and a system of eight differential equations with eight unknowns with initial conditions is derived. The validation of the model is shown in one example.

***Keywords*—**Delay differential equations (DDE), elastic cable, sliding masses

I. Introduction

TWO bodies sliding along a cable represent an engineering problem related to special structures. In practice we have encountered this problem when designing zip-lines for adrenalin amusement parks. The situation with two delayed masses occurs when two persons for some reason follow each other on the zip-line.

It has been shown in [1] that the problem of a mass sliding along a cable is a coupled system where there is no equilibrium without the sliding mass because the cable imposes nonlinear constraints onto dynamic equations of mass movement. This paper is an extension of the previous work in [1] with an addidtional mass included in the analysis. The added mass slides along the cable with a delay, after the first mass has already been released from the support. The analytical model of two masses sliding along a cable is derived and solved.

This coupled problem is modeled as a system of delayed differential equations (DDE). There are works developing special finite elements suited for cable structures like [4] but without the capability to solve delayed problems. The geometry of the problem is relatively simple so there is no need for space discretization, and kinematic relations are built into the system of DDEs. As a consequence, a DDE system can be solved using only time discretization, e.g. Runge-Kutta, but for simplicity we have adopted Mathematica solver [3] as a black box.

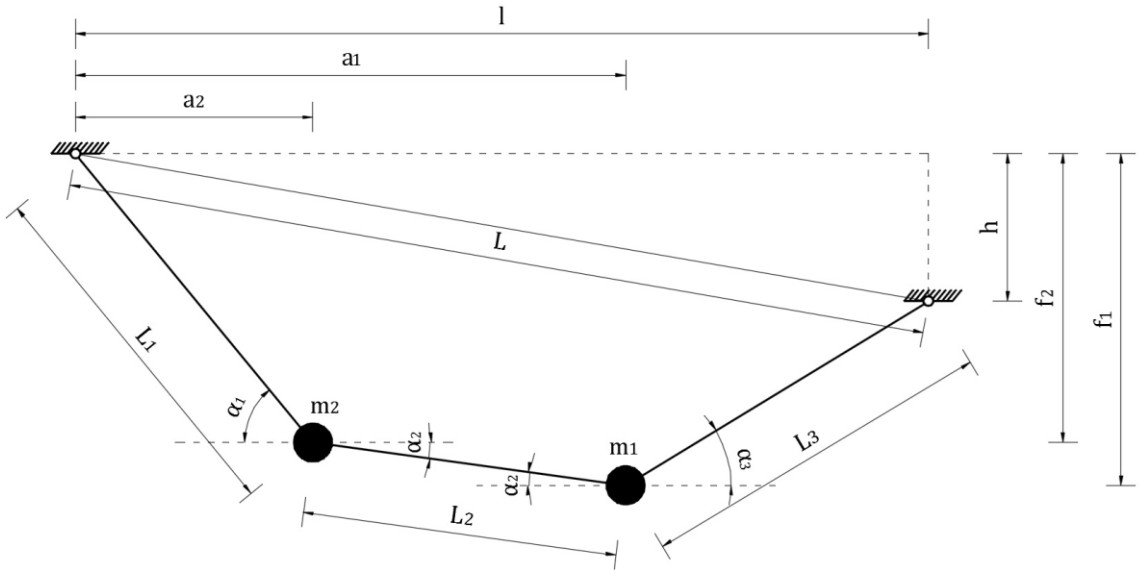


Fig. 1 Two masses sliding along a cable (phase two)

This work was supported in part by the University of Rijeka through Project no. 13.05.1.1.02 and by the European Union through Grant IPA2007/HR/16IPO/001-040513. The authors are thankful for this support.

T. Rukavina (corresponding author, phone: +38551265993, e-mail: tea.rukavina@gradri.uniri.hr) and I. Kožar (e-mail: ivica.kozar @gradri.uniri.hr) are with the Department for Computer Modeling of Materials and Structures, Faculty of Civil Engineering, University of Rijeka, Radmile Matejčić 3, 51000 Rijeka, Croatia.

II. THE MODEL

To form a model of two masses sliding along a cable (Fig. 1), we could use a system of delay differential equations [2]. Since there exists a delay τ between the two masses, all the variables related to the second mass would not depend only on time t, but on time $t-\tau$. This would lead to a system of eight delay differential equations with eight unknowns and the determination of initial conditions would be a rather complicated task.

To simplify the problem, we will divide it into two phases.

In phase one, only the first mass is sliding along the cable (Fig. 2), and in phase two the second mass is added (Fig.1).

We take the assumption that the cable is straight and the self-weight of the cable is neglected. Also, friction is not taken into account.

A. Phase one

As it was said earlier, in phase one there is only one mass sliding along the cable. This phase lasts until we introduce the second mass, at time $t_1 = \tau$.

Equations for the first phase are derived in [1], so only the final system of equations with initial conditions will be listed here. The nomenclature is taken from Fig. 2, where *a* is the horizontal, and *f* the vertical position of the mass.

From [1] we obtain:

$$\begin{aligned}
&\dot{a} = u,\\
&\dot{u} = \frac{EA}{m}\left(\frac{L_1+L_2}{L}-1\right)\left(\frac{l-a}{L_2}-\frac{a}{L_1}\right),\\
&\dot{f} = v,\\
&\dot{v} = \frac{G}{m}-\frac{EA}{m}\left(\frac{L_1+L_2}{L}-1\right)\left(\frac{f}{L_1}+\frac{f_1-h}{L_2}\right).
\end{aligned} \tag{1}$$

Regarding the initial conditions, an assumption is made for a_0, and f_0 is obtained from the following equation:

$$G - EA\left(\frac{L_1+L_2}{L}-1\right)\left(\frac{f}{L_1}+\frac{f_1-h}{L_2}\right)=0. \tag{2}$$

Also, both horizontal and vertical mass velocities are initially equal to zero. So, it can be written:

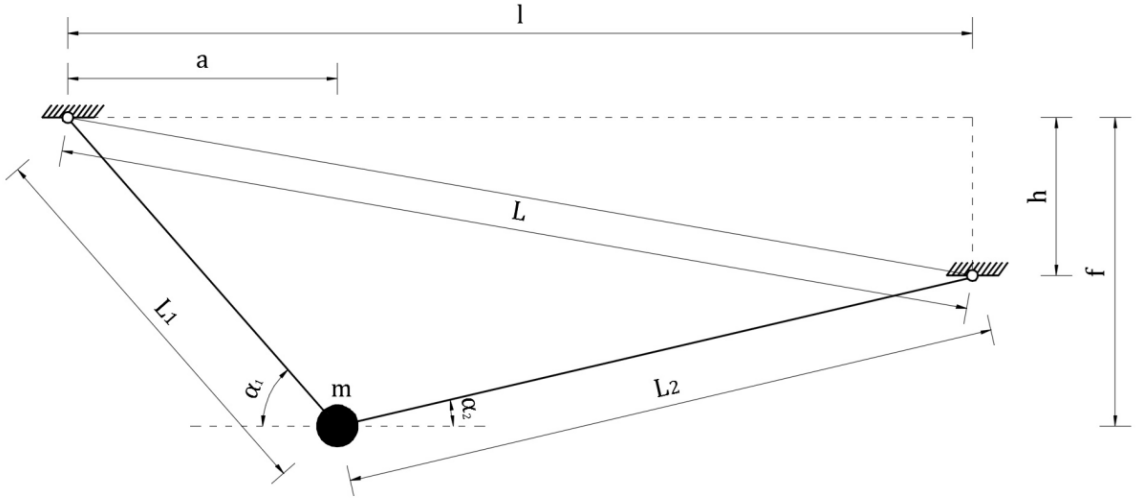


Fig. 2 One mass sliding on the cable (phase one)

$$\begin{aligned}
&u_0 = 0,\\
&v_0 = 0.
\end{aligned} \tag{3}$$

When we compute this system of four differential equations with for unknowns, a solution at time $t=\tau$ is found. We obtain the values for a_τ, f_τ, u_τ, and v_τ that will become the initial conditions for the first mass in phase two.

B. Phase two

In phase two, the second mass is added, while the first mass continues to slide along the cable. Of course, this will lead to differential equations that will depend on variables related to both masses. We assume that the time starts again from zero, so it will be named t_2. We can start from the dynamic balance equations of two masses sliding along the cable (Fig. 3). The force *T* is constant along the rope.

In Fig. 3 we can see the forces that act on each mass, so by adding up the forces along the *x* and *y* axes, we obtain the following equations for the first mass:

$$\begin{aligned}
&-T\cos\alpha_2 + T\cos\alpha_3 = m_1\ddot{a}_1\\
&-T\sin\alpha_2 - T\sin\alpha_3 + m_1 g = m_1\ddot{f}_1,
\end{aligned} \tag{4}$$

and for the second mass:

$$\begin{aligned} -T\cos\alpha_1 + T\cos\alpha_2 &= m_2\ddot{a}_2 \\ -T\sin\alpha_1 + T\sin\alpha_2 + m_2 g &= m_2\ddot{f}_2. \end{aligned} \tag{5}$$

Of course, a_1 and a_2 are horizontal positions of the first and second mass, and f_1 and f_2 are their vertical positions.

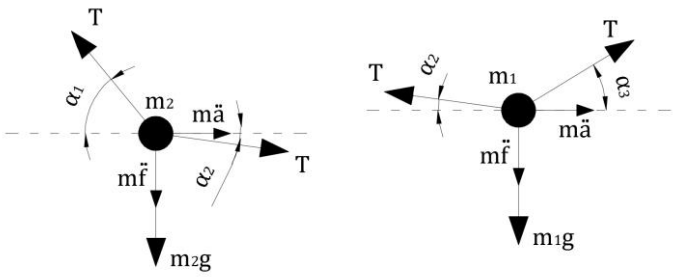


Fig. 3 Dynamic equilibrium of the sliding masses

Angles are calculated from the geometry:

$$\begin{aligned} &\cos\alpha_1 = a_2/L_1, && \sin\alpha_1 = f_2/L_1, \\ &\cos\alpha_2 = (a_1-a_2)/L_2, && \sin\alpha_2 = (f_1-f_2)/L_2, \\ &\cos\alpha_3 = (l-a_1)/L_3, && \sin\alpha_3 = (f_1-h)/L_3. \end{aligned} \tag{6}$$

We introduce another equation that is related to the length of the rope. Since we have an elastic cable, the following relation must be satisfied:

$$L_1 + L_2 + L_3 = L + \Delta L. \tag{7}$$

Since ΔL is the elongation of the cable that is equal to:

$$\Delta L = \frac{TL}{EA}, \tag{8}$$

we can transform (7) to obtain:

$$T = EA\left(\frac{L_1+L_2+L_3}{L} - 1\right). \tag{9}$$

The lengths of the cable are obtained from geometric relations:

$$\begin{aligned} L_1 &= \sqrt{a_2^{\,2} + f_2^{\,2}}, \\ L_2 &= \sqrt{(a_1-a_2)^2 + (f_1-f_2)^2}, \\ L_3 &= \sqrt{(l-a_1)^2 + (f_1-h)^2}. \end{aligned} \tag{10}$$

If we substitute (9) and (6) into (4) an (5), with a few transformations we get:

$$\begin{aligned} \ddot{a}_1 = \dot{u}_1 &= \frac{EA}{m_1}\left(\frac{L_1+L_2+L_3}{L} - 1\right)\left(\frac{l-a_1}{L_3} - \frac{a_1-a_2}{L_2}\right) \\ \ddot{f}_1 = \dot{v}_1 &= \frac{G_1}{m_1} - \frac{EA}{m_1}\left(\frac{L_1+L_2+L_3}{L} - 1\right)\left(\frac{f_1-f_2}{L_2} + \frac{f_1-h}{L_3}\right), \end{aligned} \tag{11}$$

and:

$$\begin{aligned} \ddot{a}_2 = \dot{u}_2 &= \frac{EA}{m_2}\left(\frac{L_1+L_2+L_3}{L} - 1\right)\left(\frac{a_1-a_2}{L_2} - \frac{a_2}{L_1}\right) \\ \ddot{f}_2 = \dot{v}_2 &= \frac{G_2}{m_2} + \frac{EA}{m_2}\left(\frac{L_1+L_2+L_3}{L} - 1\right)\left(\frac{f_1-f_2}{L_2} + \frac{f_2}{L_1}\right). \end{aligned} \tag{12}$$

With four additional equations:

$$\dot{a}_1 = u_1,\ \dot{f}_1 = v_1,\ \ \dot{a}_2 = u_2,\ \dot{f}_2 = v_2, \tag{13}$$

we form a system of eight differential equations with eight unknowns. The unknowns are a_1, f_1, u_1, v_1, a_2, f_2, u_2, and v_2.

We have to determine the initial conditions for both masses. As it was mentioned earlier, the initial conditions for the first mass are the final results from phase one, so we can write:

$$a_{10} = a_\tau,\ f_{10} = f_\tau,\ u_{10} = u_\tau,\ v_{10} = v_\tau. \tag{14}$$

For the second mass we determine the initial conditions in a similar way it was done for the first mass in phase one. We assume an initial value for a_{20}, and we determine f_{20} from the second equation in (11) when $\ddot{f}_2 = 0$:

$$G_2 + EA\left(\frac{L_1+L_2+L_3}{L} - 1\right)\left(\frac{f_1-f_2}{L_2} + \frac{f_2}{L_1}\right) = 0. \tag{15}$$

In this case the horizontal and vertical velocities of the second mass are also equal to zero: $u_{20} = 0$ and $v_{20} = 0$.

This completes the defining of the model of two masses sliding along a cable.

Of course, when in phase two all the variables related to the second mass are assumed to be zero, we obtain the same results as in [1], when only one mass was taken into account.

III. Example

We take the same geometric and material properties of the cable as in the example with a longer cable in [1]:

$$\begin{aligned} &l = 6000\ m \\ &h = 600\ m \\ &L = 6032\ m \\ &EA = 6\cdot 10^6\ N \end{aligned}$$

Also, the masses are equal:

$$m_1 = m_2 = 1500\ kg$$

The total analysis time in phase one is $t_1 = \tau = 10s$. The total analysis time in phase two is taken to be slightly before

the first mass reaches the end of the cable. In this case we take $t_2 = 15s$. So, the total analysis time for both phases is $t = t_1 + t_2 = 25s$.

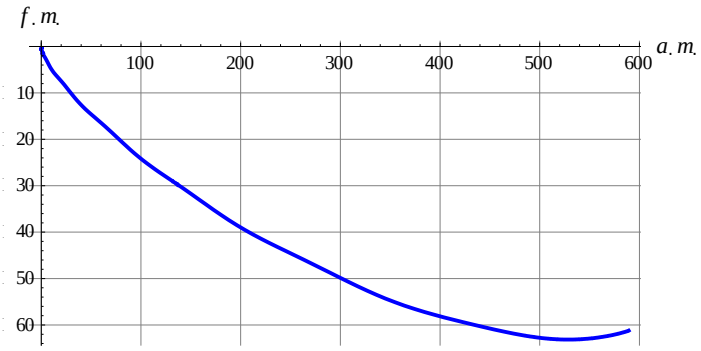


Fig. 5 The path of the first mass in both phases

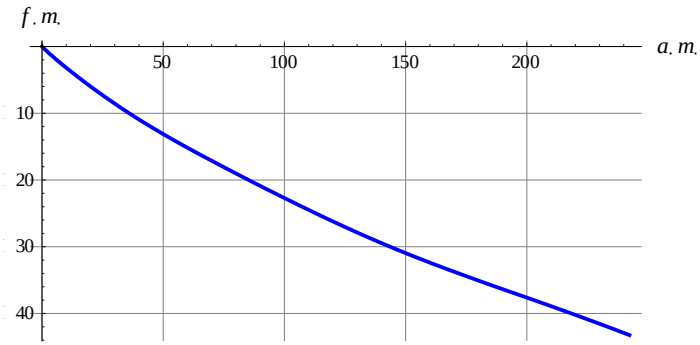


Fig. 5 The path of the second mass in phase two

Fig. 4 shows the path of the first mass in both phases, and Fig. 5 shows the path of the second mass in phase two.

We can see that in $t_2 = 15s$, the first mass reaches the end of the cable, while the second mass reaches approximately $a \cong 242m$ and $f_2 \cong 43m$.

The tension force in the cable for phase one is shown on Fig. 6, and for phase two in Fig. 7.

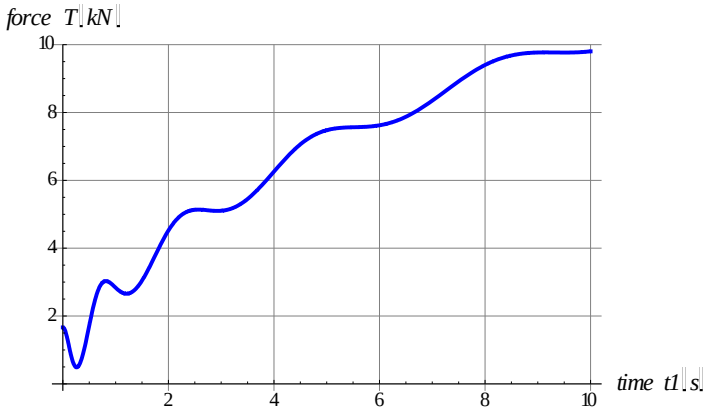


Fig. 6 Tension force in the cable for phase one

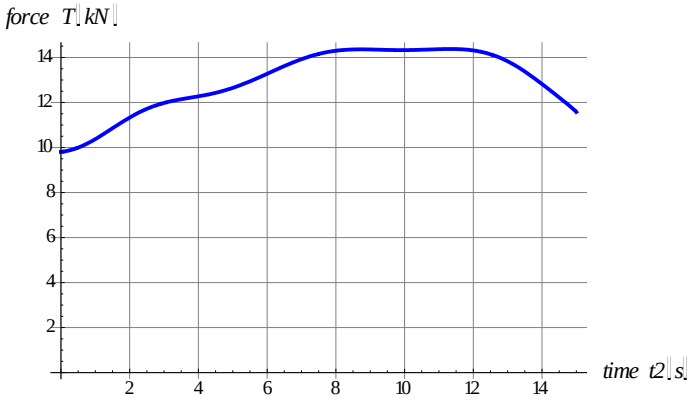


Fig. 7 Tension force in the cable for phase two

IV. Conclusion

The application of the analytical model formulated as a system of DDEs and its solution without the usual finite element discretization has proven successful in this case of non-linear and somewhat exclusive type of structure.

Authors plan to extend this model by adding pendulums on which masses will be attached. That will make it possible to describe engineering structures in a more realistic way because usually the center of the mass is dislocated from the axis of the cable.

References

[1] I. Kožar, N. Torić Malić, *Analysis of Body Sliding along Cable*. Coupled Systems Mechanics, Vol. 3, No. 3, 2014, pp. 291-304

[2] L. F. Shampine, I. Gladwell, S. Thompson, "Solving ODEs with MATLAB," Cambridge University Press, 2003

[3] Wolfram Language and System, Documentation Center, 2015, http://reference.wolfram.com/language/

[4] A. Andreu, L. Gil, P. Roca, *A new deformable catenary element for the analysis of cable net structures*. Computers&Structures, Vol. 84, 2006, pp. 1882-1890

ISBN: 978-1-61804-292-7

The sum over $E_{a,b}$

A. Chillali, A. Tadmori and M. Ziane

***Abstract*—**Let d is a positive integer. In this article we will study the elliptic curve defined over the ring $\mathbb{F}_{2^d}[\mathcal{E}]$; $\mathcal{E}^2 = 0$. More precisely we will give many various explicit formulas describing the binary operations calculus in $E_{a,b,c}$.

***Keywords*—**Elliptic Curves, Finite Ring, Cryptography.

I. Introduction

LET d be an integer, we consider the quotient ring $A = \frac{\mathbb{F}_{2^d}[X]}{(X^2)}$ where $\mathbb{F}_{2^d}$ is the finite field of order 2^d. Then the ring A is identified to the ring $\mathbb{F}_{2^d}[\mathcal{E}]$ with $\mathcal{E}^2 = 0$ ie: see [1] and [2], $A = \{ a_0 + a_1.\mathcal{E} \mid a_0; a_1 \in \mathbb{F}_{2^d} \}$.

We consider the elliptic curve over the ring A which is given by equation: $Y^2Z + cXYZ = X^3 + aX^2Z + bZ^3$,

where a, b and c are in A and c^6b is invertible in A, but we can take c = 1; see, [3].

II. Notations

Let a, $b \in A$ such that b is invertible in A and c = 1. We denote the elliptic curve over A by $E_{a,b}(A)$ and we write:

$E_{a,b}(A) = \{ [X : Y : Z] \in P_2(A) \mid Y^2Z + XYZ = X^3 + aX^2Z + bZ^3 \}$.

If $b_0 \in \mathbb{F}_{2^d}\backslash\{0\}$ and $a_0 \in \mathbb{F}_{2^d}$, we also write:

$E_{a_0,b_0}(\mathbb{F}_{2^d}) = \{ [X : Y : Z] \in P_2(\mathbb{F}_{2^d}) \mid Y^2Z + XYZ = X^3 + a_0X^2Z + b_0Z^3 \}$.

III. Classification of elements of $E_{A,B}(A)$

Let $[X : Y : Z] \in E_{a,b}(A)$, where X, Y and Z are in A. We have two cases for Z:

• Z invertible: then $[X : Y : Z] = [XZ^{-1} : Y Z^{-1} : 1]$; hence we take just [X:Y:1].

• Z non invertible: So $Z = z_1\varepsilon$, see [4], in this cases we have tow cases for Y.

- Y invertible: Then $[X : Y : Z] = [XY^{-1} : 1 : ZY^{-1}]$; so we just take $[X : 1 : z_1\varepsilon]$; then is verified the equation of $E_{a,b}(A)$: $Y^2Z + XYZ = X^3 + aX^2Z + bZ^3$,

so we can write:

$$a = a_0 + a_1\varepsilon$$

$$b = b_0 + b_1\varepsilon$$

$$X = x_0 + x_1\varepsilon$$

We have: $z_1\varepsilon + (x_0 + x_1\varepsilon).z_1\varepsilon = (x_0 + x_1\varepsilon)^3 + (a_0 + a_1\varepsilon).(x_0 + x_1\varepsilon)^2.z_1\varepsilon + (b_0 + b_1\varepsilon).{z_1}^3\varepsilon^3$

Which implies that

$$z_1\varepsilon + x_0z_1\varepsilon = {x_0}^3 + ({x_0}^2x_1 + a_0{x_0}^2z_1)\varepsilon$$

Then

$$(z_1 + x_0z_1)\varepsilon = {x_0}^3 + ({x_0}^2x_1 + a_0{x_0}^2z_1)\varepsilon$$

Since $(1, \varepsilon)$ is a base of the vector space A over $\mathbb{F}_{2^d}$, then $x_0 = 0$, so $X = x_1\varepsilon$ and $z_1\varepsilon = 0$ (ie $z_1 = 0$) hence $[X: 1: z_1\varepsilon] = [x_1\varepsilon : 1 :0]$.

- Y non invertible: then we have $Y = y_1\varepsilon$, so $X = x_0 + x_1\varepsilon$ is invertible so we take $[X: Y: Z] \sim [1: y_1\varepsilon: z_1\varepsilon]$ thus $1 + a.z_1\varepsilon = 0$, ie $1 + a_0z_1\varepsilon = 0$ which is absurd.

Proposition 1:

Every element of $E_{a,b}(A)$, is of the form $[X: Y: 1]$ or $[x\varepsilon: 1: 0]$, where $x \in \mathbb{F}_{2^d}$ and we write:

$E_{a,b}(A) = \{[X:Y:1] \in P_2(A) \mid Y^2 + XY = X^3 + aX^2 + b\} \cup \{[x\varepsilon: 1: 0] \mid x \in \mathbb{F}_{2^d}\}$. [1].

IV. Explicit Formulas

We consider the canonical projection π defined by:

$$\pi: \mathbb{F}_{2^d}[\varepsilon] \mapsto \mathbb{F}_{2^d}$$
$$x_0 + x_1\varepsilon \to x_0$$

We have π is a morphism of ring.

* Let π_2 the mapping defined by :

$$\pi_2: E_{a,b}(A) \mapsto E_{a_0,b_0}(\mathbb{F}_{2^d})$$
$$[X: Y: Z] \to [\pi(X): \pi(Y): \pi(Z)]$$

The mapping π_2 is a surjective homomorphism of groups.

Theorem1:

Let $P = [X_1: Y_1: Z_1]$, $Q = [X_2: Y_2: Z_2]$ in $E_{a,b}(A)$ then $P + Q = [X_3: Y_3: Z_3]$:

• If $\pi_2(P) = \pi_2(Q)$ then :

- ✓ $X_3 = X_1Y_1Y_2 + X_2{Y_1}^2Y_2 + {X_2}^2{Y_1}^2 + X_1{X_2}^2Y_1 + a\,{X_1}^2X_2Y_2 + a\,X_1{X_2}^2Y_1 + a\,{X_1}^2{X_2}^2 + b\,X_1{Y_1}{Z_2}^2 + b\,X_2Y_2{Z_1}^2 + b\,{X_1}^2{Z_2}^2 + b\,Y_1{Z_2}^2Z_1 + b\,Y_2{Z_1}^2Z_2 + b\,X_1{Z_2}^2Z_1$
- ✓ $Y_3 = {Y_1}^2{Y_2}^2 + X_2{Y_1}^2Y_2 + a\,X_1{X_2}^2Y_1 + a^2\,{X_1}^2{X_2}^2 + b\,{X_1}^2X_2Z_2 + b\,X_1{X_2}^2Z_1 + b\,X_1Y_1{Z_2}^2 + b\,{X_1}^2{Z_2}^2 + ab\,{X_2}^2{Z_1}^2 + b\,Y_1{Z_2}^2Z_1 + b\,X_1{Z_2}^2Z_1 + ab\,X_1{Z_2}^2Z_1 + ab\,X_2{Z_1}^2Z_2 + b^2{Z_1}^2{Z_2}^2$
- ✓ $Z_3 = {X_1}^2X_2Y_2 + X_1{X_2}^2Y_1 + Y_1{}^2Y_2Z_2 + Y_1Y_2{}^2Z_1 + {X_1}^2{X_2}^2 + X_2{Y_1}^2Z_2 + {X_1}^2Y_2Z_2 + a\,{X_1}^2Y_2Z_2 + a\,{X_2}^2Y_1Z_1 + {X_1}^2{X_2}^2Z_2 + a\,X_1{X_2}^2Z_1 + b\,Y_1{Z_2}^2Z_1 + b\,Y_2{Z_1}^2Z_2 + b\,X_1{Z_2}^2Z_1$

• If $\pi_2(P) \neq \pi_2(Q)$ then :

A. Chillali is with the USMBA, LST, FPT, Taza, Morocco,
e-mail: abdelhakim.chillali@usmba.ac.ma.
A. Tadmori., is with UMP, FSO , Oujda.
M. Ziane., is with UMP, FSO , Oujda.

- ✓ $X_1 = X_1Y_2{}^2Z_1 + X_2Y_1{}^2Z_2 + X_1{}^2Y_2Z_2 + X_2{}^2Y_1Z_1 + a\ X_1{}^2X_2Z_2 + a\ X_1X_2{}^2Z_1 + b\ X_1Z_2{}^2Z_1 + b\ X_2Z_1{}^2Z_2$
- ✓ $Y_3 = X_1{}^2X_2Y_2 + X_1X_2{}^2Y_1 + Y_1{}^2Y_2Z_2 + Y_1Y_2{}^2Z_1 + X_1{}^2Y_2Z_2 + X_2{}^2Y_1Z_1 + a\ X_1{}^2Y_2Z_2 + a\ X_2{}^2Y_1Z_1 + a\ X_1{}^2X_2Z_2 + a\ X_1X_2{}^2Z_1 + b\ Y_1Z_2{}^2Z_1 + b\ Y_2Z_1{}^2Z_2 + b\ X_1Z_2{}^2Z_1 + b\ X_2Z_1{}^2Z_2$
- ✓ $Z_3 = X_1{}^2X_2Z_2 + X_1X_2{}^2Z_1 + Y_1{}^2Z_2{}^2 + Y_2{}^2Z_1{}^2 + X_1Y_1Z_2{}^2 + X_2Y_2Z_1{}^2 + a\ X_1{}^2Z_2{}^2 + a\ X_2{}^2Z_1{}^2$

Proof:

Using the explicit formulas in W.Bosma and H.Lenstras article see, [5], we prove the theorem. ■

V. Main results

Let $a = a_0 + a_1\varepsilon,\ b = b_0 + b_1$

Lemma1:

Let $P = [x_1\varepsilon: 1: 0]$ and $Q = [t_1\varepsilon: 1: 0]$ two points in $E_{a,b}(A)$ then: $P + Q = [(x_1 + t_1)\varepsilon: 1 + t_1\varepsilon: 0]$

Proof :

As $\pi_2(P) = \pi_2(Q)$, then by applying the formula (1) in theorem, we find the result. . ■

The following lemmas may be proved by using the explicit formulas in [5, p. 236—238].

Lemma 2:

Let $P = [x_1\varepsilon: 1: 0]$ and $Q = [t_0 + t_1\varepsilon: h_0 + h_1\varepsilon: 1]$ two points in $E_{a,b}(A)$, then :

$$P + Q = [t_0 + t_1\varepsilon: (x_1t_0{}^2 + h_1)\varepsilon + h_0: 1 + x_1\varepsilon]$$

Lemma3:

Let $P = [x_0 + x_1\varepsilon: y_1\varepsilon: 1]$ and $Q = [x_0 + t_1\varepsilon: h_1\varepsilon: 1]$ two points in $E_{a,b}(A)$ then :

$$P + Q = [(h_1a_0x_0{}^3 + y_1a_0x_0{}^3 + a_1x_0{}^4 + y_1b_0x_0 + h_1b_0x_0 + y_1x_0{}^3 + x_1b_0 + h_1b_0 + b_1x_0{}^2 + y_1b_0 + x_0b_1)\varepsilon + b_0x_0{}^2 + a_0x_0{}^4 + x_0b_0: (x_1a_0b_0 + a_1b_0x_0{}^2 + x_1b_0 + a_0b_1x_0{}^2 + b_0x_0{}^2x_1 + x_0b_1 + y_1b_0 + y_1a_0x_0{}^3 + t_1a_0b_0 + y_1b_0x_0 + b_0x_0{}^2t_1 + x_0{}^2b_1)\varepsilon + x_0{}^2b_0 + a_0b_0x_0{}^2 + b_0{}^2 + x_0b_0 + a_0{}^2x_0{}^4: (a_1x_0{}^3 + h_1x_0{}^2 + a_0x_1x_0{}^2 + y_1a_0x_0{}^2 + h_1a_0x_0{}^2 + h_1x_0{}^3 + x_0{}^2t_1 + b_0x_1 + y_1b_0 + b_1x_0 + y_1x_0{}^3 + h_1b_0)\varepsilon + a_0x_0{}^3 + x_0{}^4 + x_0{}^3 + b_0x_0]$$

Lemma4:

Let $P = [x_0 + x_1\varepsilon: y_0 + y_1\varepsilon: 1]$ and $Q = [x_0 + t_1\varepsilon: h_1\varepsilon: 1]$ two points in $E_{a,b}(A)$, where $y_0 \neq 0$ Then :

$$P + Q = [(a_0x_0{}^2t_1 + a_0x_0{}^2x_1 + x_0{}^2y_1 + h_1x_0{}^2 + b_0t_1 + t_1y_0{}^2 + b_0x_1)\varepsilon + x_0{}^2y_0 + x_0y_0{}^2: (x_0{}^2x_1y_0 + x_0{}^2y_1 + y_1x_0{}^3 + h_1a_0x_0{}^2 + y_1a_0x_0{}^2 + h_1b_0 + a_0x_1x_0{}^2 + b_0t_1 + h_1x_0{}^3 + b_1y_0 + h_1x_0{}^2 + a_1x_0{}^2y_0 + b_0x_1 + y_1b_0 + a_0x_0{}^2t_1 + h_1y_0{}^2)\varepsilon + a_0x_0{}^2y_0 + x_0{}^2y_0 + b_0y_0 + x_0{}^3y_0: (x_0{}^2x_1 + h_1x_0 + x_0{}^2t_1 + x_0y_1 + x_1y_0)\varepsilon + x_0y_0 + y_0{}^2]$$

Lemma5:

Let $P = [x_0 + x_1\varepsilon: y_0 + y_1\varepsilon: 1]$; $Q = [x_0 + t_1\varepsilon: y_0 + h_1\varepsilon: 1]$ two points of $E_{a,b}(A)$,where $y_0 \neq 0$, then :

$$P + Q = [(y_1x_0{}^3 + h_1a_0x_0{}^3 + y_1a_0x_0{}^3 + a_1x_0{}^4 + y_1b_0x_0 + h_1b_0x_0 + b_1x_0{}^2 + y_1b_0 + h_1b_0 + x_0b_1 + x_1b_0 + y_0{}^3x_1 + y_0{}^3t_1 + h_1y_0{}^2x_0 + y_1y_0{}^2x_0 + b_0x_1y_0 + b_0t_1y_0 + x_1x_0{}^2y_0 + a_0x_0{}^2t_1y_0 + a_0x_0{}^2x_1y_0)\varepsilon + b_0x_0{}^2 + a_0x_0{}^4 + x_0b_0 + x_0{}^3y_0 + x_0{}^2y_0{}^2: (b_0x_0{}^2t_1 + b_0x_0{}^2x_1 + x_0{}^2b_1 + a_0b_1x_0{}^2 + a_1b_0x_0{}^2 + y_1b_0 + x_0b_1 + x_1b_0 + y_1a_0x_0{}^3 + y_1b_0x_0 + x_1a_0b_0 + t_1a_0b_0 + t_1y_0{}^3 + y_0b_1 + x_0y_0{}^2h_1 + a_1x_0{}^3y_0 + b_0y_0x_1 + b_1y_0x_0 + a_0x_1x_0{}^2y_0)\varepsilon + a_0x_0{}^3y_0 + y_0{}^4 + x_0y_0{}^3 + y_0b_0 + x_0b_0 + b_0{}^2 + a_0b_0x_0{}^2 + a_0{}^2x_0{}^4 + x_0{}^2b_0 + b_0y_0x_0: (h_1x_0{}^3 + a_0x_1x_0{}^2 + a_1x_0{}^3 + b_0x_1 + b_1x_0 + h_1x_0{}^2 + h_1a_0x_0{}^2 + y_1a_0x_0{}^2 + x_0{}^2t_1 + y_1x_0{}^3 + y_1b_0 + h_1b_0 + x_0{}^2t_1y_0 + x_0{}^2x_1y_0 + h_1y_0{}^2 + y_1y_0{}^2 + t_1y_0{}^2)\varepsilon + x_0y_0{}^2 + x_0{}^4 + a_0x_0{}^3 + x_0{}^2y_0 + b_0x_0 + x_0{}^3]$$

Lemma6:

Let $P = [x_0 + x_1\varepsilon: y_0 + y_1\varepsilon: 1]$; $Q = [t_0 + t_1\varepsilon: h_0 + h_1\varepsilon: 1]$ two points in $E_{a,b}(A)$, where $x_0 \neq t_0$, or $y_0 \neq h_0$, then :

$$P + Q = [(t_0{}^2y_1 + h_1x_0{}^2 + a_0x_0{}^2t_1 + a_1x_0{}^2t_0 + a_0x_1t_0{}^2 + a_1x_0t_0{}^2 + b_1x_0 + b_1t_0 + b_0x_1 + b_0t_1 + t_1y_0{}^2 + x_1h_0{}^2)\varepsilon + x_0{}^2h_0 + t_0{}^2y_0 + a_0x_0{}^2t_0 + a_0x_0t_0{}^2 + b_0x_0 + x_0h_0{}^2 + t_0y_0{}^2 + b_0t_0: (a_0x_0{}^2t_1 + b_0x_1 + b_1x_0 + h_1x_0{}^2 + h_1a_0x_0{}^2 + y_1b_0 + h_1b_0 + b_0t_1 + h_1y_0{}^2 + b_1y_0 + y_1h_0{}^2 + b_1h_0 + x_0{}^2t_0h_1 + x_0{}^2t_1h_0 + x_0t_0{}^2y_1 + x_1t_0{}^2y_0 + t_0{}^2y_1 + a_1x_0{}^2h_0 + a_0t_0{}^2y_1 + a_1t_0{}^2y_0 + b_1 + a_1x_0{}^2t_0 + a_0x_1t_0{}^2 + a_1x_0t_0{}^2)\varepsilon + t_0{}^2y_0 + b_0x_0 + x_0t_0{}^2y_0 + x_0{}^2h_0 + x_0{}^2t_0h_0 + a_0x_0{}^2t_0 + a_0x_0t_0{}^2 + b_0y_0 + y_0h_0{}^2 + b_0t_0 + b_0h_0 + y_0{}^2h_0 + a_0t_0{}^2y_0 + a_0x_0{}^2h_0: (x_0{}^2t_1 + t_1h_0 + a_1x_0{}^2 + t_0h_1 + x_1t_0{}^2 + a_1t_0{}^2 + x_0y_1 + x_1y_0)\varepsilon + a_0t_0{}^2 + t_0h_0 + y_0{}^2 + x_0y_0 + x_0{}^2t_0 + x_0t_0{}^2 + h_0{}^2 + a_0x_0{}^2]$$

Acknowledgment

The authors would like to thank University of Mohammed First Oujda and FPT of Taza in MOROCCO for its valued support.

References

[1] A. Chillali, The j-invariant over $E_{3^d}{}^n$, Int.j.Open problems Compt. Math.Vol.5, No 4,December 2012,ISSN 1998-6262,

Copyright ICSRS Publication, (WWW.i-csrs.org,pp.106-111, 2012).

[2] A. Chillali, , Cryptography over elliptic curve of the ring $\mathbb{F}_q[\varepsilon], \varepsilon^4 = 0$ World Academy of science Engineering and Technology,78 (2011),pp.848-850

[3] A. Chillali, Elliptic curve over ring, International Mathematical Forum, Vol.6, no.31, 2011 pp.1501-1505

[4] A. Tadmori, A. Chillali and M. Ziane, Elliptic Curves Over SPIR of characteristic Two, proceeding of the 2013 international conference on applied mathematics and Computational Methode, www.europment.org/library/2013/AMCM-05.

[5] A. Tadmori, A. Chillali and M. Ziane, Normal Form of the elliptic Curves over the finite ring, Journal of Mathematics and system Sience, 4 (2014) 194-196.

[6] A. Tadmori, A. Chillali and M. Ziane, Coding over elliptic curves in the ring of characteristic two, International journal of Applied Mathemathics and Informatics, (Volume 8. 2014).

[7] J.H. SILVERMAN, The Arithmetic of Elliptic curves, Graduate Texts in Mathematcs, Springer, Volume 106(1985).2,19,20,21

[8] J.H. ~SILVERMAN, Advanced Topics in the Arithmetic of Elliptic curves, Graduate Texts in Mathematcs, Volume 151, Springer,(1994).

[9] W. Bosma and H. Lenstra, Complete system of two addition laws for elliptic curved, Journal of Number theory, (1995).

A multiagent proposal for a parallel system

A. Arteta, Juan Castellanos, C.Nuria Gomez

***Abstract*—** Cell computing works on computational models based on cell's behavior to process the information in a faster way than usual. Recently, numerous biological models have been developed and implemented. In particular cell membrane systems are structures that simulate the behavior and the evolution of membrane systems found in Nature. This paper proposes an application capable to supervise a cells membrane model through a multiagents system (MAS). By using multiagent systems the membrane models functionality improves. This creates the possibility of having scenarios where biological systems and multiagent systems work together to enhance the performance in bio-inspired models.

***Keywords*—** Multiagent systems, membrane computing, supervised learning, biological models, P-systems.

I. Introduction

Natural computing is the science which develops new computational models based on the way Nature works. These computational models are classified in three fields:

- Neuronal networks.
- Genetic Algorithms
- Biomolecular computation.

Membrane cells computing are included in bimolecular computation. A new logical computational device appears: The P-system. These P-systems are able to simulate the behavior of the membranes on living cells when absorbing nutrients, producing chemical reactions, dissolving membranes, etc)
Membrane computing formally represents, through the use of P-systems, the processes that take place inside of the living cells.
This works shows a multiagent system that it is able to conduct the functionality of a given P-system to resolve complex problems in fast way.

Contents of this paper are classified as follows:

- Introduction to cell membranes theory.
- Components of a particular multiagent system
- Proposal for a multiagent system which controls a transaction P-system
  - Example and code
- Conclusions and further work

II. CELL MEMBRANES THEORY

This section introduces the paradigm of the P-systems (Membrane computing). A P-system is a computational model inspired by the way the living cells interact with each other through their membranes. The elements of the membranes are called objects. A region / membrane can either contain objects or other membranes.

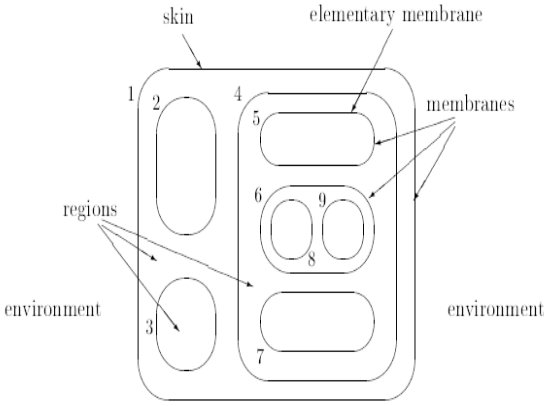


Fig. 1. The membrane's structure (left) represented in tree shape (right)

According to Păun 's definition, a transition P System of degree n, n > 1 is a construct: [[1]]

$$\Pi = \left(V, \mu, \omega_1, .., \omega_n, (R_1, \rho_1), .. (R_n, \rho_n), i_0\right)$$

Where:

1. V is an alphabet; its elements are called objects;
2. μ is a membrane structure of degree n, with the membranes and the regions labeled in a one-to-one manner with elements in a given set ; in this section we always use the labels 1,2,..,n;
3. $\omega_i \; 1 \le i \le n$, are strings from V^* representing multisets over V associated with the regions 1,2,..,n of μ

4. $R_i \, 1 \le i \le n$, are finite set of evolution rules over V associated with the regions 1,2,..,n of μ; ρ_i is a partial order over $R_i \, 1 \le i \le n$, specifying a priority relation among rules of R_i. An evolution rule is a pair (u,v) which we will usually write in the form $u \to v$ where u is a string over V and v=v' or v=v' δ where v' is a string over $(V \times \{here, out\}) \cup (V \times \{in_j \; 1 \le j \le n\})$, and δ is a special symbol not in. The length of u is called the radius of the rule $u \to v$
5. i_o is a number between 1 and n which specifies the output membrane of Π

Let *U* be a finite and not an empty set of objects and N the set of natural numbers. A *multiset of objects* is defined as a mapping:

$M : V \to \mathrm{N}$

$a_i \to u_1$

Where a_i is an object and u_i its multiplicity.

As it is well known, there are several representations for multisets of objects.

$M = \{(a_1, u_1), (a_2, u_2), (a_3, u_3)...\} = a_1^{u_1} \cdot a_2^{u_2} \cdot a_n^{u_n} .....$

Evolution rule with objects in *U* and targets in *T* is defined by $r = (m, c, \delta)$ where

$m \in M(V), c \in M(VxT)$ *and* $\delta \in \{to\ dissolve, not\ to\ dissolve\}$

From now on *'c'* will be referred to as the consequent of the evolution rule *'r'*

The *set of evolution rules* with *objects* in *V* and targets in *T* is represented by R *(U, T)*.

Rules are represented as:

$x \to y \quad or \quad x \to y\delta$ where x is a multiset of objects in M((*V*)xTar) where Tar ={here, in, out} and y is the consequent of the rule. When δ is equal to "dissolve", then the membrane will be dissolved making its set of evolution rules disappear.

P-systems evolve, which makes it change upon time; therefore it is a dynamic system. Every time that there is a change on the p-system a new transition is generated. The step from one transition to another one is defined as an evolutionary step, and the set of all evolutionary steps is named computation. Processes within the p-system will be acting in a massively parallel and non-deterministic manner. (Similar to the way the living cells process and combine information).

The whole info is processed successfully if:

1 The halt status is reached.

2 No more evolution rules can be applied.

III. MULTIAGENT SYSTEM TECHNOLOGY

Multi agent system can reach goals that are impossible for single agent systems (one agent system).
The main properties of multiagent systems are:

- Proactivity
- Autonomy

Formally speaking an agent is a real or virtual entity that:

1. Is able to act within a given environment.
2. Is able to communicate with other agents.
3. Have their own resources.
4. Is able to retrieve information and to (at least partially) know the environment.
5. Can reproduce.

According to the definition of agent, a multiagent system is defined as a system of computers which containing the following elements:

1. An environment *E*, which is a space.
2. A set of objects $O \in E$.
3. A set of agentes $A \in O$.
4. A set of relations *R* between objects and agents.
5. A set of operations that allows to the agents interact with the objects.

IV. AGENTS LEADING THE MEMBRANE MODEL: PROPOSAL

Once p-systems and multiagent systems are described separately, this section shows a way to create a multiagent system that supervise and control the membranes operations during the computation process. This agent supervised system will be referred as MSSA (Membrane system supervised by agents) from now on. As there are different components in a membrane system it is necessary now to establish how the multiagent system can manage the whole model. The following proposal is inspired in the model in [7] .

- In a p-system, given a set of membranes $M = \{m_i \mid i \in \mathrm{N}, 1 \le i \le n\}$ where m_i is a membrane, For any membrane mi it is necessary to define a single agent, . This can be defined as an

injective function.

$f_{agent} : M \to A$

$f_{agent}(m_i) = a_i \ \forall i \in \mathrm{N}\ i \le n,$ n number of membranes

- $a_i \in A$ In this way, every single agent is in charge of a single membrane. The agent is called membrane-agent
- The Multiset of objects within the region enclosed by the membrane $m_i \in M$ and the rules to to be applied on them are supervised by the membrane-agents
- All agents relate and communicate with each other.

In order to set up the synchronism in our system, a synchronization agent called $a_{sync} \in A$ is needed. This agent ensures a proper synchronization between the membrane agents.

The Multiagent system has to supervise the 2 major processes occurring in the membranes model. These are:

1. Dynamic behavior of the p-system (Computation and communication)
2. Synchronism between membranes.

Let us define each agent.

$f_{name} : \mathrm{N} \to Agent\ name$

1. $f_{name}(i) = a_i$
2. $f_{resource} : \{\omega_1, \omega_2, .., \omega_n\} \times \{R_1, R_2, .., R_n\} \to \mathrm{Resources}$
3. $f_{operation} : dynamictransition \to Agent\ behavior$

Every agent a_i is linkedto a set of resources called Res_i and set of operations .

Formally speaking, the multiagent system associated to a p-system with n membranes will have the set of resources as the union of all the objects, and the union of all the set of evolution rules included in every membrane i.e

$$\mathrm{Re}\,sources = \left[\bigcup_{i=1}^{n} R_i\right] \cup V$$

where R_i is the set of evolution rules which are included in the membrane i.

The multiagent system contains the agents a_i, resources R_i and a set of operations Op_i

Now let us define the multiagent system to evolve in order to control the transition P-system. This uses an operator that returns the status of the transition of a p-system to a specific time. In order to do this we create a new resource called *sync* which is defined as:

1. An integer (computing step)
2. A letter (Status)

The initial transition status is the integer 0
The System status for every step is defined as a letter (A,B or C), meaning as follows :
A. Rules election,
B. Objects consumption,.
C. Communication between membranes
The synchronizing agent ensures:

$sync_1 = 3A \quad sync_2 = 2B, \quad sync_3 = 4C$

$sync_i = sync_j \ \forall i \ne j \ i, j \in \mathrm{N}$

Initially. $sync_i = 0 \quad \forall i \le n \quad i \in \mathrm{N}$.

Example:

Let us have three membranes m_1, m_2, m_3 and m_1 containsm_2 whichcontainsm_3. The multiagent system has 3 membrane agents a_1, a_2, a_3 where a_1 represented as follows:

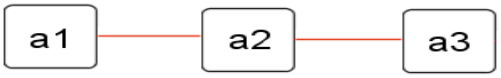


Fig. 2. Three membrane agents.

Below there is a diagram describing the relationship between a membrane system (left) and the supervisor Multiagent system.

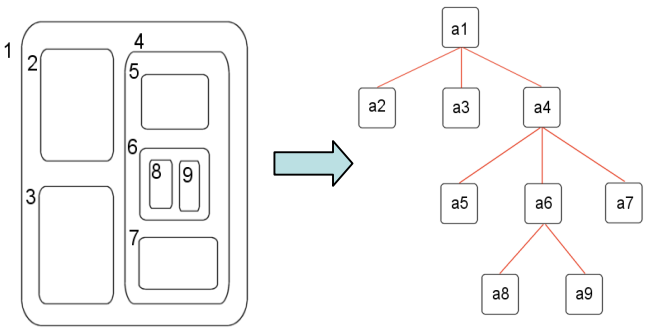


Fig. 3. MMAS description

Example of a computation process supervised by a Multiagent system

For this example a 3 membrane p-system has been chosen. This p-system is able to calculate the square of a given number. [5]. There are 2 evolution steps described. (the initial one and the final one) The P-system acts according the Multiagent system instructions.

a) Components

The P-system has:

- A set of membranes M= $\{m_1, m_2, m_3\}$
- An Alfabet V={a,b,c,d,e,f}
- A set of multiset of objects M(V)= $\{\omega_1 = \{\}, \omega_2 = \{\}, \omega_3 = \{af\}\}$ where ω_i is the multiset of objects within the region delimited by the membrane $m_i \quad \forall i \in \mathrm{N}\, i \leq 3$.
- A Multiset of evolution rules R(U,T)= $\{R_1(U,T)=\{e \rightarrow e_{out}\}, R_2(U,T)=\{b \rightarrow d, d \rightarrow de, (ff \rightarrow f > f \rightarrow \delta)\}, R_3(U,T)=\{a \rightarrow ab, a \rightarrow b\delta, f \rightarrow ff\}\}$

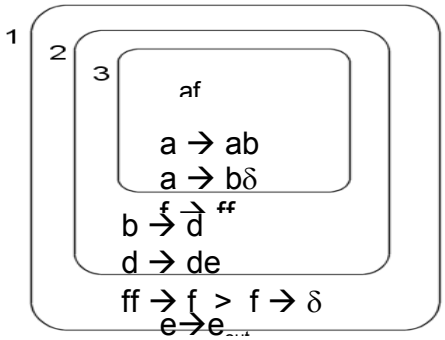


Fig 3 P-system calculating a random square number.

The set of agents A={ $\{a_1, a_2, a_3, a_{sync}\}$

The resources used by the agent a_i are referred as Res_i, these are the multiset of objects and set of evolution rules.

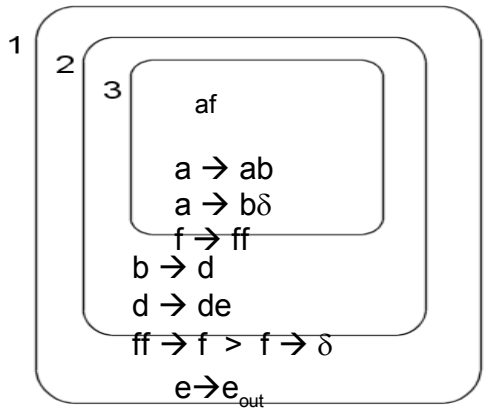
Now we go step by step

- Rules election
- Objects consumption
- Communication Stage

Initial Transition Status $sync_i = 0 \quad \forall i \leq 3 \quad i \in \mathrm{N}$

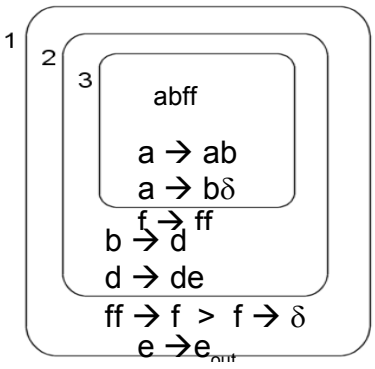
This condition is checked by the agent a_{sync}

In the transition status 1), the p-system evolves, In the Region 3 the rule number 1 and number three are applied.

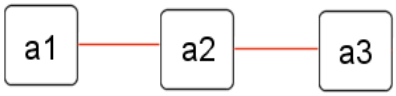


The agent a_{sync} makes sure that $sync_i = 1A \quad \forall i \leq 3 \quad i \in \mathrm{N}$. In every region the candidate rules to be applied are analyzed. Then every agent selects from its resources the rules to be applied.

In the mean time a_{sync} ensures the synchronization between regions. $sync_i = 1B \quad \forall i \leq 3 \quad i \in \mathrm{N}$. Here the agents execute the action "apply rules". agent a_3 makes the P-systems choose r_1 *and* r_3 $\mathrm{Re}\, s_3$ .After applying rules a_{sync} assures all the agents are synchronized. i.e. $sync_i = 1C \quad \forall i \leq 3 \quad i \in \mathrm{N}$. After the first computing step, the P-system looks like this:



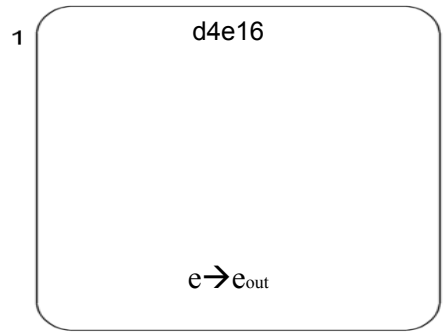
P-system co



The resources used by the agents are:

$$\left\{\bigcup_{i=1}^{n} \mathrm{Re}\, s_i\right\} = \left\{\begin{array}{l}\{[abff\,],[(a,ab),(a,b\delta),(f,ff)\},\\ \{[,[(b,\,d),\,(d,\,de),\,(ff,\,f),\,(f,\,\delta)]\},\,\{[[(e,\,e_{out})]\}\end{array}\right\}$$

In the end the program ends like follows:



Note. The number (4) besides the object (d) indicates the multiplicity of the object

The system returns 16= 4^2 . The agents have supervised every single computing step and have ensured a minimum number of operations to optimize the functionality of the membrane system. The main difference with the standard P-system is that elections are chosen in an intelligent way by the agents who supervise the membrane model

Example and code.

Rules R

Objects w

Dissolve = false

While NOT finished

waitsync

//EVOLVE

R' = Rules_election(w,R by agent)

w' = Rules_Application (w,R,P) by agent

waitsync

//COMUNNICATION between agents

w = COMMUNICATION(w')

dissolve = finish(w)

waitsync: Sinchronization between agents

Rules_election: Selection of the rules

Rules application: Application of the rules

communication (w) Communication between agents and exchange of the objects w

Finish: The computation is finished

V. CONCLUSIONS

This paper contributes with an implementation of a multiagent system that manages and supervises a cell membrane model. The agents along with their relations and resources are able to modify the membrane system functionality based on the agent's configuration. Therefore, the main idea of this work is to define a new model created from the original membrane computing technology. This update involves a multiagent system which is the one who takes care of the process. The most interesting part of this work is that a new way to define a biological model in terms of a multiagent supervised system has been created.

It is important to stress that the membrane systems described here is a generic one. The rules election, priorities between rules, etc are not fully described as this is not the main purpose of it. Moreover, the Multiagent system technology has been generally described to understand better the concepts of systems supervised by agents. In order to understand it better refer to [6] "An Introduction to Multiagent Systems", Wiley, 2002. or [8] "Multi-Agent Systems. An Introduction to Distributed Artificial Intelligence.

The code provided here is a proposal but it is not the main goal of this work. By formally defining a multiagent system it would be possible to take full advantage of the Multiagent system technology and apply it into cells membrane system or any other biological model.

Refer to [2] for checking links between Multiagents and P-systems.

Membranes agents are independent and autonomous which can modify the entire functionality of the membrane systems. The example provided shows how the membrane system can improve performance when a proper set of agents is chosen. The agents are the intelligent entities who supervise the entire membrane system and optimize the functionality.

Thus, the whole idea of this proposal is to improve and take all the possible advantages of the Multi agents systems to apply them into biological systems and make them work better in terms of performance.

.

REFERENCES

[1] "Computing with Membranes", Journal of Computer and System Sciences, 61(2000), and Turku Center of Computer Science-TUCS Report nº 208, 1998.

[2] A Proposal of Multi-Agent Simulation System for Membrane Computing Devices Giovanni Acampora, Member, IEEE and Vincenzo Loia Member, IEEE 2007 IEEE Congress on Evolutionary Computation (CEC 2007).

[3] Membranes as Multi-agent Systems: an Application to Dialogue Modelling Gemma Bel-Enguix and Dolores Jimenez Lopez Bel-Enguix. G., Lopez, D.,T., 2006, in IFIP International Federation for Infonnation Processing, Volume 218 Professional.

[4] l "Algorithm for Application of Evolution Rules based on linear diofantic equations" Synasc 2008, Timisoara Romania September 2008[1]

[5] A. Syropoulos, E.G. Mamatas, P.C. Allilones, K.T. Sotiriades "A

["Structures and Bio-language to Simulate Transition P Systems on Digital Computers," Multiset Processing

[6] An Introduction to MultiAgent Systems", Wiley, 2002.

[7] Analysys of a P-System under a Mutiagent System perspetive, International Book Series "Information Science and Computing" pag 117 Varna Bulgaria 2009

[8] [8] "Multi-Agent Systems. An Introduction to Distributed Artificial Intelligence," Addison Wesley, 1999.pp. 876—880. Available: http://www.halcyon.com/pub/journals/21ps03-vidmar

Exact solutions of the nonlinear Schrödinger equation by generalizing exp-function method

Sheng Zhang and Zhao-Yu Wang

***Abstract*—**In this paper, a generalized exp-function method is proposed for constructing exact solutions of the complex nonlinear partial differential equations. As one application of the generalized method, the nonlinear Schrödinger equation is considered and new single-, double- and three-wave solutions with parameters are obtained, from which a uniform formula of *N*-wave solution is derived. Thanks to the arbitrariness of the included parameters, not only possess these obtained multiwave solutions enrich structures like the breather solutions and envelope solutions, but also high-wave solution can give all the low-wave solutions. It is shown that the generalized exp-function method combined with appropriate ansätz may provide with a straightforward, effective and alternative method for constructing multiwave solutions of some other complex nonlinear partial differential equations.

***Keywords*—**Exp-function method, Multiwave solution, Nonlinear Schrödinger equation, Breather solution, Envelope solution.

I. Introduction

Searching for exact solutions of nonlinear partial differential equations (PDEs) plays an important role in the study of some nonlinear phenomena involved in many fields from physics to biology, chemistry, mechanics, etc. In the past several decades, there has been significant progression in the development of many methods for solving nonlinear PDEs, such as those in [1]-[14]. With the development of soliton theory, finding multiwave solutions of nonlinear PDEs has gradually developed into a significant direction in nonlinear science. Since proposed by He and Wu in 2006, the exp-function method [15] has been applied to many kinds of nonlinear equations [16]-[30]. More and more studies show that the exp-function method is available for many nonlinear PDEs and can be used to construct multiple types of exact solutions due to its more general ansätz with free parameters.

The present paper is motivated by the desire to generalize the exp-function method to construct multiwave solutions of the nonlinear Schrödinger equation [31]:

$$\mathrm{i}u_t + u_{xx} + |u|^2 u = 0, \tag{1}$$

where i denotes the imaginary number unit, $|u|$ is the modules of u.

II. Methodology

In this section, we describe the basic idea of the generalized exp-function method with a general ansätz for multiwave solutions of the given complex nonlinear PDE, say, in two real variables x and t:

$$P(u, u_t, u_x, u_{tx}, u_{tt}, u_{xx}, \cdots) = 0, \tag{2}$$

where P is a polynomial of u and its derivatives, otherwise, a suitable transformation can transform Eq. (2) into such an equation. The generalized exp-function method for single-wave solution is based on the assumption that Eq. (2) has a solution in the form:

$$u(x,t) = \frac{\sum_{i_1=0}^{p_1}\sum_{i_2=0}^{p_2} a_{i_1 i_2} \mathrm{e}^{i_1\xi_1 + i_2\xi_1^*}}{\sum_{j_1=0}^{q_1}\sum_{j_2=0}^{q_2} b_{j_1 j_2} \mathrm{e}^{i_1\xi_1 + i_2\xi_1^*}}, \tag{3}$$

where $\xi_1 = k_1 x + c_1 t + \omega_1$, $\xi_1^* = k_1^* x + c_1^* t + \omega_1^*$ denotes the complex conjugate of ξ_1; $a_{i_1 i_2}$, $b_{j_1 j_2}$, c_1 or c_1^*, k_1 or k_1^* are unknown complex constants to be determined; ω_1 or ω_1^* is an arbitrary complex constant; and the real values of p_1, p_2, q_1, q_2 can be determined by balancing the linear term of highest order in Eq. (2) with the highest order nonlinear term.

In order to seek *N*-wave solution for arbitrary integer $N > 1$, we generalize Eq. (3) to the following form:

$$u(x,t) = \frac{\sum_{i_1=0}^{p_1}\sum_{i_2=0}^{p_2}\cdots\sum_{i_{2N}=0}^{p_{2N}} a_{i_1 i_2 \cdots i_{2N}} \mathrm{e}^{\sum_{g=1}^{N}(i_g\xi_g + i_{g+N}\xi_g^*)}}{\sum_{j_1=0}^{q_1}\sum_{j_2=0}^{q_2}\cdots\sum_{j_{2N}=0}^{q_{2N}} b_{j_1 j_2 \cdots j_{2N}} \mathrm{e}^{\sum_{g=1}^{N}(j_g\xi_g + j_{g+N}\xi_g^*)}}, \tag{4}$$

where $\xi_g = k_g x + c_g t + \omega_g$, $\xi_g^* = k_g^* x + c_g^* t + \omega_g^*$ is the complex conjugate of ξ_g; $a_{i_1 i_2 \cdots i_{2N}}$, $b_{j_1 j_2 \cdots j_{2N}}$, c_g or c_g^*, k_g or k_g^* are undetermined constants; ω_g or ω_g^* are arbitrary

This work was supported by the Natural Science Foundation of Liaoning Province (L2012404) of China, the PhD Start-up Funds of Bohai University (bsqd2013025) and Liaoning Province of China (20141137), the Liaoning BaiQianWan Talents Program (2013921055) and the Natural Science Foundation of China (11371071).

S. Zhang is with the School of Mathematics and Physics, Bohai University, Jinzhou 121013, PR China (corresponding author to provide phone: 086-416-3400149; e-mail: szhangchina@ 126.com).

Z. Y. Wang is with the School of Mathematics and Physics, Bohai University, Jinzhou 121013, PR China (e-mail: 1174833500@qq.com)

complex constants; and the real values of p_1, p_2, $\cdots$, p_{2N}, q_1, q_2, $\cdots$, q_{2N} are embedded integers.

If give $N = 2$ to Eq. (4), one has:

$$u(x,t) = \frac{\sum_{i_1=0}^{p_1}\sum_{i_2=0}^{p_2}\sum_{i_3=0}^{p_3}\sum_{i_4=0}^{p_4} a_{i_1 i_2 i_3 i_4} \mathrm{e}^{\sum_{g=1}^{2}(i_g\xi_g + i_{g+2}\xi_g^*)}}{\sum_{j_1=0}^{q_1}\sum_{j_2=0}^{q_2}\sum_{j_3=0}^{q_3}\sum_{j_4=0}^{q_4} b_{j_1 j_2 j_3 j_4} \mathrm{e}^{\sum_{g=1}^{2}(j_g\xi_g + j_{g+2}\xi_g^*)}}, \quad (5)$$

which can be used to construct double-wave solution of Eq. (2).

When $N = 3$, Eq. (4) gives:

$$u(x,t) = \frac{\sum_{i_1=0}^{p_1}\sum_{i_2=0}^{p_2}\sum_{i_3=0}^{p_3}\sum_{i_4=0}^{p_4}\sum_{i_5=0}^{p_5}\sum_{i_6=0}^{p_6} a_{i_1 i_2 i_3 i_4 i_5 i_6} \mathrm{e}^{\sum_{g=1}^{3} i_g (i_g\xi_g + i_{g+3}\xi_g^*)}}{\sum_{j_1=0}^{q_1}\sum_{j_2=0}^{q_2}\sum_{j_3=0}^{q_3}\sum_{j_4=0}^{q_4}\sum_{j_5=0}^{q_5}\sum_{j_6=0}^{q_6} b_{j_1 j_2 j_3 j_4 j_5 j_6} \mathrm{e}^{\sum_{g=1}^{3}(j_g\xi_g + j_{g+3}\xi_g^*)}}, \quad (6)$$

which is effective for obtaining three-wave solution of Eq. (2). Substituting Eq. (5) into Eq. (2), then equating to zero each coefficient of the same order power of the exponential functions yields a set of equations. Solving the set of equations, we can determine the double-wave solution, and the following three-wave solution by means of Eq. (6), provided they exist.

III. Multiwave solutions

In this section, let us apply the generalized exp-function method described in Section 2 to the nonlinear Schrödinger equation (1).

Firstly, we suppose that Eq. (1) has a single-wave solution in the form:

$$u = \frac{a_0 + a_1 \mathrm{e}^{\xi_1} + a_2 \mathrm{e}^{\xi_1^*} + a_3 \mathrm{e}^{\xi_1+\xi_1^*}}{1 + b_1 \mathrm{e}^{\xi_1} + b_2 \mathrm{e}^{\xi_1^*} + b_3 \mathrm{e}^{\xi_1+\xi_1^*}}. \quad (7)$$

Substituting Eq. (7) into Eq. (1), and using Mathematica, then equating to zero each coefficient of the same order power of $\mathrm{e}^{\theta\xi_1+\vartheta\xi_1^*}$ ($\theta,\vartheta = 1,2,3,4$) yields a set of equations for a_0, a_0^*, a_1, a_1^*, a_2, a_2^*, a_3, a_3^*, a_4, a_4^*, b_1, b_1^*, b_2, b_2^*, b_3, b_3^*, c_1, c_1^*, k_1, k_1^*, three simplest equations of which read:

$$a_1^2 a_2^* b_1 = 0,\ a_1^* a_2^2 b_2 = 0,\ a_3^2 a_3^* b_3 = 0. \quad (8)$$

Supposing $b_3 \neq 0$, from the third one of Eq. (8) we have $a_3 = 0$. In this case, the foregoing set of equations give two simple equations:

$$a_2 b_3^2 b_3^* (k_1^2 - \mathrm{i}c_1) = 0,\ a_1 b_3^2 b_3^* (k_1^{*2} - \mathrm{i}c_1^*) = 0, \quad (9)$$

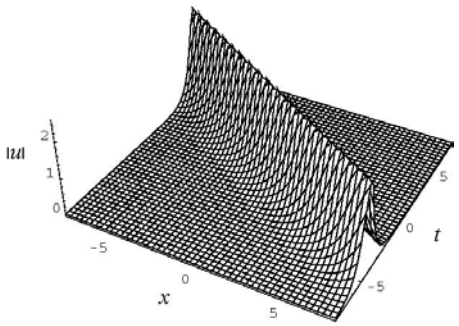
If $a_1 \neq 0$, Eq. (9) means that $c_1 = \mathrm{i}k_1^2$ and $a_2 = 0$. At the same time, $a_2 = 0$ justly solves the first two equations of Eq. (8). Thus the foregoing set of equations are further simplified and give a simple equation:

$$-2a_1 b_1 b_3 b_3^* k_1^{*2} = 0, \quad (10)$$

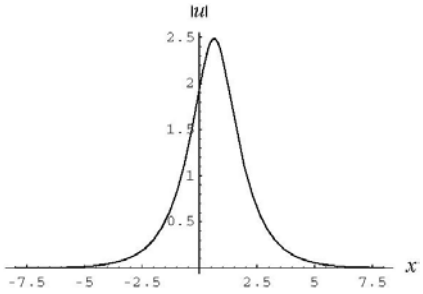
which let one choose $b_1 = 0$. Otherwise, $k_1 = 0$ and hence only a trivial solution of Eq. (1) can be obtained if it exists. This is not the one we expect. Substituting $b_1 = 0$ into the foregoing set of equations, we have $a_0^* a_1 = 0$ and then $a_0 = 0$ because that $a_1 \neq 0$ is supposed beforehand. Using $a_0 = 0$ to simplify the foregoing set of equations, we have

$$-2a_1 b_2^2 k_1 k_1^* = 0, \quad (11)$$

which shows $b_2 = 0$ is the only choice.

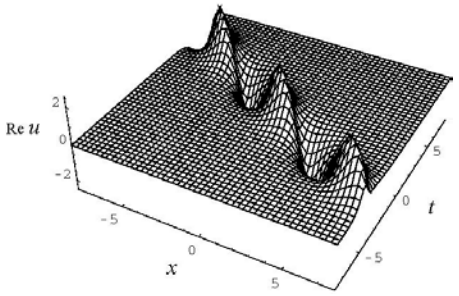


(1a) $x \in [-8,8]$, $t \in [-8,8]$

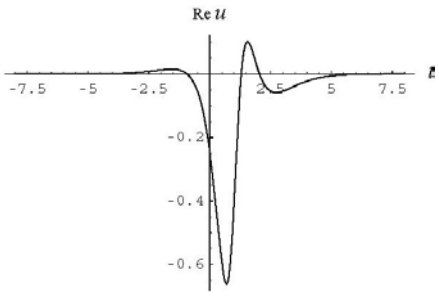


(1b) $x \in [-8,8]$, $t = 0$

Fig. 1. Modulus of single-wave solution (15).



(2a) $x \in [-8,8]$, $t \in [-8,8]$



(2b) $x = -1$, $t \in [-8,8]$

Fig. 2. Real part of single-wave solution (15).

Through a series of algebraic simplifications, the foregoing set of equations give the last two equations:

$$a_1^2 a_1^* - 2a_1 b_3 k_1^2 - 4a_1 b_3 k_1 k_1^* - 2a_1 b_3 k_1^{*2} = 0, \quad (12)$$

$$a_1^2 a_1^* b_3 - 2a_1 b_3 b_3^* k_1^2 - 4a_1 b_3 b_3^* k_1 k_1^* - 2a_1 b_3 b_3^* k_1^{*2} = 0 . \quad (13)$$

Solving Eqs. (12) and (13), we have

$$b_3 = \frac{a_1 a_1^*}{2(k_1 + k_1^*)^2}, \quad (14)$$

and therefore obtain the single-wave solution of Eq. (1):

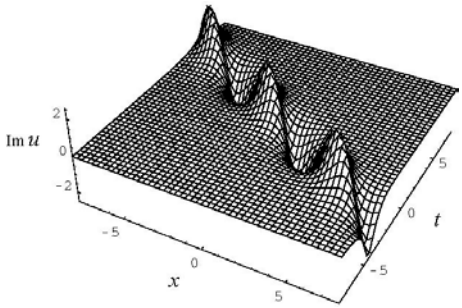
$$u = \frac{a_1 \mathrm{e}^{\xi_1}}{1 + a_1 a_1^* \mathrm{e}^{\xi_1 + \xi_1^* + A_{13}}}, \quad (15)$$

with

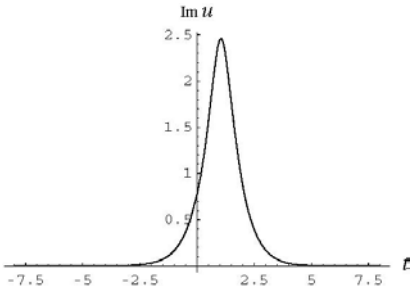
$$\xi_1 = k_1 x + \mathrm{i} k_1^2 t + \omega_1 , \quad \xi_1^* = k_1^* x - \mathrm{i} k_1^{*2} t + \omega_1^* , \quad (16)$$

$$\mathrm{e}^{A_{13}} = \frac{1}{2(k_1 + k_1^*)^2} . \quad (17)$$

where a_1 or a_1^*, k_1 or k_1^*, ω_1 or ω_1^* are arbitrary complex constants.



(3a) $x \in [-8,8]$, $t \in [-8,8]$



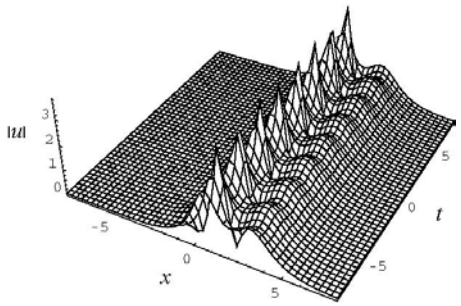
(3b) $x = -1$, $t \in [-8,8]$

Fig. 3. Imaginary part of single-wave solution.

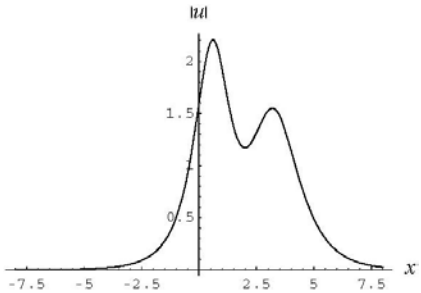
In the other hand, if $b_3 = 0$, from Eq. (8) we can see that the third equation holds automatically but $a_1 = 0$ or $a_2 = 0$. Otherwise, $b_1 = 0$ and $b_2 = 0$, this will lead to a trivial solution of Eq. (1). For the case of $a_1 = 0$, $b_1 \neq 0$ and $b_2 \neq 0$, we obtain $a_3 = 0$. Then the subsequent computation shows either $a_0 = 0$ or $c_1 = -\mathrm{i} k_1^2$, however both cases result in a trivial solution of Eq. (1). By the similar analysis, for both the case of $a_1 = 0$, $b_1 \neq 0$, $b_2 = 0$ and the case of $a_1 = 0$, $b_1 = 0$, $b_2 \neq 0$, we can obtain only a trivial solution of Eq. (1). Similarly for the case of $a_0 = 0$, $b_1 \neq 0$, $b_2 \neq 0$ or $a_2 = 0$, $b_1 \neq 0$, $b_2 = 0$ or $a_1 = 0$, $b_1 = 0$, $b_2 \neq 0$, Eq. (1) has only a trivial solution. Without loss of generality, we reconsider Eq. (9) under the assumption of $a_1 = 0$ and $b_3 \neq 0$, but there is not a non-trivial solution of Eq. (1) can be obtained.

Taking into consideration the constructional features of the single-wave solution (15), we next suppose that Eq. (1) has a double-wave solution in the form:

$$u = \frac{a_1 \mathrm{e}^{\xi_1} + a_2 \mathrm{e}^{\xi_2} + a_3 \mathrm{e}^{\xi_1 + \xi_2 + \xi_1^*} + a_4 \mathrm{e}^{\xi_1 + \xi_2 + \xi_2^*}}{1 + b_1 \mathrm{e}^{\xi_1 + \xi_1^*} + b_2 \mathrm{e}^{\xi_1 + \xi_2^*} + b_3 \mathrm{e}^{\xi_2 + \xi_1^*} + b_4 \mathrm{e}^{\xi_2 + \xi_2^*} + b_5 \mathrm{e}^{\xi_1 + \xi_2 + \xi_1^* + \xi_2^*}} . \quad (18)$$



(4a) $x \in [-8,8]$, $t \in [-8,8]$



(4b) $x \in [-8,8]$, $t = 0$

Fig. 4. Modulus of double-wave solution (24).

Substituting Eq. (18) into Eq. (1), and using Mathematica, then equating to zero each coefficient of the same order power of $\mathrm{e}^{\theta \xi_1 + \vartheta \xi_2 + \mu \xi_1^* + \rho \xi_2^*}$ $(\theta, \vartheta, \mu, \rho = 1,2,3,4)$ yields a set of equations for a_1, a_1^*, a_2, a_2^*, a_3, a_3^*, a_4, a_4^*, b_1, b_1^*, b_2, b_2^*, b_3, b_3^*, b_4, b_4^*, b_5, b_5^*, c_1, c_1^*, c_2, c_2^*, k_1, k_1^*, k_2, k_2^*. Solving the set of equations, we have:

$$a_3 = \frac{a_1 a_2 a_1^* (k_1 - k_2)^2}{2(k_1 + k_1^*)^2 (k_2 + k_1^*)^2}, \quad a_4 = \frac{a_1 a_2 a_2^* (k_1 - k_2)^2}{2(k_1 + k_2^*)^2 (k_2 + k_2^*)^2}, \quad (19)$$

$$b_1 = \frac{a_1 a_1^*}{2(k_1 + k_1^*)^2}, \quad b_2 = \frac{a_1 a_2^*}{2(k_1 + k_2^*)^2}, \quad (20)$$

$$b_3 = \frac{a_2 a_1^*}{2(k_2 + k_1^*)^2}, \quad b_4 = \frac{a_2 a_2^*}{2(k_2 + k_2^*)^2}, \quad (21)$$

$$b_5 = \frac{a_1 a_2 a_1^* a_2^* (k_1 - k_2)^2 (k_1^* - k_2^*)^2}{4(k_1 + k_1^*)^2 (k_2 + k_1^*)^2 (k_1 + k_2^*)^2 (k_2 + k_2^*)^2}, \quad (22)$$

$$c_1 = \mathrm{i} k_1^2 , \quad c_2 = \mathrm{i} k_2^2 , \quad (23)$$

and then obtain the double-wave solution of Eq. (1) as follows:

$$u = \frac{f(\xi_1, \xi_2, \xi_1^*, \xi_2^*)}{g(\xi_1, \xi_2, \xi_1^*, \xi_2^*)}, \quad (24)$$

with

$$f(\xi_1, \xi_2, \xi_1^*, \xi_2^*) = a_1 \mathrm{e}^{\xi_1} + a_2 \mathrm{e}^{\xi_2} + a_1 a_2 a_1^* \mathrm{e}^{\xi_1 + \xi_2 + \xi_1^* + A_{12} + A_{13} + A_{23}}$$
$$+ a_1 a_2 a_2^* \mathrm{e}^{\xi_1 + \xi_2 + \xi_2^* + A_{12} + A_{14} + A_{24}} , \quad (25)$$

$$g(\xi_1,\xi_2,\xi_1^*,\xi_2^*)=1+a_1a_1^*\mathrm{e}^{\xi_1+\xi_1^*+A_{13}}+a_1a_2^*\mathrm{e}^{\xi_1+\xi_2^*+A_{14}}$$
$$+a_2a_1^*\mathrm{e}^{\xi_2+\xi_1^*+A_{23}}+a_2a_2^*\mathrm{e}^{\xi_2+\xi_2^*+A_{24}}$$
$$+a_1a_2a_1^*a_2^*\mathrm{e}^{\xi_1+\xi_2+\xi_1^*+\xi_2^*+A_{12}+A_{13}+A_{14}+A_{23}+A_{24}+A_{34}}, \quad (26)$$

$$\xi_1=k_1x+\mathrm{i}k_1^2t+\omega_1,\ \xi_2=k_2x+\mathrm{i}k_2^2t+\omega_2, \quad (27)$$

$$\xi_1^*=k_1^*x-\mathrm{i}k_1^{*2}t+\omega_1^*,\ \xi_2^*=k_2^*x-\mathrm{i}k_2^{*2}t+\omega_2^*, \quad (28)$$

$$\mathrm{e}^{A_{12}}=2(k_1-k_2)^2,\ \mathrm{e}^{A_{j,l+2}}=\frac{1}{2(k_j+k_l^*)^2},\ (j,l=1,2), \quad (29)$$

$$\mathrm{e}^{A_{34}}=2(k_1^*-k_2^*)^2, \quad (30)$$

where a_1, a_1^*, a_2, a_2^*, c_1, c_1^*, c_2, c_2^*, k_1, k_1^*, k_2, k_2^*, ω_1, ω_1^*, ω_2, ω_2^* are arbitrary complex constants.

If we suppose Eq. (1) has a three-wave solution in the form:

$$u=\frac{f(\xi_1,\xi_2,\xi_3,\xi_1^*,\xi_2^*,\xi_3^*)}{g(\xi_1,\xi_2,\xi_3,\xi_1^*,\xi_2^*,\xi_3^*)}, \quad (31)$$

with

$$f(\xi_1,\xi_2,\xi_3,\xi_1^*,\xi_2^*,\xi_3^*)=a_1\mathrm{e}^{\xi_1}+a_2\mathrm{e}^{\xi_2}+a_3\mathrm{e}^{\xi_3}+a_4\mathrm{e}^{\xi_1+\xi_2+\xi_1^*}$$
$$+a_5\mathrm{e}^{\xi_1+\xi_2+\xi_2^*}+a_6\mathrm{e}^{\xi_1+\xi_2+\xi_3^*}+a_7\mathrm{e}^{\xi_1+\xi_3+\xi_1^*}+a_8\mathrm{e}^{\xi_1+\xi_3+\xi_2^*}$$
$$+a_9\mathrm{e}^{\xi_1+\xi_3+\xi_3^*}+a_{10}\mathrm{e}^{\xi_2+\xi_3+\xi_1^*}+a_{11}\mathrm{e}^{\xi_2+\xi_3+\xi_2^*}$$
$$+a_{12}\mathrm{e}^{\xi_2+\xi_3+\xi_3^*}+a_{13}\mathrm{e}^{\xi_1+\xi_2+\xi_3+\xi_1^*+\xi_2^*}+a_{14}\mathrm{e}^{\xi_1+\xi_2+\xi_3+\xi_1^*+\xi_3^*}$$
$$+a_{15}\mathrm{e}^{\xi_1+\xi_2+\xi_3+\xi_2^*+\xi_3^*}, \quad (32)$$

$$g(\xi_1,\xi_2,\xi_3,\xi_1^*,\xi_2^*,\xi_3^*)=1+b_1\mathrm{e}^{\xi_1+\xi_1^*}+b_2\mathrm{e}^{\xi_1+\xi_2^*}+b_3\mathrm{e}^{\xi_1+\xi_3^*}$$
$$+b_4\mathrm{e}^{\xi_2+\xi_1^*}+b_5\mathrm{e}^{\xi_2+\xi_2^*}+b_6\mathrm{e}^{\xi_2+\xi_3^*}+b_7\mathrm{e}^{\xi_3+\xi_1^*}+b_8\mathrm{e}^{\xi_3+\xi_2^*}$$
$$+b_9\mathrm{e}^{\xi_3+\xi_3^*}+b_{10}\mathrm{e}^{\xi_1+\xi_2+\xi_1^*+\xi_2^*}+b_{11}\mathrm{e}^{\xi_1+\xi_2+\xi_1^*+\xi_3^*}$$
$$+b_{12}\mathrm{e}^{\xi_1+\xi_2+\xi_2^*+\xi_3^*}+b_{13}\mathrm{e}^{\xi_1+\xi_3+\xi_1^*+\xi_2^*}+b_{14}\mathrm{e}^{\xi_1+\xi_3+\xi_1^*+\xi_3^*}$$
$$+b_{15}\mathrm{e}^{\xi_1+\xi_3+\xi_2^*+\xi_3^*}+b_{16}\mathrm{e}^{\xi_2+\xi_3+\xi_1^*+\xi_2^*}+b_{17}\mathrm{e}^{\xi_2+\xi_3+\xi_1^*+\xi_3^*}$$
$$+b_{18}\mathrm{e}^{\xi_2+\xi_3+\xi_2^*+\xi_3^*}+b_{19}\mathrm{e}^{\xi_1+\xi_2+\xi_3+\xi_1^*+\xi_2^*+\xi_3^*}. \quad (33)$$

By the similar manipulations mentioned above, we can obtain

$$a_4=\frac{a_1a_2a_1^*(k_1-k_2)^2}{2(k_1+k_1^*)^2(k_2+k_1^*)^2},\ a_5=\frac{a_1a_2a_2^*(k_1-k_2)^2}{2(k_1+k_2^*)^2(k_2+k_2^*)^2}, \quad (34)$$

$$a_6=\frac{a_1a_2a_3^*(k_1-k_2)^2}{2(k_1+k_3^*)^2(k_2+k_3^*)^2},\ a_7=\frac{a_1a_3a_1^*(k_1-k_3)^2}{2(k_1+k_1^*)^2(k_3+k_1^*)^2}, \quad (35)$$

$$a_8=\frac{a_1a_3a_2^*(k_1-k_3)^2}{2(k_1+k_2^*)^2(k_3+k_2^*)^2},\ a_9=\frac{a_1a_3a_3^*(k_1-k_3)^2}{2(k_1+k_3^*)^2(k_3+k_3^*)^2}, \quad (36)$$

$$a_{10}=\frac{a_2a_3a_1^*(k_2-k_3)^2}{2(k_2+k_1^*)^2(k_3+k_1^*)^2},\ a_{11}=\frac{a_2a_3a_2^*(k_2-k_3)^2}{2(k_2+k_2^*)^2(k_3+k_2^*)^2}, \quad (37)$$

$$a_{12}=\frac{a_2a_3a_3^*(k_2-k_3)^2}{2(k_2+k_3^*)^2(k_3+k_3^*)^2}, \quad (38)$$

$$a_{13}=\frac{a_1a_2a_3a_1^*a_2^*(k_1-k_2)^2(k_1-k_3)^2(k_2-k_3)^2(k_1^*-k_2^*)^2}{4(k_1+k_1^*)^2(k_2+k_1^*)^2(k_3+k_1^*)^2(k_1+k_2^*)^2(k_2+k_2^*)^2(k_3+k_2^*)^2} \quad (39)$$

$$a_{14}=\frac{a_1a_2a_3a_1^*a_3^*(k_1-k_2)^2(k_1-k_3)^2(k_2-k_3)^2(k_1^*-k_3^*)^2}{4(k_1+k_1^*)^2(k_2+k_1^*)^2(k_3+k_1^*)^2(k_1+k_3^*)^2(k_2+k_3^*)^2(k_3+k_3^*)^2} \quad (40)$$

$$a_{15}=\frac{a_1a_2a_3a_2^*a_3^*(k_1-k_2)^2(k_1-k_3)^2(k_2-k_3)^2(k_2^*-k_3^*)^2}{4(k_1+k_2^*)^2(k_2+k_2^*)^2(k_3+k_2^*)^2(k_1+k_3^*)^2(k_2+k_3^*)^2(k_3+k_3^*)^2} \quad (41)$$

$$b_1=\frac{a_1a_1^*}{2(k_1+k_1^*)^2},\ b_2=\frac{a_1a_2^*}{2(k_1+k_2^*)^2},\ b_3=\frac{a_1a_3^*}{2(k_1+k_3^*)^2}, \quad (42)$$

$$b_4=\frac{a_2a_1^*}{2(k_2+k_1^*)^2},\ b_5=\frac{a_2a_2^*}{2(k_2+k_2^*)^2},\ b_6=\frac{a_2a_3^*}{2(k_2+k_3^*)^2}, \quad (43)$$

$$b_7=\frac{a_3a_1^*}{2(k_3+k_1^*)^2},\ b_8=\frac{a_3a_2^*}{2(k_3+k_2^*)^2},\ b_9=\frac{a_3a_3^*}{2(k_3+k_3^*)^2}, \quad (44)$$

$$b_{10}=\frac{a_1a_2a_1^*a_2^*(k_1-k_2)^2(k_1^*-k_2^*)^2}{4(k_1+k_1^*)^2(k_2+k_1^*)^2(k_1+k_2^*)^2(k_2+k_2^*)^2}, \quad (45)$$

$$b_{11}=\frac{a_1a_2a_1^*a_3^*(k_1-k_2)^2(k_1^*-k_3^*)^2}{4(k_1+k_1^*)^2(k_2+k_1^*)^2(k_1+k_3^*)^2(k_2+k_3^*)^2}, \quad (46)$$

$$b_{12}=\frac{a_1a_2a_2^*a_3^*(k_1-k_2)^2(k_2^*-k_3^*)^2}{4(k_1+k_2^*)^2(k_2+k_2^*)^2(k_1+k_3^*)^2(k_2+k_3^*)^2}, \quad (47)$$

$$b_{13}=\frac{a_1a_3a_1^*a_2^*(k_1-k_3)^2(k_1^*-k_2^*)^2}{4(k_1+k_1^*)^2(k_3+k_1^*)^2(k_1+k_2^*)^2(k_3+k_2^*)^2}, \quad (48)$$

$$b_{14}=\frac{a_1a_3a_1^*a_3^*(k_1-k_3)^2(k_1^*-k_3^*)^2}{4(k_1+k_1^*)^2(k_3+k_1^*)^2(k_1+k_3^*)^2(k_3+k_3^*)^2}, \quad (49)$$

$$b_{15}=\frac{a_1a_3a_2^*a_3^*(k_1-k_3)^2(k_2^*-k_3^*)^2}{4(k_1+k_2^*)^2(k_3+k_2^*)^2(k_1+k_3^*)^2(k_3+k_3^*)^2}, \quad (50)$$

$$b_{16}=\frac{a_2a_3a_1^*a_2^*(k_2-k_3)^2(k_1^*-k_2^*)^2}{4(k_2+k_1^*)^2(k_3+k_1^*)^2(k_2+k_2^*)^2(k_3+k_2^*)^2}, \quad (51)$$

$$b_{17}=\frac{a_2a_3a_1^*a_3^*(k_2-k_3)^2(k_1^*-k_3^*)^2}{4(k_2+k_1^*)^2(k_3+k_1^*)^2(k_3+k_1^*)^2(k_3+k_3^*)^2}, \quad (52)$$

$$b_{18}=\frac{a_2a_3a_2^*a_3^*(k_2-k_3)^2(k_2^*-k_3^*)^2}{4(k_2+k_2^*)^2(k_3+k_2^*)^2(k_3+k_2^*)^2(k_3+k_3^*)^2}, \quad (53)$$

$$b_{19}=[a_1a_2a_3a_1^*a_2^*a_3^*(k_1-k_2)^2(k_1-k_3)^2(k_2-k_3)^2(k_1^*-k_2^*)^2$$
$$\times(k_1^*-k_3^*)^2(k_2^*-k_3^*)^2]/[8(k_1+k_1^*)^2(k_2+k_1^*)^2$$
$$\times(k_3+k_1^*)^2(k_2+k_1^*)^2(k_2+k_2^*)^2(k_2+k_3^*)^2$$
$$\times(k_3+k_1^*)^2(k_3+k_2^*)^2(k_3+k_3^*)^2], \quad (54)$$

$$c_1=\mathrm{i}k_1^2,\ c_2=\mathrm{i}k_2^2,\ c_3=\mathrm{i}k_3^2. \quad (55)$$

With the help of Eqs. (32)–(55), the three-wave solution (31) can be finally determined.

If we continue to construct the *N*-wave solutions for any $N\geq 4$, the following similar manipulations becomes rather complicated since equating to zero the coefficients of the exponential functions implies a highly nonlinear system [19].

ISBN: 978-1-61804-292-7

Fortunately, by analyzing the obtained solutions (15), (21) and (26) and introducing the following notations:

$$\xi_j = k_j x + \mathrm{i}k_j^2 t + \omega_j, \tag{56}$$

$$\xi_{N+j} = \xi_j^* = k_j^* x - \mathrm{i}k_j^{*2} t + \omega_j^*, \ (j = 1,2,\cdots,N), \tag{57}$$

$$\mathrm{e}^{A_{jl}} = 2(k_j - k_l)^2, \ (j < l = 2,3,\cdots,N), \tag{58}$$

$$\mathrm{e}^{A_{j,N+l}} = \frac{1}{2(k_j + k_l^*)^2}, \ (j,l = 1,2,\cdots,N), \tag{59}$$

$$\mathrm{e}^{A_{N+j,N+l}} = \mathrm{e}^{A_{j,l}^*} = 2(k_j^* - k_l^*)^2, \ (j < l = 2,3,\cdots,N), \tag{60}$$

$$a_{N+j} = a_j^*, \ (j = 1,2,\cdots,N), \tag{61}$$

we obtain a uniform formula of the *N*-wave solution of Eq. (1):

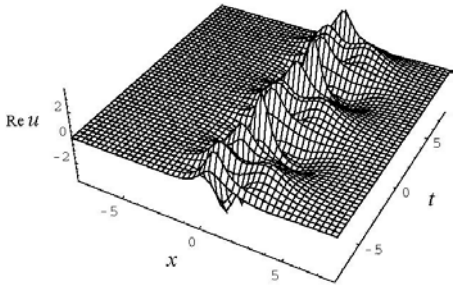
$$u = \frac{\sum_{\mu=0,1} B_1(\mu) \prod_{j=1}^{2N} a_j^{\mu_j} \mathrm{e}^{\sum_{j=1}^{2N} \mu_j \xi_j + \sum_{j=1}^{2N} \mu_j \mu_l A_{jl}}}{\sum_{\mu=0,1} B_2(\mu) \prod_{j=1}^{2N} a_j^{\mu_j} \mathrm{e}^{\sum_{j=1}^{2N} \mu_j \xi_j + \sum_{j=1}^{2N} \mu_j \mu_l A_{jl}}}, \tag{62}$$

where the summation $\Sigma_{\mu=0,1}$ refers to all combinations of each $\mu_j = 0,1$ for $j = 1,2,\cdots,N$, $B_1(\mu)$ and $B_2(\mu)$ denote that when we select all the possible combinations $\mu_j = 0,1$ for $j = 1,2,\cdots,N$ the following conditions hold, respectively

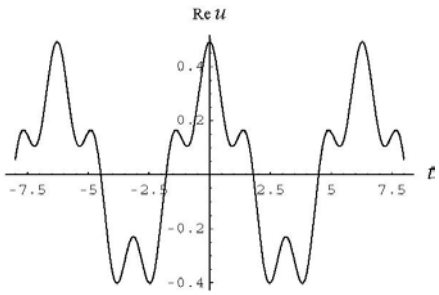
$$\sum_{j=1}^{N} \mu_j = \sum_{j=1}^{N} \mu_{N+j}, \ \sum_{j=1}^{N} \mu_j = \sum_{j=1}^{N} \mu_{N+j} + 1. \tag{63}$$

We would like to note that solutions (15), (24), (31) and (62) with arbitrary constants are more general than the ones constructed by the existing methods, for example, Hirota's bilinear method [31] as stated. To the best of our knowledge, these obtained solutions have not been reported in literature. In addition, we can easy to see that the obtained multiwave solutions with arbitrary constants have another advantage over the ones in [31], that is the high-wave solution can give all the low-wave solutions. Or, more specifically, the three-wave solution (26) given $a_3 = 0$ turns into the double-wave solution (21). Giving $a_2 = 0$ and $a_3 = 0$ to the three-wave solution (26) or giving $a_2 = 0$ to the double-wave solution (21), we can reach the single-wave solution (15). Similarly, the *N*-wave solution (47) can give any low-wave solution as long as these constants $a_j = 1$ ($j = 1,2,\cdots,N$) are properly selected.

Figs. 1-12 display the modulus, real part and imaginary part of the single-wave solution (15), double-wave solution (24) and three-wave solution (31) respectively, which propagate along x-axis. In Figs. 1-3, the parameters are selected as $a_1 = 1 + 2\mathrm{i}$, $k_1 = 1 - 0.8\mathrm{i}$, $\omega_1 = 0$. In Figs. 4-6, the parameters are selected as $a_1 = 1$, $a_2 = 1$, $k_1 = 1$, $k_2 = 2$, $\omega_1 = 0$, $\omega_2 = 0$. In Figs. 7-9, the parameters are selected as the same as Figs. 4-6 except for the different $k_1 = 1 - 0.5\mathrm{i}$, $k_2 = 2 + 0.05\mathrm{i}$. In Figs. 10-12, the parameters are selected as $a_1 = 1$, $a_3 = 1$, $a_2 = 1$, $k_1 = 1 - 0.2\mathrm{i}$, $k_2 = 2 + \mathrm{i}$, $k_3 = 3$, $\omega_1 = 0$, $\omega_2 = 0$, $\omega_3 = 0$. It is easy to see from Figs. 3-12 that the "breather" phenomena have occurred at different locations in the process of propagation of the solutions (15), (24) and (31). Besides, Figs. 8-12 show that some envelopes have been shaped in the interaction of the double-wave solution (24) and three-wave solution (31). It is due to the arbitrariness of the included parameters, these obtained multiwave solutions possess enrich structures.

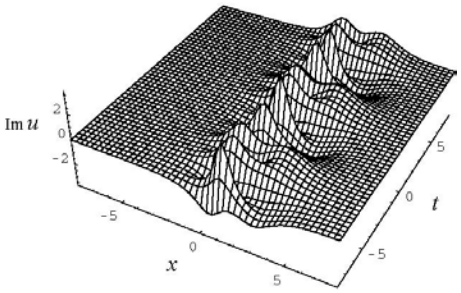


(5a) $x \in [-8,8]$, $t \in [-8,8]$

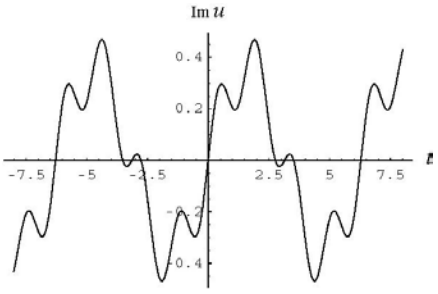


(5b) $x = -1$, $t \in [-8,8]$

Fig. 5. Real part of double-wave solution (24).



(6a) $x \in [-8,8]$, $t \in [-8,8]$



(6b) $x = -1$, $t \in [-8,8]$

Fig. 6. Imaginary part of double-wave solution (24).

Remark 1. All solutions (15), (24) and (31) obtained above have been checked with Mathematica by putting them back into Eq. (1).

IV. Conclusions

In summary, the single-wave solution (15), double-wave

solution (24), three-wave solution (31) and the uniform formula of *N*-wave solution (62) of the nonlinear Schrödinger equation (1) have been obtained due to to the generalization of the exp-function method presented in this paper. Even if these obtained solutions can be constructed by some a future improved version of Hirota's bilinear method [31], the proposed method with the help of Mathematica for generating single-, double-wave and three-wave solutions (15), (24) and (31) is more simple and straightforward.

Generally speaking, when we use Hirota's bilinear method [31], the considered equation must be reduced to the so-called Hirota bilinear form of one or more new dependent variables by means of a suitable transformation and the defined bilinear operator. For Eq. (1), we may take a rational transformation:

$$u = \frac{G}{F},\ F = F(x,t),\ G = G(x,t), \tag{64}$$

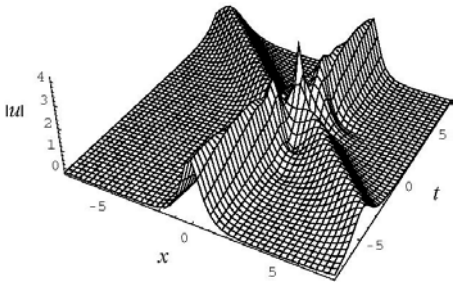
then it is reduced to the so-called Hirota bilinear forms:

$$(\mathrm{i}D_t + D_x^2)G \cdot F = 0,\ D_x^2 F \cdot F = GG^*, \tag{65}$$

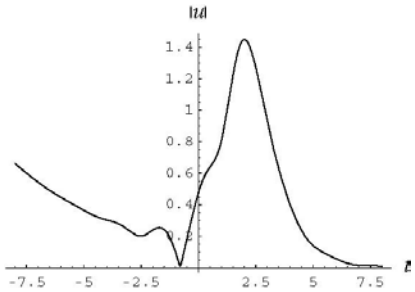
where D_x and D_t are the bilinear operators [31].

Secondly, expanding each of new dependent variables in infinite series of a formal expansion parameter, we spit the Hirota bilinear form into a system of linear differential equations, from which we truncate the infinite series by selecting some appropriate exponential function solutions of the obtained differential equations. Finally, we use the selected exponential function solutions to determine the new variables and hence the multi-wave solutions of the given equation.

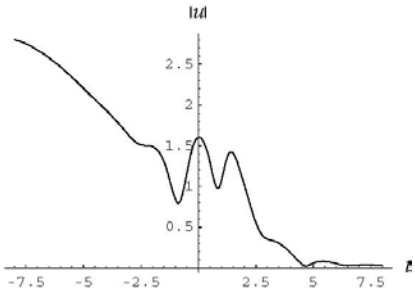
Compared with Hirota's bilinear method, the generalized exp-function method presented in this paper does not take above steps in constructing multiwave solutions. We note that there is still not a general rule for us to take in the selection of a suitable transformation such as (64). Besides, the obtained multi-wave solutions constructed by the generalized exp-function method contain some free parameters so that the high-wave solution degenerate into all the low-wave solutions, which are more general than the ones through Hirota's bilinear method. In this sense, we may conclude that the generalized exp-function method has the advantage of simplicity and effectiveness and may provide us with a straightforward and applicable mathematical tool for generating multiwave solutions or testing its existence and can be extended to some other complex nonlinear PDEs in mathematical physics.



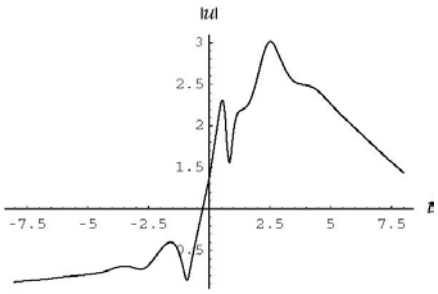
(7a) $x \in [-8,8]$, $t \in [-8,8]$



(7b) $x = -1$, $t \in [-8,8]$

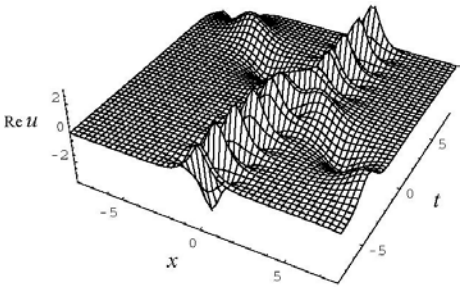


(7c) $x = 0$, $t \in [-8,8]$

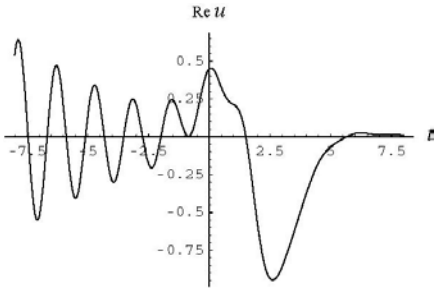


(7d) $x = 2$, $t \in [-8,8]$

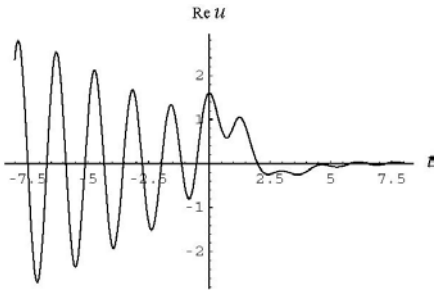
Fig. 7. Imaginary part of double-wave solution.



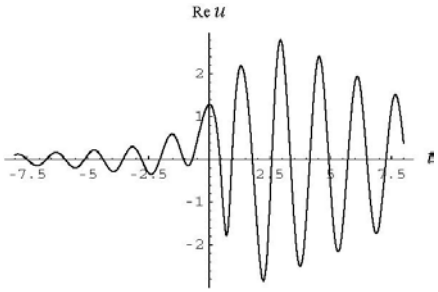
(8a) $x \in [-8,8]$, $t \in [-8,8]$



(8b) $x = -1$, $t \in [-8,8]$

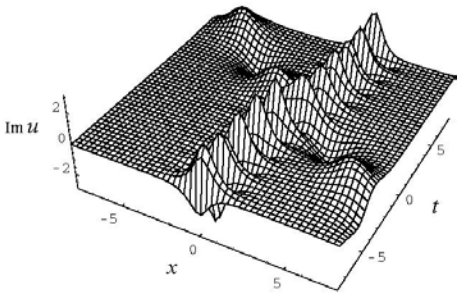


(8c) $x = 0$, $t \in [-8,8]$

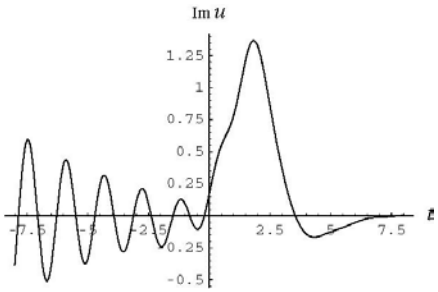


(8d) $x = 2$, $t \in [-8,8]$

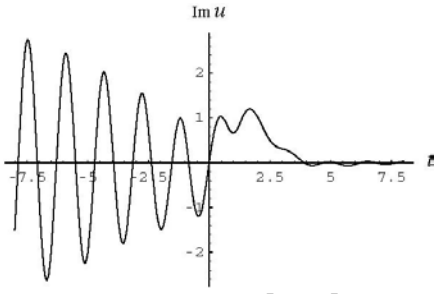
Fig. 8. Imaginary part of double-wave solution (24).



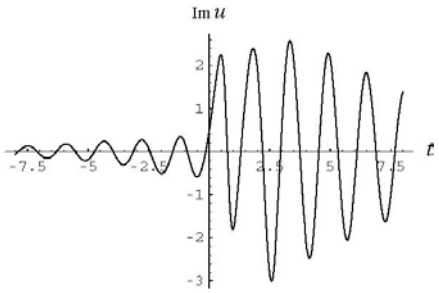
(9a) $x \in [-8,8]$, $t \in [-8,8]$



(9b) $x = -1$, $t \in [-8,8]$

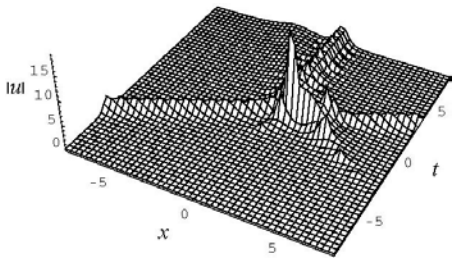


(9c) $x = 0$, $t \in [-8,8]$

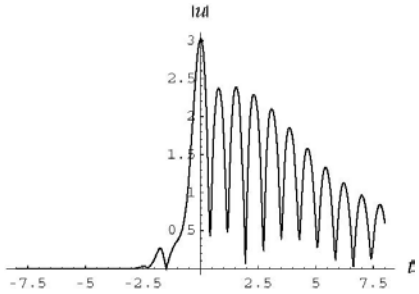


(9d) $x = 2$, $t \in [-8,8]$

Fig. 9. Imaginary part of double-wave solution.

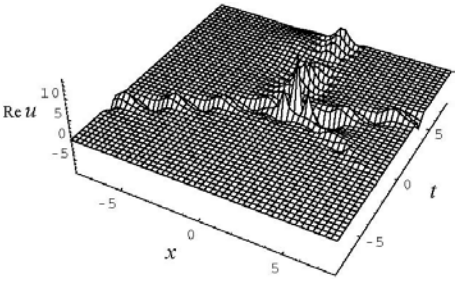


(10a) $x \in [-8,8]$, $t \in [-8,8]$

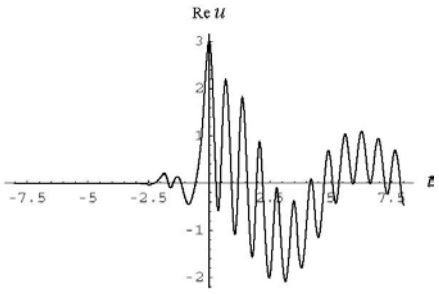


(10b) $x = 0$, $t \in [-8,8]$

Fig. 10. Imaginary part of three-wave solution.



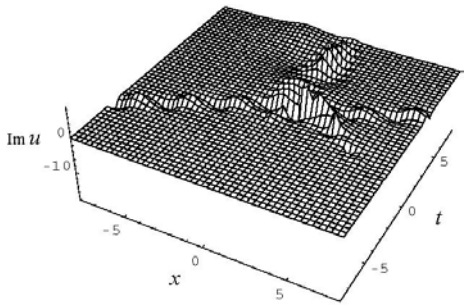
(11a) $x \in [-8,8]$, $t \in [-8,8]$



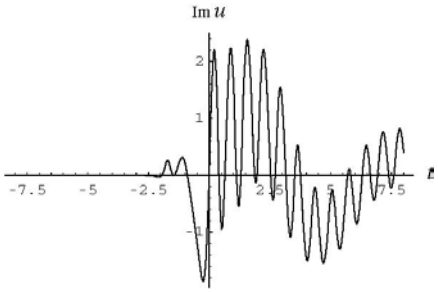
(11b) $x = 0$, $t \in [-8,8]$

Fig. 11. Imaginary part of three-wave solution.

ISBN: 978-1-61804-292-7



(12a) $x \in [-8,8]$, $t \in [-8,8]$



(12b) $x = 0$, $t \in [-8,8]$

Fig. 12. Imaginary part of three-wave solution.

REFERENCES

[1] C. S. Garder, J. M. Greene, M. D. Kruskal, and R. M. Miura, "Method for solving the Korteweg-de Vries equation," *Phys. Rev. Lett.* vol. 19, no. 12, pp. 1095–1097, Nov. 1967.

[2] R. Hirota, "Exact solution of the Korteweg-de Vries equation for multiple collisions of solitons," *Phys. Rev. Lett.* vol. 27, no. 18, pp. 1192–1194, Nov. 1971.

[3] M. R. Miurs, *Bäcklund Transformation.* Berlin, Springer, 1978.

[4] J. Weiss, M. Tabor, and G. Carnevale, "The Painlevé property for partial differential equations," *J. Math. Phys.* vol. 24, no. 3, pp. 522–526, Mar. 1983.

[5] M. L. Wang, "Solitary wave solutions for a variant Boussinesq equations," *Phys. Lett. A* vol. 199, no. 3-4, pp. 169–172, Mar. 1995.

[6] E.G. Fan, "Travelling wave solutions in terms of special functions for nonlinear coupled evolution systems," *Phys. Lett. A* vol. 300, no. 2-3, pp. 243–249, Jul. 2002.

[7] E. G. Fan and H. H. Dai, "A direct approach with computerized symbolic computation for finding a series of traveling waves to nonlinear equations," *Comput. Phys. Commun.* vol. 153, no.1, pp. 17–30, Jun. 2003.

[8] E. Yomba, "The modified extended Fan sub-equation method and its application to the (2+1)-dimensional Broer–Kaup–Kupershmidt equation," *Chaos Soliton. Fract.* vol. 27, no. 1, pp. 187–196, Jan. 2007.

[9] S. Zhang and T. C. Xia, "A generalized auxiliary equation method and its application to (2+1)-dimensional asymmetric Nizhnik–Novikov –Vesselov equations," *J. Phys. A: Math. Theor.* vol. 40, no. 2, pp. 227–248, Jan. 2007.

[10] S. Zhang and H. Q. Zhang, "Variable-coefficient discrete tanh method and its application to (2+1)-dimensional Toda equation." *Phys. Lett. A* vol. 373, no. 33, pp. 2905–2910, Aug. 2009.

[11] E. G. Fan, K. W. Chow, and J. H. Li, "On doubly periodic standing wave solutions of the coupled higgs field equation," *Stud. Appl. Math.* vol. 128, no. 1, pp. 86–105, Jan. 2012.

[12] W. X. Ma and J. H. Lee, "A transformed rational function method and exact solutions to 3+1 dimensional Jimbo–Miwa equation," *Chaos Soliton. Fract.* vol. 42, no. 3, pp. 1356–1363, Mar. 2009.

[13] Z. Y. Yan, "Localized analytical solutions and parameters analysis in the nonlinear dispersive Gross–Pitaevskii mean-field GP (m,n) model with space-modulated nonlinearity and potential," *Stud. Appl. Math.* vol. 132, no. 3, pp. 266–284, Apr. 2014.

[14] C. Q. Dai, X. G. Wang, and G. Q. Zhou, "Stable light-bullet solutions in the harmonic and parity-time-symmetric potentials," *Phys. Rev. A* vol. 89, no. 1, 013834(7pp.), Jan. 2014.

[15] J. H. He and X. H. Wu, "Exp-function method for nonlinear wave equations," *Chaos Soliton. Fract.* vol. 30, no. 3, pp. 700–708, Nov. 2006.

[16] S. Zhang, "Application of Exp-function method to a KdV equation with variable coefficients," *Phys. Lett. A* vol. 365, no. 5-6, pp. 448–453, Jun. 2007.

[17] S. D. Zhu, "Exp-function method for the hybrid-lattice system," *J. Nonlinear Sci. Numer. Simul.* vol. 8, no. 3, pp. 461–464, Sep. 2007.

[18] C. Q. Dai and J. L. Chen, "New analytic solutions of stochastic coupled KdV equations," *Chaos Soliton. Fracat.* vol. 42, no. 4, pp. 2200-2207, Nov. 2009.

[19] V. Marinakis, "The Exp-function method find *n*-soliton solutions," *Z. Naturforsch. A* vol. 63, no. 10-11, pp. 653–656, Oct. 2008.

[20] S. Zhang and H. Q. Zhang, "Exp-function method for *N*-soliton solutions of nonlinear evolution equations in mathematical physics," *Phys. Lett. A* vol. 373, no. 30, pp. 2501–2505, Jul. 2009.

[21] A. Ebaid, "Exact solitary wave solutions for some nonlinear evolution equations via Exp-function method," *Phys. Lett. A* vol. 365, no. 3, pp. 213–219, May 2012.

[22] S. Zhang, "Exact solutions of a KdV equation with variable coefficients via Exp-function method," *Nonlinear Dyn.* vol. 52, no. 1-2, pp. 11–17, Apr. 2008.

[23] Boz and A. Bekir, "Application of exp-function method for (3+1)-dimensional nonlinear evolution equations," *Comput. Math. Appl.* vol. 56, no. 5, pp. 1451–1456, Sep. 2008.

[24] X. H. Wu and J. H. He, "Solitary solutions, periodic solutions and compacton-like solutions using the Exp-function method,"*Comput. Math. Appl.* vol. 54, no. 7-8, pp. 966–986, Oct. 2007.

[25] J. H. He and M. A. Abdou, "New periodic solutions for nonlinear evolution equations using Exp-function method," *Chaos Soliton. Fract.* vol. 34, no. 5, pp. 1421–1429, Dec. 2007.

[26] J. H. He and L. N. Zhang, "Generalized solitary solution and compacton-like solution of the Jaulent-Miodek equations using the Exp-function method," *Phys. Lett. A* vol. 372, no. 7, 1044–1047, Feb. 2007.

[27] S. Zhang and Y. Y. Zhou, "Multiwave solutions for the Toda lattice equation by generalizing Exp-Function method," *IAENG Int. J. Appl. Math.* vol. 44, no. 4, 177–182, Feb. 2014.

[28] L. Zhao, D. J. Huang, and S. G. Zhou, "A new algorithm for automatic computation of solitary wave solutions to nonlinear partial differential equations based on the Exp-function method," *Appl. Math. Comput.* vol. 219, no. 4, pp. 1890–1896, Nov. 2012.

[29] S. Zhang, J. Wang, A. X. Peng, and B. Cai, "A generalized Exp-function method for multiwave solutions of sine-Gordon equation," *Pramana J. Phys.* vol. 81, no. 5, pp. 763–773, Nov. 2013.

[30] I. Aslan, "On the application of the Exp-function method to the KP equation for *N*-soliton solutions," *Appl. Math. Comput.* vol. 219, no. 6, pp. 2825–2828, Nov. 2012.

[31] R. Hirota, *The Direct Method in Soliton Theory.* New York, Cambridge University Press, 2004.

Optimization of Truss Structures Using Genetic Algorithms with Domain Trimming (GADT)

Samer Barakat and Omar Nassif

***Abstract*—** This paper presents evolutionary-based least-weight optimization procedure for designing truss structures. A modified version of Genetic Algorithm with Domain Trimming (GADT) is developed and presented herein. The DADT is used for solving the nonlinear constrained optimization problems. In this optimum design formulation, the objective function is the material weight of the truss; the design variables are the cross-sections of the truss members; the constraints are the stresses in members and the displacements of the joints. The constraints were handled using non-stationary dynamically modified penalty functions. One classical truss optimization example is presented herein to demonstrate the efficiency of the GADT algorithm. The test problem includes a 10-bar planar truss subjected to two load conditions. The result shows that the GADT method is very efficient in finding the best discovered optimal solutions, which are better of the results of other structural optimization methods..

***Keywords*—**Truss Structural Optimization, Genetic Algorithm, Domain Trimming, Constraint Handling.

I. Introduction

Obtiaing optimal designs that satisfy multiple conflicting criteria, such as minimum cost and maximum performance, is one of the most influential factors in modern structural design. Most structural designs are considered constrained optimization problems that can be solved to identify the design values of structural performance. The optimum solutions might be linearly and/or nonlinearly constrained in the design space. In the presence of multiple optima and non-smooth constraints in the design variable space, it is difficult to obtain a set of optimum values using local optimization techniques. On the other hand, this difficulty has geared the research towards relatively new and innovative evolutionary based optimization techniques [1] such as the Genetic Algorithm (GA) [2], Ant Colony Optimization (ACO) [3], Particle Swarm Optimizer (PSO) [4], Shuffled Complex Evolution (SCE) [5]-[7], Harmony Search [8], and Hybrid Methods [9]-[11]. These approaches are investigated and used in recent years for optimizing structural designs and are proven superior to local search techniques. Many structural optimization problems involve problem-specific constraints applicable to the solutions limiting the feasible search space Compared to other constraint handling techniques the use of penalty functions is relatively simple and easy to implement. This study presents the development and implementation of the GADT to achieve superior optimization. The capabilities of the developed optimization tool are demonstrated on two classical truss optimization problems being challenging with unknown global and multiple local minima.

S. B. Author is with the University of Sharjah, Sharjah, UAE (corresponding author phone: +971-6-5050958; fax: +971-6-5585173; e-mail: sbarakat@sharjah.ac.ae).

T. C. Author is with Department of Civil Engineering, University of Tennessee, USA (e-mail: onassif@vols.utk).

II. Genetic Algorithm with Domain Trimming (GADT)

A. Development of the GADT Optimization Technique

To begin GA optimization, a population of solution alternatives Np (population size) are randomly generated using a uniform probability distribution; each solution of the GA consists of a combination of variables (x1, x2, x3, …, xn) which has its own fitness value. In cases where the optimization is performed to find the minimum weight for a given problem, a function F*j(X) = total mass + penalty has to be minimized. Solution alternatives that yield low F*j(X) values for the objective function would have better fitness as long as they are not violating the problem constraints. Populations of solutions are represented by chromosomes. The design variables stored in the chromosome can be either discrete (selected from a pool of defined values) or continuous (selected from a continuous range of variables). In this research, the vector of variables contains continuous values. Once F*j(X) for every solution j in the initial population is computed, a fitness value is assigned to each solution j using Eq.(1), Solutions with F*j(X) less than F*ave of the population are considered unfit and are eliminated by assigning them a fitness value of zero:

$$F_j(\boldsymbol{X}) = \begin{cases} \mathrm{F}^*_{\mathrm{ave}} - F^*_j(\boldsymbol{X}) & \text{for } F^*_j(\boldsymbol{X}) < \mathrm{F}^*_{\mathrm{ave}} \\ 0 & \text{for } F^*_j(\boldsymbol{X}) \geq \mathrm{F}^*_{\mathrm{ave}} \end{cases} \quad \mathrm{j} = 1,2,\dots,\mathrm{N_p} \qquad (1)$$

where ($\boldsymbol{X}$) is the vector of the design variables.

The three basic operations of a GA, reproduction, crossover, and mutation, are used to improve the fitness of each population from one generation (iteration) to the next. The reproduction operation selects the better fit designs, copies them, and places them into a mating pool allowing each to mate and reproduce. The roulette wheel selection method is used in this study for its simplicity and popularity (Fig. 3). This method assigns for each fitness function value a portion of unity that will be used as the reproduction probability, $\boldsymbol{P_r}$. If for example, the best fit solution in a certain population has a

ISBN: 978-1-61804-292-7

P_r value of 10% and the population size N_p =100 then the mating pool will have approximately 10 copies of this solution.

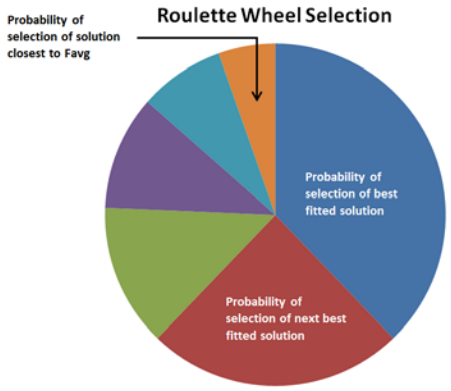


Fig. 3: Pie chart of probability of selection of fit parents

After the reproduction operation is performed, the crossover operation mates the selected designs to create more fit offspring solutions. The uniform crossover operation is used to combine genetic information between two parent solutions. Uniform crossover selects two parent solutions at a time from the mating pool and swaps variables corresponding to zeros in a binary vector known as a mask. The mask is the same length as all variable vectors and consists of a preselected percentage of randomly arranged zeros (%c). This percentage has an impact on the speed of convergence to an optimum solution. Each mask within a population is different, so the number of unique, randomly generated masks is equal to half of the number of solutions (parents) in the population multiplied by the total number of generations (populations).

Since GA mimics the natural selection of the fittest treats through multiple generations, it is inherently vulnerable to *Genetic Drift* (continuous survival of "unfit" but "lucky" individuals, and their genes, from one generation to the next). The mutation operation is used to minimize the effect of genetic drift and add diversity to the search space by randomly changing a variable in a design solution. During mutation the value of any chromosome variable may be changed to a randomly selected variable. Another counter measure of genetic drift used for continuous variables is to use blending techniques [12], a general blending formula would be of the form:

$$x_{new} = \beta x_{mn} + (1-\beta)x_{dn} \quad (2)$$

Where, x_{new} = the nth variable of the offspring chromosome of the crossover, β = random number on the interval [0,1], x_{mn} = the nth in the mother chromosome, and x_{dn} = the nth in the father chromosome.

One of the main advantages for the GA optimization is its relative insensitivity to local minima (not very susceptible to being trapper in local minima). However, this may also be considered as one of its limitations: being a *"low-resolution"* technique especially if the population size is relatively limited. i.e. convergence may sometime occur to points in the neighborhood of, but not exactly at, global minima. The reason for this is that initially, the solution variables are randomly selected from a pre-specified range (domain); those variables do not change (except for mutation, which occurs at small probability) but rather change place from one solution to another. In addition, a poor choice of the solution domain (e.g.: [1, 1000] while the optimal value is at a value of 2) further aggravates the issue. Blending techniques, while acting to reduce this disadvantage, introduce a new random parameter (β), it may also have a negative effect as it may negate the strong traits of the parents.

In this study, a new technique is developed to alleviate this inherent limitation of GA. The technique involves trimming the domain then re-initiating the GA so that the probability of selecting the optimal solution is improved. For example, trimming a domain from [1, 1000] to [1, 10] then re-initiating GA would increase the initial probability of selection of the optimal solution by 200 times, when everything being constant. Trimming is done as a percentage of original domain size, and continuous trimming then GA re-initiation goes on until the optimum solution is found. One can think of trimming in terms of evolution theory as when a catastrophe occurs in nature eliminating the majority of the population leaving only the elite survivals to restart the evolution process. Fig.4 simulates this technique on a single variable chromosome (taking trimming as 90% of domain size):

1- After the GA converged to a certain objective function value, the elite 10 chromosomes (in this case, variables) are taken regardless of which generation they're in, since the range between the minimum and maximum variable is more than 90%, no trimming occurs.

2- As the GA converged for the second time, the range of values are now less than 90% and trimming will happen, but rather than being centered, the trimming range is biased toward the mean by a ratio: (max-mean)/(max-min), the trimmed domain is now less than 90% of the original as some of the trimming range lies outside the original domain.

3,4,5 - Because the domain has been trimmed, the chances of getting to the global optimum are higher, resulting in the elite values getting closer to each other, trimming will stop when range of values exceeds 90% of the domain or if the GA found the optimum solution to a satisfactory precision.

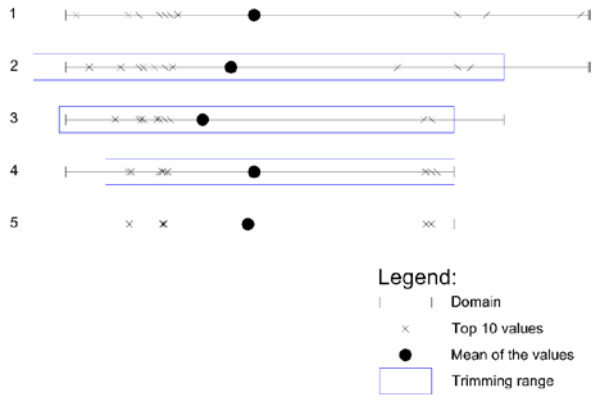


Fig. 4: Illustration of the domain trimming method on a single variable (trimming is 90%)

A question arises that how can it be sure that the optimal solution is inside the trimmed domain; the answer is that it doesn't guarantee that, however, the worst case scenario is that it will get results comparable to conventional GA since

trimming doesn't happen until GA converges. Experimenting on sample problems showed that the technique gives better results than conventional GA. Fig. 5a-d illustrates the logical steps for the GADT technique via flowcharts.

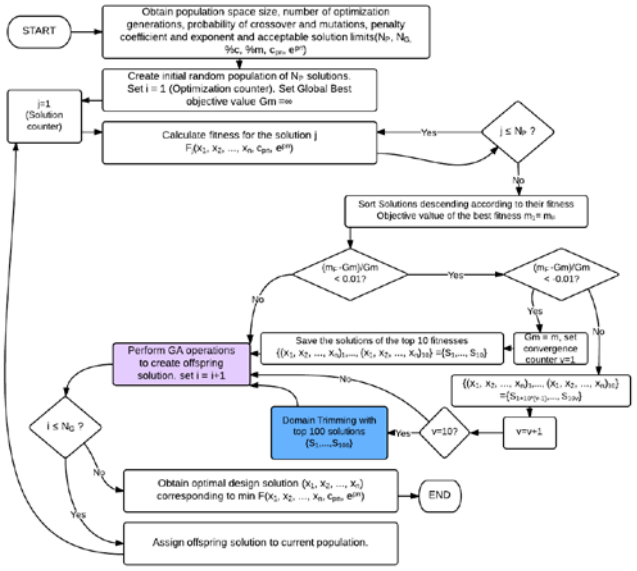


Fig. 5a: Flow chart of general GADT optimization

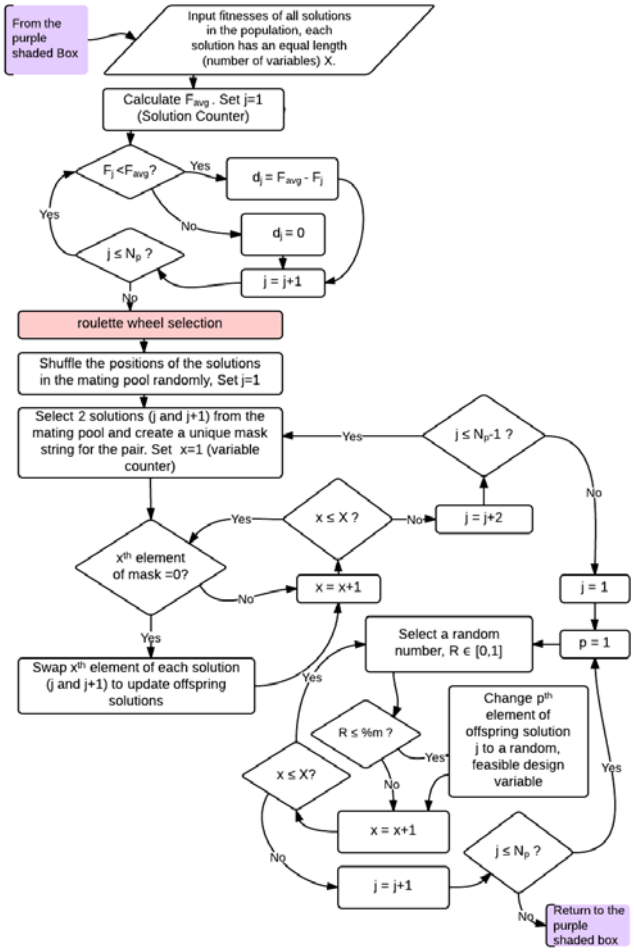


Fig. 5b: Flow chart of Reproduction, Crossover and Mutation

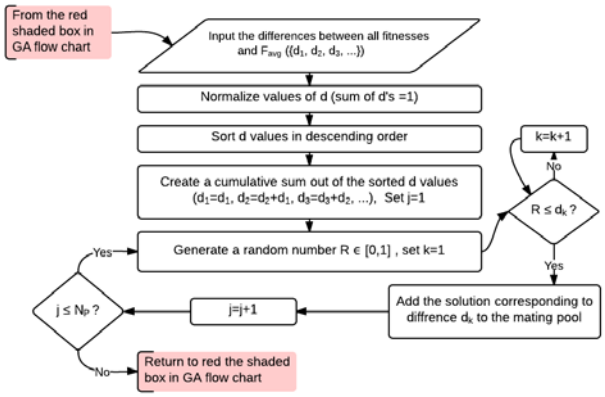


Fig. 5c: Flow chart of Roulette wheel selection

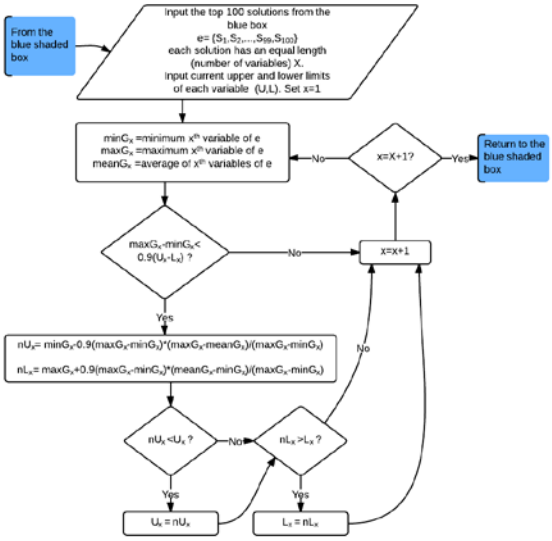


Fig. 5d: Flowchart of Domain Trimming

To demonstrate the GADT algorithm performance, one of the standard test functions in optimization problems is considered: the Six-Hump Camelback function problem (43):

$$f(x,y) = \left(4 - 2.1x^2 + x^{\frac{4}{3}}\right)x^2 + xy + (-4 + 4y^2)y^2 \quad (3)$$

With boundaries: x ∈ [-1.5, 1.5] and y ∈ [-2, 2], the function takes the form shown in Fig. 6; it has six minima, two of them are global minima with a value of: -1.0316 located at the two points: (-0.0898, 0.7126) and (0.0898, -0.7126).

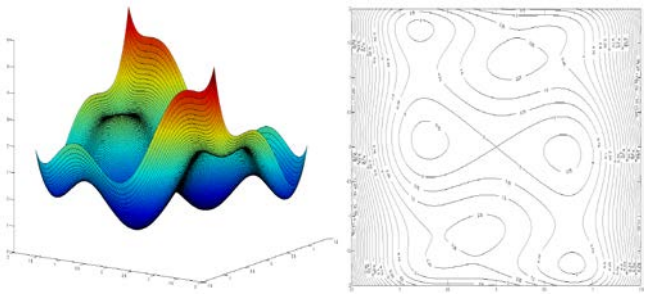


Fig. 6: Six-hump Camelback function. a) 3D plot b) Contour plot

ISBN: 978-1-61804-292-7

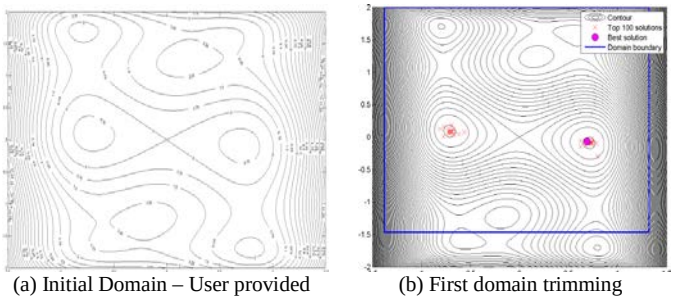
Initially, the GA starts off with the domain provided by the user and iterates until there is a convergence, i.e. the fitness values are close to each other and the improvement in best fitness in subsequent iterations is less than a specified value, say, 1%. At this stage, the GA has identified a number of local minima and further iterations will have little effect on finding the global minimum.

Convergence to a series of *"not necessarily most fit"* solutions means that the GA is approaching the vicinity of the global optimum but cannot discover the fittest solution. This may be due to elimination of some of the fittest solutions as they are combined with lesser-fit solutions, or inclusion of residual unfit solutions within the available solution domain. At this point, refinement of the available solution domain improves the chances of discover the fittest solution. GADT technique removes the most unfit solutions from the domain and re-initiates GA within the trimmed domain. Occasionally, some fittest solutions are dismissed as unfit as result of being combined with lesser-fit solutions. To minimize the probability of this potential pitfall, two measures are incorporated in the GATD: (a) the initial domain considers a large enough population to allow higher probability for discovering the most combinatory possibilities, and (b) the domain trimming limited to a high percentage (e.g. 90%) of the domain from the previous step. This means that while some values are clearly identified as unfit, they are still retained in the domain for the subsequent GA re-initiation. The motivation for this retention it exhaust all possibilities of producing fit solutions before completely eliminating part of the domain. The impact on speed of discovery of optimum solutions is obvious. However, it is justified by the significantly improved successful discovery rate. The method can also be helpful to identify what range of variables to look for in subsequent searches.

Fig.7 shows the fittest 100 overall solutions throughout the GADT; the box drawn on the contour map represents the updated (trimmed) domain, eliminating 10% of the original domain at each step. The continuation of domain trimming results in identifying new local minima. As more local minima are identified, they begin to compete with each other. Eventually, the trimming technique will exclude the values of low fitted local minima and will deem them as unfit results. This helps the GA to focus more on only the top-fitted solutions ultimately enhance the discovery rate (precision) of global optima solutions. Fig.7d shows the final domain after 20 domain trimming iterations; it is noted that the domain cannot get trimmed further as the two global minima reside on the trimmed boundaries of the last step. It is worth mentioning that the final domain is only 3.5% of the initial full domain in more precise (virtually exact) optimum solution: f(x, y) = -1.0316284533464 (0.000000014% error).

B. GADT Technique Robustness

To ensure GADT robustness, domain trimming is subject to further criteria preventing any potential adverse consequences. Domain trimming will not resume in any of the following scenarios: (a) if the range of values that yield fit results in the current domain is more than 90% of the previous domain; in other words, domain trimming will not commence until the feasible values get condensed into less than 90% of the domain. And (b) if after trimming, the GA yields solutions worse than previous GA initiation. Although the latter was not encountered during any of the algorithm testing sessions for any of the presented problems here, the criterion is set in place for potential future problem-specific complications. The effect of initial domain size is discussed in a subsequent section.



(a) Initial Domain – User provided (b) First domain trimming

(c) after 10 domain trimming (d) Final solution domain

Fig.7: Illustration of the GADT algorithm applied to the Six-hump Camelback function at different iterations of domain trimming.

C. Constraints Handling

The use of penalty functions is very popular in handling constraints enabling the solution of constrained problems as unconstrained. The solutions that violate any constraints are penalized in order to characterize non-feasible solutions by high objective function values. Non-stationary (dynamic) penalty functions typically exhibit superior performance to stationary (static) penalty functions. In its generality, a non-stationary penalty function is defined as:

$$f(X) = F(X) + p(X, c_{pn}, e_{pn}) \tag{4}$$

Where *F(x)* is the original objective function of the constrained optimization problem; $p(\bullet)$ is a dynamically modified penalty value, defined as:

$$p(X, c_{pn}, e_{pn}) = \begin{cases} (c_{pn}\, r_i)^{e_{pn}}, & r \geq 1 \\ 0, & r < 1 \end{cases} \tag{5}$$

Where, $\boldsymbol{r_i}$ is the individual member's performance criteria (in a structural design problem, this is often taken as the demand-to-capacity ratio, or utilization ratio to the satisfaction of the relevant design code); while $\boldsymbol{c_{pn}}$ and $\boldsymbol{e_{pn}}$ are the penalty coefficient and exponent, respectively. As their name implies, they provide means to penalize the optimization objective if the $\boldsymbol{r_i}$ exceeds unity. Both $\boldsymbol{c_{pn}}$ and $\boldsymbol{e_{pn}}$, with different severity, will penalize unfit solutions minimizing their probability of re-appearing subsequent generation and feasible solution domain.

D. Truss Structural Optimization

The mathematical form of the optimization problem for truss structure can be expressed as follows:

$$\text{Find} \qquad A^T = \{A_1,\ A_2,\ \ldots\ldots\ A_n\} \tag{6}$$

$$\text{To Minimize} \qquad F = W(A) = \rho \sum_{i=1}^{n} L_i A_i \tag{7}$$

$$\text{Subject to} \qquad g_j^L \le g_j(A) \le g_j^U \qquad j = 1,2,\ldots m \tag{8}$$

$$\text{and} \qquad A_i^{\min} \le A_i \le A_i^{\max} \qquad i = 1,2,\ldots n \tag{9}$$

Where A_i = the design variable i (member i cross-sectional area), n = the number of the design variables, $W(A)$ = the objective function (the structural weight), ρ = the material density, L_i = the member length, m = the number of inequality constraints (g), $A_i^{\min}$ and $A_i^{\max}$ are the lower and the upper bound of the i^{th} variable respectively. The lower and upper bounds posed by Eq.(8) on the constraints include truss member stresses and joint displacements.

E. Example: Cantilever 10-Bar Planar Truss Structure

The GADT is tested against a classical global optimization problem: the *Cantilever 10-Bar Planar Truss Structure* optimization problem. This 10-dimensional problem has been investigated by many researchers, and has been well-established as an optimization benchmark problem known for being challenging with unknown global and multiple local minima. A schematic of the cantilever 10-bar planar truss structure can be found in Fig.8.

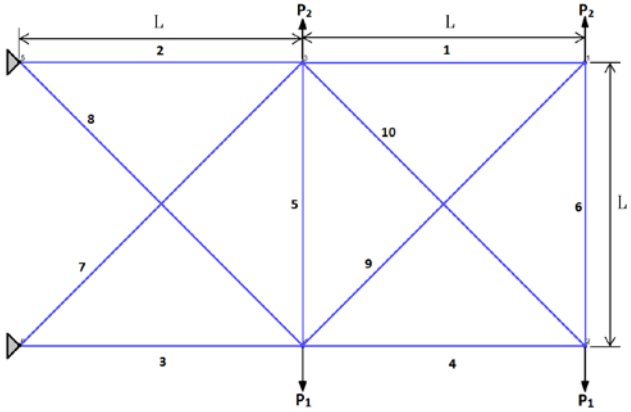


Fig.8: 10-Bar planar cantilever truss model

The assumed material density is 0.1 lb/in^3 (2767.990 kg/m^3), the length L is 360 in (914.4 cm) and the modulus of elasticity is 10,000 ksi (68,950 MPa). The stress and deflection limitations on the members are ±25.0 ksi (172.375 MPa) and ± 2.00 in (5.08 cm), respectively. Cross-sectional areas are allowed to vary between 0.1 in^2 and 35 in^2 (0.6452 cm^2 and 225.806 cm^2). No member grouping is utilized, resulting in each member having a potentially unique cross-sectional area (A_1 to A_{10}). This truss optimization problem has 10 design variables and 32 (10 tension stresses, 10 compression stresses, and 12 displacements) constraints. Two loading cases are studied: Case 1, when P_1 = 100 kips (444.8 kN) and P_2 = 0 kips and Case 2, when P_1 = 150 kips (667.2 kN) and P_2 = 50 kips (222.4 kN).

Table 2: Optimization results for the 10-bar planar truss Load Case 1

Member Areas- cm^2	Mathematical programming methods (MPM)								
	This work	[7]	[9]	[10]		[11]	[8]	[13]	
A1	195.486	196.583	196.646	198.116	197.674	195.560	193.990	194.518	197.871
	30.300	30.470	30.480	30.708	30.640	30.310	30.068	30.150	30.670
A2	0.652	0.645	0.645	0.645	0.645	0.645	0.645	0.658	0.645
	0.101	0.100	0.100	0.100	0.100	0.100	0.100	0.102	0.100
A3	147.756	149.627	150.050	154.275	148.778	150.074	149.720	146.518	153.290
	22.902	23.192	23.258	23.913	23.061	23.260	23.207	22.710	23.760
A4	97.465	98.265	97.973	95.006	96.898	98.264	97.858	98.517	94.129
	15.107	15.231	15.186	14.726	15.019	15.230	15.168	15.270	14.590
A5	0.652	0.645	0.645	0.645	0.645	0.645	0.645	0.658	0.645
	0.101	0.100	0.100	0.100	0.100	0.100	0.100	0.102	0.100
A6	3.871	3.368	3.527	0.645	3.813	3.549	3.458	3.510	0.645
	0.586	0.522	0.547	0.100	0.591	0.550	0.536	0.544	0.100
A7	137.840	135.595	48.079	55.140	48.327	48.261	48.142	139.098	55.342
	21.365	21.017	7.452	8.547	7.491	7.480	7.462	21.560	8.578
A8	48.420	48.420	13.195	135.551	136.180	139.428	136.950	48.652	135.935
	7.505	7.453	2.045	21.010	21.108	20.920	21.228	7.541	21.070
A9	0.652	0.645	138.985	134.262	139.131	139.428	139.550	0.645	135.226
	0.101	0.100	21.543	20.811	21.565	21.610	21.630	0.100	20.960
A10	137.795	139.259	0.645	0.645	0.645	0.645	0.645	138.388	0.645
	21.358	21.585	0.100	0.100	0.100	0.100	0.100	21.450	0.100
Weight kN	22.46	22.51	22.51	22.58	22.51	22.50	22.51	22.50	22.58
lb	5049.83	5060.85	5060.87	5076.68	5061.40	5055.00	5057.37	5057.88	5076.85
σ_{max} Mpa	172.37	172.37	172.37	140.53	172.11	172.37	172.37	172.37	140.30
ksi	25.0	25.0	25.0	20.4	25.0	25.0	25.0	25.0	20.3
Δ_{max} cm	5.08	5.08	5.079995	5.079999	5.079925	5.08	5.08	5.08	5.0797
in	2.0	2.0	1.9999981	1.99999944	1.9999703	2.0	2.0	2.0	1.999859

Tables 1 and 2 give the best discovered optimum solutions along with the corresponding minimum weight for the two cases 1 and 2, respectively, benchmarked against optimal designs by other published studies. It should be noted that the best discovered solution was found after 10000 iterations for case 1 and 4000 iterations for case 2; initial population size, prior to trimming, is taken as 500 solutions.

The optimal solutions found by the GADT meet all of the problem constraints and the comparisons in Tables 1 and 2 show that the GADT provides superior results. Fig.9 shows convergence histories for loading cases 1 and 2. Notice that convergence plateaus after around 800 iterations; at which point, the trimming has progressed such that it has little further effect on the GA signifying the elimination of most or all unfit values.

Table 3: Optimization results for the 10-bar planar truss Load Case 2

Member Areas- cm^2 in^2	Optimal cross-sectional areas							
	This work	[7]	[11]	[8]	[14]	[13]	[15]	[16]
A1	153.313	151.804	148.720	149.997	156.707	151.933	162.513	166.513
	23.764	23.530	23.052	23.250	24.290	23.550	25.190	25.810
A2	0.652	0.645	0.645	0.658	0.645	0.645	2.342	0.645
	0.101	0.100	0.100	0.102	0.100	0.100	0.363	0.100
A3	161.505	163.099	165.170	166.002	150.647	163.164	164.002	175.680
	25.033	25.280	25.601	25.730	23.350	25.290	25.420	27.230
A4	91.052	92.710	97.671	93.613	88.129	92.645	92.452	107.419
	14.113	14.370	15.139	14.510	13.660	14.360	14.330	16.650
A5	0.684	0.645	0.645	0.645	0.645	0.645	2.690	0.645
	0.106	0.100	0.100	0.100	0.100	0.100	0.417	0.100
A6	12.819	12.709	12.703	12.755	12.703	12.709	20.284	13.058
	1.987	1.970	1.969	1.977	1.969	1.970	3.144	2.024
A7	83.148	82.774	78.748	81.354	80.903	82.645	94.258	91.741
	12.888	12.830	12.206	12.610	12.540	12.810	14.610	14.220
A8	80.174	79.936	81.084	78.774	81.742	79.936	77.936	82.452
	12.427	12.390	12.568	12.210	12.670	12.390	12.080	12.780
A9	0.645	0.645	0.645	0.645	0.645	0.645	3.310	0.645
	0.100	0.100	0.100	0.100	0.100	0.100	0.513	0.100
A10	130.663	131.160	131.160	131.354	141.741	131.225	130.708	142.837
	20.253	20.330	20.330	20.360	21.970	20.340	20.260	22.140
Weight kN	20.78	20.81	20.81	20.78	20.88	20.81	21.78	22.52
lb	4668.78	4676.92	4676.05	4668.81	4691.84	4676.96	4895.60	5059.70
σ_{max} Mpa	172.40	172.40	172.40	172.40	172.40	172.40	161.37	172.40
ksi	25.00	25.00	25.00	25.00	25.00	25.00	23.30	25.0
Δ_{max} cm	5.08	5.08	5.08	5.08	5.08	5.08	5.08	4.62
in	2.00	2.00	2.00	2.00	2.00	2.00	2.00	1.82

ISBN: 978-1-61804-292-7

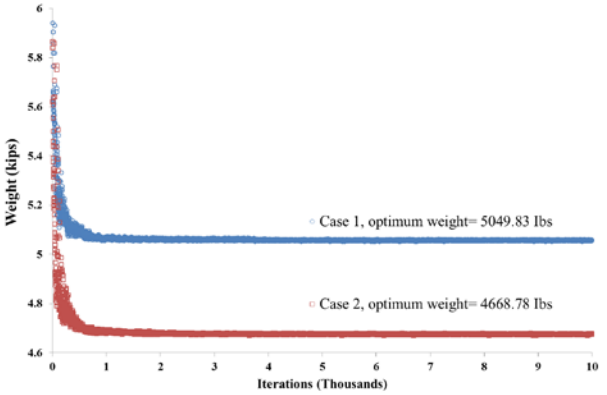


Fig.9: Path to optimum solution for load cases 1 and 2 on the 10-bar truss.

Figs.9 and 10 compare the first 300 iterations of the optimization with and without domain trimming for the loading case 1. Note that for the shown number of iterations the difference between the different discovered solutions is relatively marginal. Figs.11 and 12 compare the convergence histories of GA with and without domain trimming for a population size of 500 going through 2000 iterations. Notice that in some subsequent iterations, the conventional GA identifies a higher minimum weight as the best discovered solution, as compared to a previous iteration. This primarily a result of the mutation process, which eventually is filtered out with enough iterations. This effect is not exhibited in presence of domain trimming since the mutation is unable to select the excluded/trimmed unfit values.

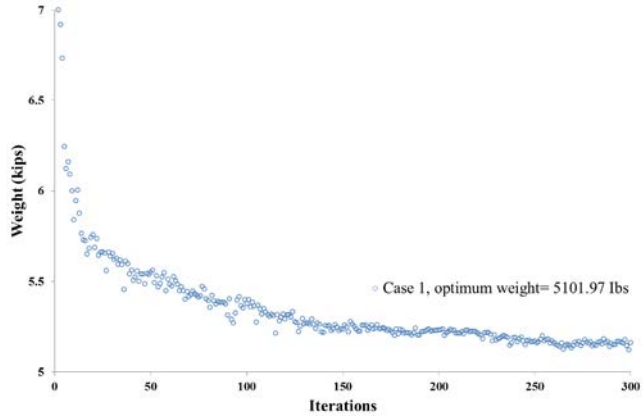


Fig.9: First 300 iterations of a conventional GA optimization on 10-bar truss.

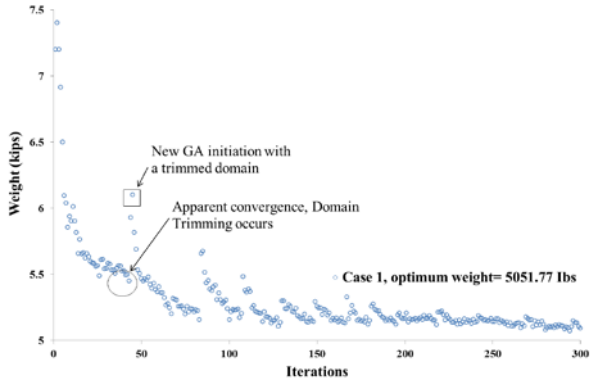


Fig.10: First 300 iterations of GADT optimization on 10-bar truss.

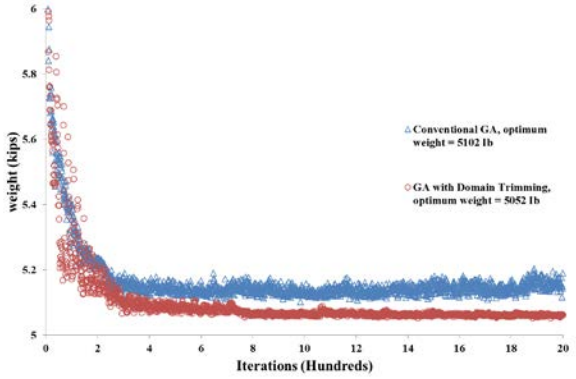


Fig.11: Path to optimum solution for both GA and GADT for load case 1.

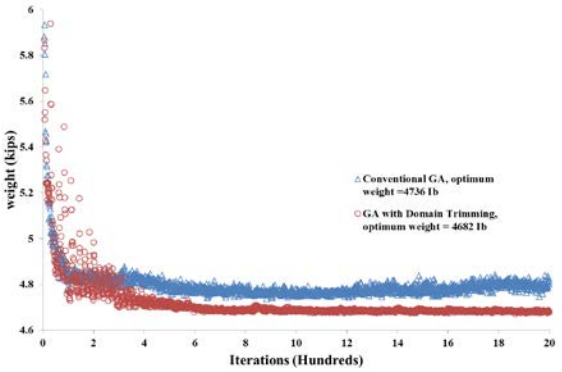


Fig.12: Path to optimum solution for both GA and GADT for load case 2

Since GA optimization is an evolutionary method, it relies on generating populations of solutions in order to find the optimum amongst them. To promote better discovery of optimum results, an adequate population size should be selected. Unfortunately, this parameter is problem-specific. The number of variables in a solution, diversity of values and initial domain selection are some of the factors determining the appropriate population size. The inclination to start with a very large population size in order to give better discovery chances, is sometimes counterproductive when the "low-resolution" limitations of the GA are signified. However, selecting too small of population size will result in convergence failure for the GA for lack of sufficient representative solutions for selection. The GADT technique provides mitigating to this population size effect since its re-initiation gives a new chance for previously undiscovered variables to be selected in the new population. To demonstrate this feature, the 10-bar planar truss problem was repeated with different population sizes. Population sizes chosen for the test were 20, 50, 100, 200, 300, 400 and 500 while all optimizations ran for 2000 iterations. The histogram in Figure 13 shows the effect of population size on finding the optimum solution. It can be seen that increasing the population size improves the precision of finding the optimum solution, although slightly. Population sizes as low as 50 solutions per population yielded acceptable results which are less than 0.5% away from their counterpart with a population size of 500. However, when selecting a population size of 20, the method fails and the discovered solution is far off from the optimum.

ISBN: 978-1-61804-292-7

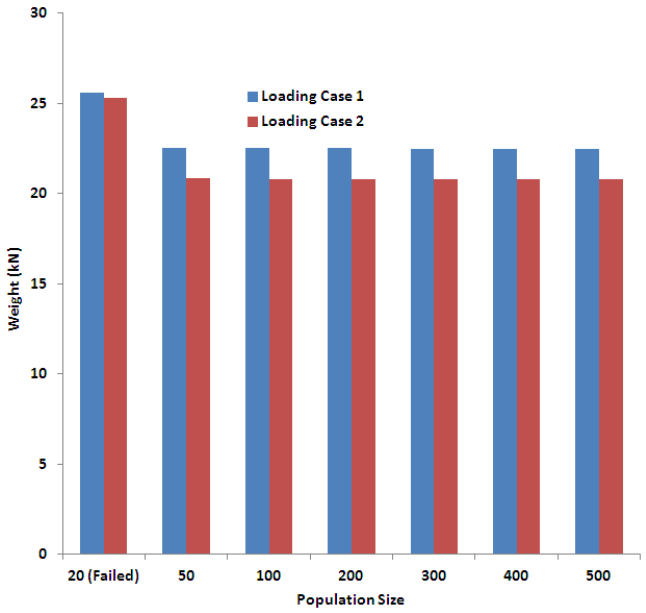


Fig.13: Sensitivity of GADT Algorithm to population size.

III. Summary and Conclusions

A modified version of Genetic Algorithm with Domain Trimming (GADT) is developed and presented in this study. The innovative technique is used to solve a least-weight optimization problem in the design of truss structures. Prior to implementation of the innovative GADT technique on the truss problem, a well-established Six-Hump Camelback function benchmark problem is used for demonstration purposes. Then, the GADT is tested against a well-established challenging benchmark problem; it is tested against a classical global optimization problem: the Cantilever 10-Bar Planar Truss Structure optimization problem. This benchmark problem is a 10-dimensional optimization problem with unknown local and global minima. The GADT performed superiorly in the demonstration as compared to recently published optimum solutions in the literature. In this optimum design formulation, the objective function is the material weight of the supporting truss; the design variables the cross-sectional areas of the truss members; the constraints are the stresses in members and the displacements of the joints. The GADT handles the problem-specified constraints using 'non-stationary penalty functions' method. The results show that the (GADT) method is efficient in finding the best discovered optimal solution. The optimal solutions found by the GADT meet all of the problem constraints and the comparisons with the published literature show that the GADT provides superior results.

References

[1] Lagaros ND, Papadrakakis M, Kokossalakis G. Structural optimization using evolutionary algorithms. Comput Struct. 2002; 80(7-8):571–89.

[2] Balling RJ, Briggs RR, Gillman K. Multiple Optimum Size/Shape/Topology Designs for Skeletal Structures Using a Genetic Algorithm. J Struct Eng. American Society of Civil Engineers; 2006; 132(7):1158–65.

[3] Perez RE, Behdinan K. Particle swarm approach for structural design optimization. Comput Struct. Pergamon Press, Inc.; 2007; 85(19-20):1579–88.

[4] Luh G-C, Lin C-Y. Structural topology optimization using ant colony optimization algorithm. Appl Soft Comput. Elsevier Science Publishers B. V.; 2009; 9(4):1343–53.

[5] Barakat, Samer A., and Salah Altoubat. "Application of evolutionary global optimization techniques in the design of RC water tanks." *Engineering Structures* 31.2 (2009): 332-344.

[6] Barakat, Samer, and Hisham Ibrahim. "Application of shuffled complex evolution global optimization technique in the design of truss structures."*Modeling, Simulation and Applied Optimization (ICMSAO), 2011 4th International Conference on*. IEEE, 2011.

[7] Barakat, Samer A. "Shuffled complex evolution optimizer for truss structure optimization." Computing in Civil and Building Engineering, Proceedings of the International Conference. Vol. 30. 2010.

[8] Lee, Kang Seok, and Zong Woo Geem. "A new structural optimization method based on the harmony search algorithm." Computers & Structures 82.9, 781-798, 2004.

[9] Kaveh, A., and S. Malakouti Rad. "Hybrid genetic algorithm and particle swarm optimization for the force method-based simultaneous analysis and design."Iranian Journal of Science and Technology, Transaction B: Engineering 34.B1, 15-34, 2010.

[10] Csébfalvi, Anikó. "A hybrid meta-heuristic method for continuous engineering optimization." Civil Engineering 53.2, 93-100, 2009.

[11] Kaveh, A., and S. Talatahari. "A hybrid particle swarm and ant colony optimization for design of truss structures." Asian Journal of Civil Engineering 9.4, 329-348, 2008.

[12] N. J. Radcliffe, Forma Analysis and Random Respectful Recombination, Proceedings of the Fourth International Conference on Genetic Algorithms Morgan Kaufmann, San Mateo, CA, 1991, 222-229.

[13] Schmit Jr LA, Miura H, "Approximation concepts for efficient structural synthesis," NASA CR-2552, Washington, DC: NASA, 1976.

[14] SCHMIT JR LA, FARSHI B., 1974. Some approximation concepts for structural synthesis, AIAA J; 12(5):692–9.

[15] VENKAYYA VB., "Design of optimum structures." Computer &Structures 1971, 1(1–2):265–309.

[16] DOBBS MW, NELSON RB., 1976. Application of optimality criteria to automated structural design, AIAA J; 14(10):1436–43.

A Model of the Universe According to the Virial Theorem

Hasan Arslan

Abstract— The model of the universe is defined according to the virial theorem. The previous works are searched. The Energy of the universe and the distance of it are related to each other by an equation derived from the virial theorem. The universe is considered to be under a periodic motion.

***Keywords*—**Virial theorem, Hubble's constant, Accelerating universe, Model of the Universes.

I. Introduction

The article is designed as follows. The derivation of the virial theorem is given in section II both in classical case and in quantum mechanical case according to the previous works [1-7]. In section III, the model of the universe is expressed. The conclusion is given in section IV.

II. The Virial Theorem

In the references [1-7] the virial theorem is derived as follows:

$$G = \sum_i \vec{p}_i . \vec{r}_i \tag{1}$$

here G is considered to be a quantity, product of the momentum and the position of the particle in a stable system. Taking the derivative of Equation (1), we get:

$$\frac{dG}{dt} = \sum_i \left(\frac{d\vec{p}_i}{dt} . \vec{r}_i + \vec{p}_i . \frac{d\vec{r}_i}{dt} \right) \tag{2}$$

The second term on the right hand of the equation (2) can be written as [4, 5]

$$\sum_i \vec{p}_i \frac{d\vec{r}_i}{dt} = \sum_i \left(m\dot{\vec{r}}_i \right) \dot{\vec{r}}_i = m\dot{r}_i^2 = 2T$$
$$\text{and} \sum_i \dot{\vec{p}}_i r_i = \sum_i F_i r_i \tag{3}$$

Here T is the kinetic energy. If the quantity G is bounded in a time interval, one can write:

$$\frac{1}{\tau} \int_0^\tau \frac{dG}{dt} dt = \frac{1}{\tau} \left(G(\tau) - G(0) \right) = 0 \tag{4}$$

From Equation (2), it can be written as

$$\frac{1}{\tau} \int_0^\tau \frac{dG}{dt} dt = 2T + \sum \vec{F}_i . \vec{r}_i \tag{5}$$

for the periodical momentum. From Equations (4) and (5) we obtain;

$$2T = -\sum_i \vec{F}_i . \vec{r}_i \tag{6}$$

Equation (6) is the virial theorem in the classical case.

Now, we are looking for the quantum mechanical virial theorem. The time-dependent Schrödinger Equation is;

$$i\hbar \frac{d\psi}{dt} = H\psi. \tag{7}$$

The derivative of expectation value of an operator *A* with respect to time is;

$$i\hbar \frac{d}{dt} \langle \psi | A | \psi \rangle = \langle \psi | [H, A] \psi \rangle. \tag{8}$$

Let us choose A to be $A = \vec{r} . \vec{p}$ [1-7]. Putting this in Equation (7) and taking *A* to be time-independent, we get;

$$\langle \psi | [H, A] \psi \rangle = 0 \tag{9}$$

Then virial theorem is obtained as

$$[H, A] = \left[\frac{p^2}{2m} + V(r), \vec{r} . \vec{p} \right]$$
$$= i\hbar \vec{r} . \nabla V - \frac{i\hbar}{m} \vec{p}^2 \tag{10}$$
$$= i\hbar \vec{r} . \nabla V - 2i\hbar T = 0$$

where *T* is the kinetic energy and *V* is the potential energy. Then, the virial theorem can be written as

$$2\langle T \rangle = \langle \vec{r} . \vec{\nabla} V \rangle \tag{11}$$

The kinetic energy can be defined by

$$T = \frac{p^2}{2m} \tag{12}$$

where *p* is the momentum operator and *m* is the mass of the particle under consideration. The momentum operator is

$$p = -i\hbar \nabla \tag{13}$$

Using Equation (11) for a potential of the form $V = kr^n$, the kinetic energy is obtained as

$$\langle T \rangle = \frac{n}{2} \langle V(r) \rangle \tag{14}$$

Department of Physics, Bingöl University, Bingöl, Turkey

e-mail: hasanarslan46@yahoo.com

We note here that if $n=0$, we do not have to write Equation (14) because of the singularity, for $n\neq 0$ the equation is valid.
Due to the Equation (12) and the Equation (14), the following relation can be written:

$$V(r)=\frac{p^2}{nm}. \tag{15}$$

The mechanical energy of a closed system is conserved and is given by;

$$E=T+V \tag{16}$$

From Equation (11), we can write

$$T=\frac{1}{2}r\nabla V \tag{17}$$

Then, we can rearrange this last equation by taking

$$\nabla V=\frac{\partial V}{\partial r} \tag{18}$$

as

$$V=2T\ln r/r_0. \tag{19}$$

Then the total mechanical energy of a closed system becomes

$$E=(1+2\ln r/r_0)T \tag{20}$$

or it can be written for the distance travelled as

$$r=r_0 e^{\frac{(E-T)}{2T}}. \tag{21}$$

Here r_0 is the distance between the two spherical shell orbit of the universe, and r is that distance of it at a later time. For whom like to make a search and to learn more about on the virial theorem, I would like to recommend his/her to look the references [8- 26] and the references given therein.

III. The Construction of the Model of the Universe

The Bing Bang Theory suggest that the universe began to be formed by explotion of a very tiny, high densed, very hot point and then began to be cooled quickly and expanded too fast to the outer dimensions and formed the space of the universe where everything,like galaxies-stars- clusters- planets etc., take place in.

One of the Hubble's major discovery was based on comparing his measurements of the Cepheid-based galaxy distance determinations with measurements of the relative velocities of these galaxies. He showed that more distant galaxies were moving away from us more rapidly with speed *v* moving away from us as:

$$v = H_o d \tag{21}$$

where *d* is its distance. The constant of proportionality H_o is now named as the Hubble constant. The common unit of measuring velocity is km/sec, while the most common unit of measuring the distance to the nearby galaxies is Megaparsec (Mpc) which is equal to 3.26 million light years . Thus the units of the Hubble constant is (km/sec)/Mpc[27, 28].

This discovery is the beginning of the modern age of cosmology. Cepheid variables remain one of the best methods of measuring distances to galaxies and these variables are very important to determine the expansion rate and the age of the universe [28]. Also, one who would like to learn more about the Hubble's constant can see the studies[27- 36].

Now, the kinetic energy can be written in terms of the Hubble's constant as

$$T=\frac{1}{2}mv^2=\frac{1}{2}mH_0^2r^2 \tag{22}$$

where *r* is the distance from the inner spherical orbit of the universe to the outer spherical orbit of the universe, *m* is the mass included in this universe. Then, the total energy of the universe as a closed system is written as

$$E=\frac{1}{2}mH_0^2r^2(1+2\ln r/r_0) \tag{23}$$

Here, it is considered the universe as an closed system, and the energy of it is given by the Equation (23). In the model, the universe is expanding out through an undestroyable wall, and there are other universes between the undestroyable wall and the universe under consideration because our universe is expanding. A model of the universe is defined by Hawking as "the universe in a nutshell" [37]. Undestroyable wall is assumed to be a region with a too high energy which allow nothing to pass through or effect it in any situation. I would like to describe the motion of the universe we inside in as from the beginning of the Big Bang to outward direction till this undestroyable wall. Then it has to have an inward motion through to make ready the conditions of a new Big Bang explosion in outward direction again. Here the Big Bang can be considered in two stages. First, the Big Bang of the innermost Universe in the outward direction. Second, the Bing Bang of the outermost Universe in inward direction. And also, if our universe is expanding, then there are some of the universes outside our universe that are shrinking [37, 38], and since our universe is accelerating then there are other expanding universes inside it. The undestroyable wall covering these universes is the outer most region of them.

To simplify my model, I would like to say that the motions of the universes are like the motion of valve plungers of a vehicle, when some of them are in upward motion, the others are in downward motion. Upward motion refers to the expansions of the universes, downward motion to the shrinkage of the others. Or, it can be described as a motion of a spring moving forward and backward about its equilibrium point on a frictionless surface with a mass attached to its end. And, the universes are like the spherical shells inside each other with the regions that themselves are in motion except the outermost region of the outermost universe. The regions of each universe is covered by the two spherical orbits like that of each orbit of the electron move in. These universes are thought to have parallel spheres of shells. And, these parallel spherical shells are inside each other, one covered by the other. The parallel universes are also studied in [39-41].

IV. Conclusion

If the universes are closed or stable systems, the energy of the universes and the distances of them are related to each other for each universe by the Equation (23) derived from the virial theorem. Therefore, if they are the closed or stable systems, this equation should describe the motions of the universes. As a result, since the virial theorem describes the periodic motion of the closed systems, we can take the beginning of the Big Bang as the initial time of this periodic motion.

References

[1] Arslan, Hasan. "The Dirac Equation According to the Virial Theorem for a Poential V= n kr^n." *Advanced Studies in Theoretical Physics* 8.22 (2014): 983-989.

[2] Kuić, Domagoj. "Quantum mechanical virial theorem in systems with translational and rotational symmetry." *International Journal of Theoretical Physics* 52.4 (2013): 1221-1239.

[3] Stokes, J. Dustan, et al. "The virial theorem in graphene and other Dirac materials." *Philosophical Magazine Letters* 93.12 (2013): 672-679.

[4] Goldstein, H. "Classical Mechanics," Addison-Wesley Publishing Company, Inc., Reading, 1959.

[5] S. T. Thornton and J. B. Marion, "Classical Dynamics of Particles and Systems," Thomson Learning, Belmont, 2004.

[6] Arslan, Hasan. "The Distances in the Stable Systems Due to the Virial Theorem." *Applied Mathematics* 4.4 (2013).

[7] Hasan Arslan and Nihan Hulaguhanoglu; The Wavefunctions and Energy Eigenvalues of the Schrödinger Equation for Different Potentials Due to the Virial Theorem, SCITEED 2014, Curran Associates, Inc(2014), 57 Morehouse Lane Red Hook, NY 12571.

[8] Al-Khasawneh, Belal Yaseen, and Mohammad B. Altaie. *Investigating the Gravitational Properties of Dark Matter*. Diss. 2012.

[9] Nadareishvili, T., and A. Khelashvili. "Generalization of Hypervirial and Feynman-Hellmann Theorems for Singular Potentials." *arXiv preprint arXiv:0907.1824* (2009).

[10] Gurtler, R., and David Hestenes. "Consistency in the formulation of the Dirac, Pauli, and Schroedinger theories." *Journal of Mathematical Physics* 16.3 (2008): 573-584.

[11] Wuk Namgung. " Virial Theorem, Feynman-Hellman Theorem, and Variational Method." JKPS 1998 32:647-650

[12] Bahcall, John N. "Virial Theorem for Many-Electron Dirac Systems." *Physical Review* 124 (1961): 923-924.

[13] Ru-Zeng, Zhu, Wen Yu-Hua, and Qian Jin. "The virial theorem in refined Thomas-Fermi-Dirac theory for the interior of atoms in a solid." *Chinese Physics* 11.11 (2002): 1193-1195.

[14] Weislinger, Edmond, and Gabriel Olivier. "The classical and quantum mechanical virial theorem." *International Journal of Quantum Chemistry* 8.S8 (1974): 389-401.

[15] Weislinger, Edmond, and Gabriel Olivier. "The virial theorem with boundary conditions applications to the harmonic oscillator and to sine-shaped potentials." *International Journal of Quantum Chemistry* 9.S9 (1975): 425-433.

[16] Kalman, G., V. Canuto, and B. Datta. "Self-consistency condition and high-density virial theorem in relativistic many-particle systems." *Physical Review D* 13 (1976): 3493-3494.

[17] Rosicky, F., and F. Mark. "The relativistic virial theorem by the elimination method and nonrelativistic approximations to this theorem." *Journal of Physics B: Atomic and Molecular Physics* 8.16 (1975): 2581.

[18] Barshalom, A., and J. Oreg. "The relativistic virial theorem in plasma EOS calculations." *High Energy Density Physics* 5.3 (2009): 196-203.

[19] Arslan, Hasan. "The Dirac Equation with the Scattered Electron Including Extra Potential Energy Comes from the Virial Theorem." *Journal of Modern Physics* 4.4 (2013).

[20] Arslan, Hasan. "A Unified Equation of Interactions." *Open Journal of Microphysics* 1 (2011): 28.

[21] Lucha, Wolfgang, and Franz F. Schöberl. "Relativistic virial theorem." *Physical review letters* 64.23 (1990): 2733-2735.

[22] Shabaev, V. M. "Virial relations for the Dirac equation and their applications to calculations of H-like atoms." *arXiv preprint physics/0211087* (2002).

[23] Semay, Claude. "Virial theorem for two-body Dirac equation." *Journal of mathematical physics* 34.5 (1993): 1791-1793.

[24] March, N. H. "The Viral Theorem for Dirac's Equation." *Physical Review* 92 (1953): 481-482.

[25] Rose, M. E. and T. A. Welton. "The virial theorem for a Dirac particle." *Physical Review* 86.3 (1952): 432.

[26] Balinsky, A. A., and W. D. Evans. "On the virial theorem for the relativistic operator of Brown and Ravenhall, and the absence of embedded eigenvalues." *Letters in Mathematical Physics* 44.3 (1998): 233-248.

[27] Freedman, Wendy L., and Barry F. Madore. "The Hubble Constant." *arXiv preprint arXiv:1004.1856* (2010).

[28] Freedman, Wendy L., and Long Long Feng. "Determination of the Hubble constant." *Proceedings of the National Academy of Sciences* 96.20 (1999): 11063-11064.

[29] R. K. Mishra and Amritbir Singh, *Int. J. Pure Appl. Sci. Technol.*, 5(1) (2011), pp. 1-8.

[30] L.V.E. Koopmans, T.Treu, C.D. Fassnacht, R.D. Blandford, G. Surpi; THE HUBBLE CONSTANT FROM THE GRAVITATIONAL LENS B1608+656, ApJ, 2001.

[31] C.R. Keeton, C.S. Kochanek; Determining the Hubble Constant from the Gravitational Lens PG 1115+080, arXiv:astro-ph/9611216v2

[32] The Age of the Universe, http://cosmos.phy.tufts.edu/~zirbel/ast9/handouts/Age-of-universe.PDF.

[33] Brian Conway 08354502, Determination of Hubble's constant through least square's fitting method, http://www.maths.tcd.ie/~conwaybr/labs/astro_comp_ex1.pdf.

[34] Changjun Gao, A Model of Nonsingular Universe, www.mdpi.com/journal/entropy.

[35] Dr Lisa Jardine-Wright, Cavendish Laboratory, Hubble's Law-Measuring The Age of the Universe, http://wwwoutreach.phy.cam.ac.uk/resources/astro/KS5/hubble/Hubbles_Law.pdf

[36] Gonzalo A. Moreno Jimenez, Theoretical Calculation of the Hubble Constant and Relation to CMB and CIB, Galilean Electrodynamics, 2008, http://www.gonzaloamoreno.com/L_SIM_HUBBLEC.pdf.

[37] Stephan Hawking, The Universe in a Nutshell, 2001. London: Bantam Press.

[38] The Royal Swedish Academy of Sciences, The accelerating Universe, 2011,http://www.nobelprize.org/nobel_prizes/physics/laureates/2011/advanced-physicsprize2011.pdf.

[39] Tegmark, Max. "Parallel universes." *Science and ultimate reality* (2004): 459-491.

[40] Ryan, Marie-Laure. "From parallel universes to possible worlds: Ontological pluralism in physics, narratology, and narrative." *Poetics Today* 27.4 (2006): 633-674.

[41] KEVITIY, TIVE WORK OF N., ALISON CHANDLER, and G. MILLER. "The promise of parallel universes." (2007).

Hasan Arslan was born in Elbistan, in 1966. He graduated from METU Physics Department with BS degree in 1993, in Ankara, Turkey. He got Msc degree from Çukurova University Physics Department in 1998, PhD degree in the same Institution in 2008 with the completed related studies in Theoretical High Energy Physics in Adana, Turkey.
Hasan worked as an English, Maths, Science teacher in private schools in Adana-Turkey in 1994. He worked as an official teacher of English in National Education Ministry of Turkey from 1998 to 2009 in Adana-Turkey. He has been working in Bingöl University Physics Department from 2009 till now. Some of his works can be listed as:
1. Aydemir, A., H. Arslan, and A. K. Topaksu. "The estimation of the Z′ gauge boson mass in E 6 models." *Physics of Particles and Nuclei Letters* 6.4 (2009): 304-308.
2. Arslan, Hasan. "A Unified Equation of Interactions." *Open Journal of Microphysics* 1.02 (2011): 28.
3. Arslan, Hasan. "The Dirac Equation According to the Virial Theorem for a Poential n kr V=." *Advanced Studies in Theoretical Physics* 8.22 (2014): 983-989.

A Numerical Investigation of a Vortex Ring in a Rotating Fluid

Watchapon Rojanaratanangkule

***Abstract*—**The evolution of an axisymmetric vortex ring in a fluid rotating along the axis of the ring propagating direction is investigated by means of direct numerical simulation (DNS). The Reynolds number and the rotation number are defined from the initial circulation and the initial ring radius and are set to 5500 and 0.1, respectively. The axial vortex is observed to shed from the vortex core, leading to the development of the primary and secondary instabilities of the vortex ring. The observed primary instability simultaneously deforms the cross-section of the vortex ring and twists the toroid of the ring about its centreline along the ring circumference.

***Keywords*—**Vortex ring, Rotating fluid, Direct numerical simulation.

I. Introduction

LARGE-SCALE vortical structures are of interest for engineers and physicists since they are the basic features of many turbulent flows. Both experimental and numerical studies have observed that such coherent structures are organised and appear in the form of hairpins, lines, loops or rings. In order to obtain a deeper understanding of turbulence physics and control turbulence phenomena, the dynamics of a vortex ring has been investigated as a simple model of the various vortical interactions through theoretical, experimental and numerical perspectives due to the simplicity of its geometry [1]. Understanding their dynamics, fundamental properties and how they are affected by and interact with various backgrounds in which they can exist (e.g. stably stratified or rotating backgrounds) will help us obtain deeper understanding of turbulence. In this work, the formation and evolution of a single vortex ring in a rotating fluid, whose axis of rotation is parallel to the direction of translation of the ring, will be investigated to explore the effect of the background rotation on the vortex ring evolution.

Verzicco *et al.* [2] investigated the dynamics of a vortex ring in a rotating fluid via numerical simulations and laboratory experiments at low Reynolds numbers ($\mathcal{O}(1000)$). Their results delineated the evolution of the ring into two regimes depending on the angular speed of the rotating system. For low-rotation regime, the dynamical structures of the ring does not differ much from that in a non-rotating fluid. One distinct effect of the background rotation is that it introduces an azimuthal (swirl) velocity to the ring leading to the appearance of an elongated axial vortex. Once the rotation rate exceeds its critical value, the Coriolis force due to the background rotation suppresses the formation of the ring. Additionally, there appears the radiation of the energy of the ring by the inertial waves. Brend & Thomas [3] performed a set of experiments to quantify the decay length of the vortex ring as a function of the rotation number. Their experimental data can be used to roughly approximate how long the vortex ring can propagate through the fluid with background rotation before it decays.

The vortex ring in a rotating fluid possesses some qualitatively similar characteristics with the ring in a non-rotating fluid but with an assigned azimuthal velocity (referred to as a vortex ring with swirl), especially the existence of the axial vortex. Recent numerical experiments of Balakrishnan [4] showed that the ring with swirl exhibits a maximum limit of the amount of swirl. If the swirl strength, measured in terms of angular impulse, is larger than the limit, the ring will rapidly eject the fluid with an azimuthal velocity from the vortex core. This rapid ejection is directed radially outward and downstream of the ring similar to a jet flow. When the Reynolds number is high enough, a helical instability develops in the ring due to the presence of swirl. This helical instability is different from an azimuthal instability (Widnall instability [5]) of the non-swirling ring in such a way that the helical instability simultaneously deforms and twists the vortex core.

The aim of the present work is to extend the study of Verzicco *et al.* [2] to a higher Reynolds number ($\mathcal{O}(5000)$) to investigate whether the ring in a rotating fluid at a low-rotation regime can develop any new features.

II. Numerical Approach

In the present work, a single vortex ring with radius R and core radius δ is considered. While the ring propagates along the positive z-direction, the computational domain is rotated in the axial direction with a constant angular velocity $\Omega_i = (0, 0, \Omega_z)$. The evolution of the ring is governed by the continuity and the incompressible Navier–Stokes equations formulated in a translating and rotating frame of reference. The governing equations in a Cartesian coordinate system, $x_i = (x, y, z)$, can be written as

$$\frac{\partial u_i}{\partial x_i} = 0\,, \tag{1}$$

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = -\frac{1}{\rho}\frac{\partial P_{\text{eff}}}{\partial x_i} + \nu \frac{\partial^2 u_i}{\partial x_j \partial x_j} - 2\epsilon_{ijk}\Omega_j u_k - \frac{\mathrm{d}U_F}{\mathrm{d}t}\delta_{3i}\,, \tag{2}$$

where $u_i = (u, v, w)$ is the velocity vector at time t, ϵ_{ijk} is the Levi–Civita symbol and δ_{ij} is the Kronecker delta. The effective pressure P_{eff} includes the thermodynamics pressure and the centrifugal force. The translating speed of the moving

This work is financially supported by Chiang Mai University.

W. Rojanaratanangkule is with the Department of Mechanical Engineering, Chiang Mai University, Chiang Mai 50200, Thailand (e-mail: watchapon.roj@eng.cmu.ac.th).

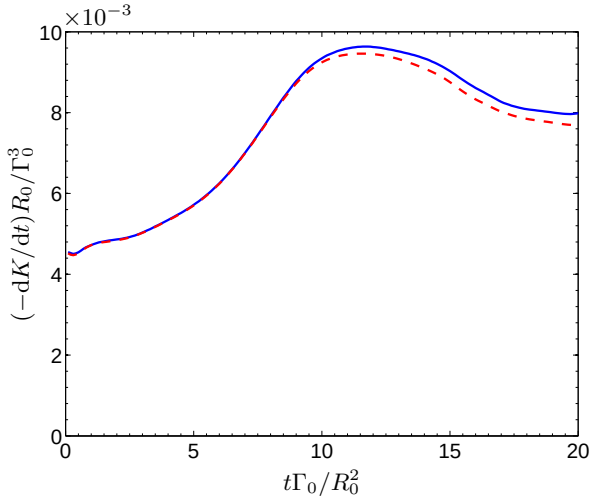


Fig. 1. History of rate of change of volume–integrated kinetic energy: —— $-\mathrm{d}K/\mathrm{d}t$; - - - - $\epsilon_K + F_K$.

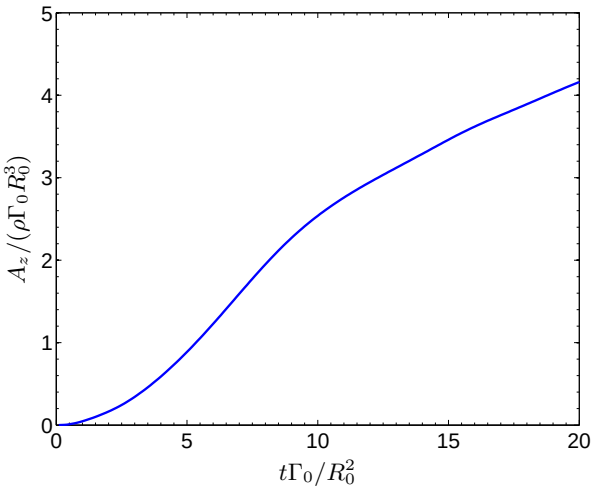


Fig. 2. Evolution of the angular impulse of a vortex ring in a rotating fluid for $Re = 5500$ and $Ro = 0.1$.

frame U_F is equivalent to the propagating velocity of the ring and is determined via a proportional-integral (PI) controller of Archer *et al.* [6]. During the calculations, the fluid properties (density ρ and kinematic viscosity ν) are kept constant.

The Navier–Stokes equations are numerically solved without any turbulence model (referred to as direct numerical simulation, DNS) on a staggered grid with second-order finite differencing in space and Adams–Bashforth stepping in time [7]. The ring with initial radius R_0, circulation Γ_0 and slenderness ratio, δ_0/R_0, of 0.2 is initiated at the centre of the domain at $x_i = (0, 0, 0)$ with a Gaussian distribution of azimuthal vorticity. Azimuthal perturbation in the form of a sum of 32 Fourier modes with amplitude of $0.0004R_0$ and random phase is applied to the ring radius. The initial velocity field can be obtained by means of a vorticity–vector stream function method (see, e.g., [8]). The Reynolds number $Re = \Gamma_0/\nu$ and the rotation number $Ro = 2\Omega_z R_0^2/\Gamma_0$ are set to 5500 and 0.1, respectively. It should be noted that the definition of our Re and Ro is different from that of Verzicco *et al.* [2], who defined the Reynolds number and the rotation number based on the centreline ejection velocity and the radius of the orifice. They demonstrated that their Reynolds/rotation number is about 1.5 times lower/higher than those defined from the initial circulation and ring radius. The simulations are performed in a cuboidal domain of size $L_x = L_y = L_z = 8R_0$ with the grid resolution of 256^3. The time-dependent uniform inflow velocity U_F together with a zero vorticity condition is employed at $z = +L_z/2$, while a zero gradient condition for the velocity field is applied at the outflow plane ($z = -L_z/2$). A periodic boundary condition is specified at the x- and y-directions.

III. Results

This section presents the results from the DNS of a single vortex ring in a rotating fluid. We verify the adequacy of the grid resolution in Sec. III-A. The effect of the background rotation on the development of the vortex ring is explored in Sec. III-B.

A. Resolution Check

The adequacy of the grid resolution used is verified by comparing the left- and the right-hand side of the volume-integrated instantaneous kinetic-energy equation, written as

$$-\frac{\mathrm{d}K}{\mathrm{d}t} = \epsilon_K + F_K\,, \tag{3}$$

where $K = 0.5\int_V \left(u_i u_i - U_F^2\right)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$ is the volume-integrated kinetic energy in a co-moving frame of reference, ϵ_K is the volume-integrated rate of kinetic energy dissipation and F_K is the net volume-integrated kinetic energy flux. It should be noted that the work done due to the Coriolis force is zero, $\boldsymbol{u}\cdot(2\boldsymbol{\Omega}\times\boldsymbol{u}) = 0$, since it is a fictitious force. The difference between the rate of change of the volume-integrated kinetic energy $\mathrm{d}K/\mathrm{d}t$ and the RHS of (3) is illustrated in Fig. 1. The difference between the two sides of (3) is less than 1% up to $t \approx 10R_0^2/\Gamma_0$, when the ring begins to develop a three-dimensional instability leading to the breakdown to turbulence. During that period, reasonable accuracy is obtained with the maximum error being less than 3%, indicating that the spatial resolution is fine enough to accurately capture all scales of the flow.

B. Effect of the Coriolis force

The strength of the swirl can be measured via an integral quantity namely angular impulse $\boldsymbol{A} = (A_x, A_y, A_z)$, defined as [9]

$$\boldsymbol{A} = \frac{\rho}{3}\int_V \boldsymbol{x}\times(\boldsymbol{x}\times\boldsymbol{\omega})\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,, \tag{4}$$

where $\boldsymbol{\omega} = \boldsymbol{\nabla}\times\boldsymbol{u}$ is the vorticity vector. The angular impulse can be interpreted as the resultant moment of the impulsive force that generates the motion from rest and it is invariant in an unbounded domain [9]. For an axisymmetric ring, it possesses only the angular impulse in the axial direction. Figure 2 displays the history of the angular impulse of the vortex ring in a rotating fluid at $Ro = 0.1$. It can be seen that the angular impulse keeps increasing due to an endless supply of the Coriolis force. This is the major different between the

ISBN: 978-1-61804-292-7

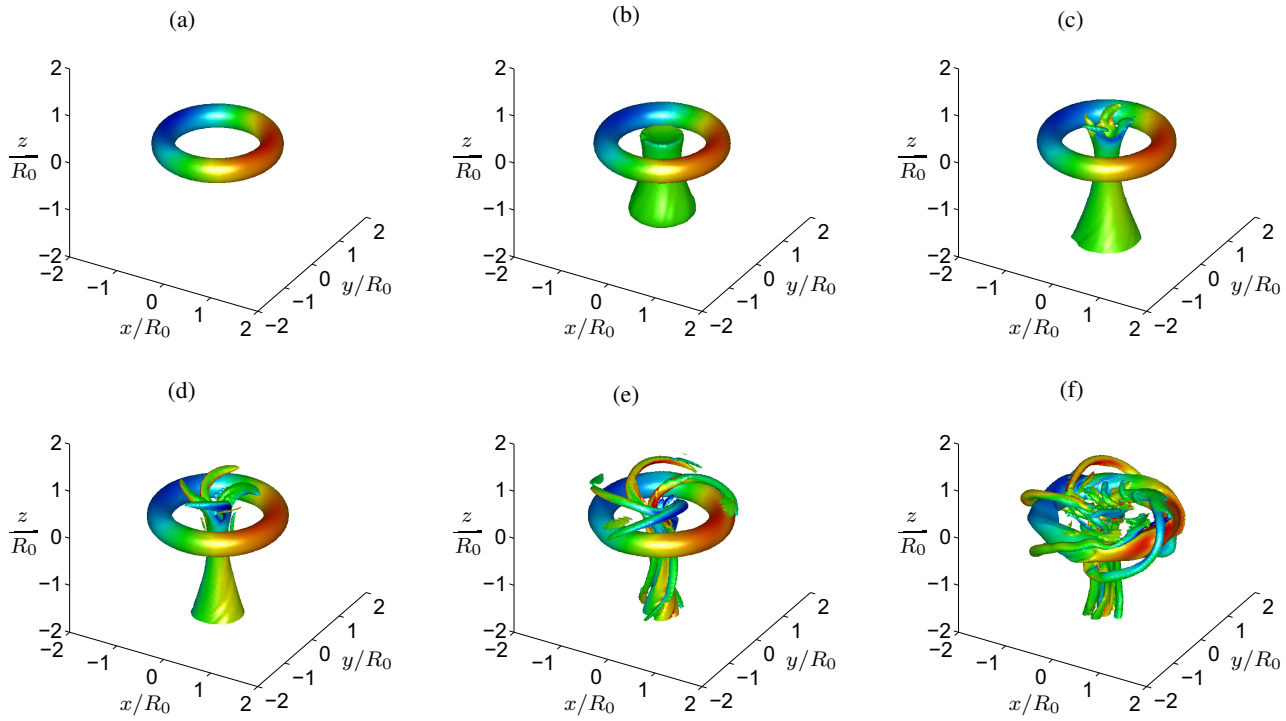


Fig. 3. Three-dimensional isosurfaces of the second invariant of the velocity gradient tensor Q, coloured by ω_y, showing the evolution of the ring in a fluid rotating in the axial direction with $Ro = 0.1$: (a) $t\Gamma_0/R_0^2 = 1$, (b) $t\Gamma_0/R_0^2 = 6$, (c) $t\Gamma_0/R_0^2 = 10$, (d) $t\Gamma_0/R_0^2 = 12$, (e) $t\Gamma_0/R_0^2 = 15$ and (f) $t\Gamma_0/R_0^2 = 20$. Surface level $QR_0^4\Gamma_0^2 = 0.005$

vortex ring in a rotating fluid and the ring with swirl since the angular impulse of the latter is conserved. It is of interest to note that the growth rate of the angular impulse seems to decrease after $t \approx 10R_0^2/\Gamma_0$, indicating that the ring would evolve with different dynamics after that time.

The evolution of the ring in a rotating fluid is visualised by means of the second invariant of the velocity gradient tensor $Q = -0.5u_{i,j}u_{j,i}$ (see, e.g., [10] for details), illustrated in Fig. 3. The contours of the azimuthal vorticity ω_θ on the plane $\theta = 0$ are also depicted in Fig. 4 to aid the analysis of the results. Initially, the Coriolis force does not have much influence on the development of the ring. Hence, the vortex ring remains axisymmetric and laminar, as displayed in Figs. 3(a) and 4(a). With time, the ring sheds the swirling wake along the axial direction, as illustrated in Fig. 3(b). This swirling wake is usually referred to as the axial vortex. It can be seen from Fig. 4(b) that the axial vortex consists of the negative and positive azimuthal vorticity at respectively the front and rear part of the vortex core. The existence of this additional vorticity structure can be explained by investigating the transport equation of the azimuthal component of vorticity. In the presence of the background rotation in the axial direction and the assumption of axisymmetric flow, the azimuthal vorticity transport equation can be written as

$$\frac{\partial \omega_\theta}{\partial t} + u_r\frac{\partial \omega_\theta}{\partial r} + u_z\frac{\partial \omega_\theta}{\partial z} + \frac{u_\theta \omega_r}{r} = \omega_r\frac{\partial u_\theta}{\partial r} + \omega_z\frac{\partial u_\theta}{\partial z} + \frac{\omega_\theta u_r}{r} + \nu\left(\boldsymbol{\nabla}^2\omega_\theta - \frac{\omega_\theta}{r^2}\right) + 2\Omega_z\frac{\partial u_\theta}{\partial z}, \quad (5)$$

where the term $2\Omega_z\partial u_\theta/\partial z$ represents the effect of the Coriolis force and illustrates that the axial gradient of azimuthal velocity plays an important role in the evolution of the azimuthal vorticity. Figure 5 displays the contour of the azimuthal velocity u_θ on the plane $\theta = 0$ at $t\Gamma_0/R_0^2 = 6$. As the flow follows the conservation of angular momentum (ru_θ), the azimuthal velocity increases as $1/r$ when the fluid particles move radially inward. However, u_θ must be zero at $r = 0$ resulting in a decrease of u_θ near the axis consistent with the flow map in Fig. 5. It can also be seen that the azimuthal velocity possesses a positive axial gradient ($\partial u_\theta/\partial z > 0$) behind the vortex core. This results in a positive azimuthal vorticity. On the other hand, the negative ω_θ at the front half of the ring is due to the negative $\partial u_\theta/\partial z$ at that region.

At $t\Gamma_0/R_0^2 = 10$, the axial vortex begins to develop an instability, as depicted in Figs. 3(c) and 4(c). This instability grows with time and leads to the appearance of the four secondary (spiral) vortices in front of the vortex core (Figs. 3d and 4d). The axial vortex is distorted by these secondary vortices while the radial length of the spiral vortices increases with time, as illustrated in Figs. 3(e) and 4(e). At $t\Gamma_0/R_0^2 = 20$, the cross-section of the vortex core begins to deform (Figs. 4f). In the mean time, the toroid of the ring is twisted about its centreline along the ring circumference (visualised by the isosurface of Q, see Figs. 3f). The simultaneous deformation and twisting of the vortex core observed in this work are qualitatively similar to the helical instability occurring in a vortex ring with swirl [4]. The further investigation of the flow instabilities induced due to the system rotation is deferred to a future study.

ISBN: 978-1-61804-292-7

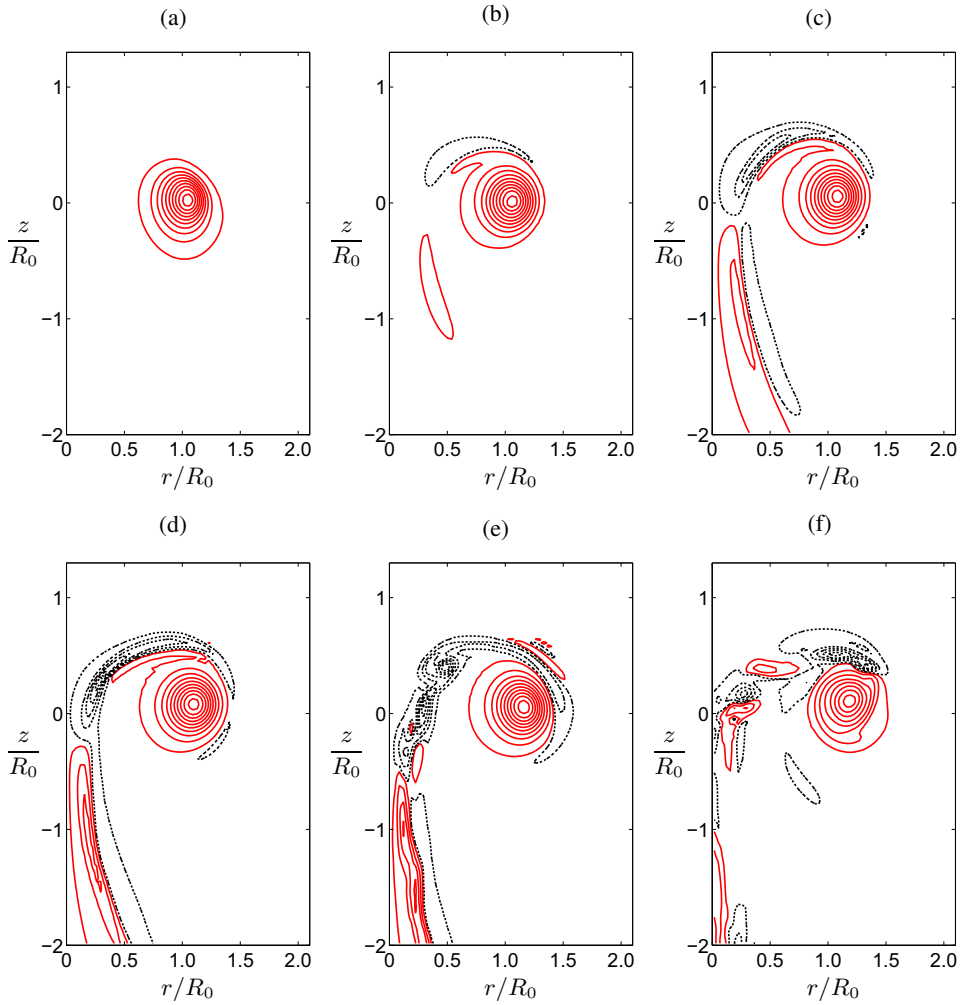


Fig. 4. Contours of azimuthal vorticity ω_θ on the plane $\theta = 0$ at $t\Gamma_0/R_0^2$ = (a) 1, (b) 6, (c) 10, (d) 12, (e) 15 and (f) 20. The lowest contour levels of $|\omega_\theta|_{\max}/40$ and equal spacing of $|\omega_\theta|_{\max}/10$ were used. Red solid lines and black dotted lines respectively show positive and negative vorticity.

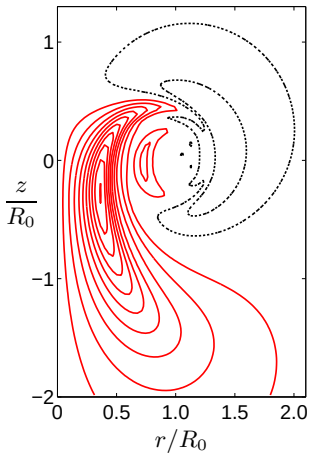


Fig. 5. Contour of azimuthal velocity u_θ on the plane $\theta = 0$ at $t\Gamma_0/R_0^2 = 6$. The lowest contour levels of $|u_\theta|_{\max}/12$ and equal spacing of $|u_\theta|_{\max}/10$ were used. Red solid lines and black dotted lines respectively show positive and negative velocity.

IV. Summary

Direct numerical simulation (DNS) has been employed to investigate the formation and evolution of a single vortex ring in a fluid that rotates along the axial direction. It is found that the swirling wake is shed downstream from the vortex core. The azimuthal vorticity of the axial vortex is positive at the rear part of the vortex core, and is negative at the front half. The appearance of the axial vortex is a result of the axial gradient of the azimuthal velocity as $\partial u_\theta/\partial z$ is the source for the azimuthal vorticity. Once the axial vortex has occurred, it develops a secondary instability appearing in the form of four spiral vortices in front of the vortex core. The primary instability is then developed, resulting in the simultaneous deformation of the vortex core cross-section and twisting of the vortex core about its centreline along the ring circumference. Future work will analyse and develop a mathematical description of the ring instability induced due to the system rotation.

Acknowledgment

The author would like to express his gratitude to the computational resources provided by the Faculty of Engineering, Chiang Mai University and the HPC services from the Large-scale Simulation Research Laboratory of National Electronics and Computer Technology center.

REFERENCES

[1] S. Shariff and A. Leonard, "Vortex rings," *Annu. Rev. Fluid Mech.*, vol. 24, pp. 235–279, 1992.

[2] R. Verzicco, P. Orlandi, A. H. M. Eisenga, G. J. F. van Heijst, and G. F. Carnevale, "Dynamics of a vortex ring in a rotating fluid," *J. Fluid Mech.*, vol. 317, pp. 215–239, 1996.

[3] M. A. Brend and P. J. Thomas, "Decay of vortex rings in a rotating fluid," *Phys. Fluids*, vol. 21, no. 4, p. 044105, 2009.

[4] S. K. Balakrishnan, *A numerical study of some vortex ring phenomena using direct numerical simulation (DNS)*. PhD thesis, University of Southampton, Southampton, UK, 2013.

[5] S. E. Widnall and C.-Y. Tsai, "The instability of the thin vortex ring of constant vorticity," *Phil. Trans. R. Soc. Lond.*, vol. 287, no. 1344, pp. 273–305, 1977.

[6] P. J. Archer, T. G. Thomas, and G. N. Coleman, "Direct numerical simulation of vortex ring evolution from the laminar to the early turbulent regime," *J. Fluid Mech.*, vol. 598, pp. 201–226, 2008.

[7] T. G. Thomas and J. J. R. Williams, "Development of a parallel code to simulate skewed flow over a bluff body," *J. Wind Eng. Ind. Aerodyn.*, vol. 67–68, pp. 155–167, 1997.

[8] W. E and J.-G. Liu, "Finite difference methods for 3D viscous incompressible flows in the vorticity–vector potential formulation on nonstaggered grids," *J. Comput. Phys.*, vol. 138, no. 1, pp. 57–82, 1997.

[9] G. K. Batchelor, *An Introduction to Fluid Dynamics*. Cambridge: Cambridge University Press, 1967.

[10] J. Jeong and F. Hussain, "On the identification of a vortex," *J. Fluid Mech.*, vol. 285, pp. 69–94, 1995.

ISBN: 978-1-61804-292-7

Unranking algorithms applied to MuPAD

X. Molinero and J. Vives

***Abstract*—We present an improvement of the implementation of some unlabeled unranking algorithms of the open-source algebraic combinatorics package MuPAD-Combinat for the computer algebra system MuPAD. We compare our implementation with the current one. Moreover, we have also developed unranking algorithms applied to some unlabeled admissible operators that are not still implemented in the package MuPAD-Combinat. These algorithms are also able to develop some structures useful to generate molecules applied to chemistry and influence graphs applied to game theory and social networks, among other topics.**

***Index Terms*—Unranking Algorithms, MuPAD, Generating Molecules, Generating Influence Games.**

Unlabeled class	Specification
Binary trees	$\mathcal{B} = Z + \mathcal{B} \times \mathcal{B}$
Unary-binary trees or Motzkin trees	$\mathcal{M} = Z + Z \times \mathcal{M} + Z \times \mathcal{M} \times \mathcal{M}$
Integer partitions	$\mathcal{P} = \mathsf{Set}(\mathsf{Seq}(Z, \text{card} \geq 1))$
Integer compositions	$\mathcal{C} = \mathsf{Seq}(\mathsf{Set}(Z, \text{card} \geq 1))$
Non-ordered rooted trees or Rooted unlabeled trees	$\mathcal{T} = Z \times \mathsf{Set}(\mathcal{T})$
Binary sequences	$\mathcal{A} = \mathsf{Seq}(Z + Z)$
Non plane ternary trees	$\mathcal{D} = Z + \mathsf{Set}(\mathcal{D}, \text{card}{=}3)$
Integer partitions with distinct parts	$\mathcal{E} = \mathsf{PowerSet}(\mathsf{Seq}(Z, \text{card}{\geq}1))$

Fig. 1. Examples of unlabeled classes and their specifications

I. Introduction

The problem of *unranking* asks for the generation of the ith combinatorial object of size n in some combinatorial class $\mathcal{A}$, according to some well defined order among the objects of size n of the class. Efficient unranking algorithms have been devised for many different combinatorial classes, like binary and Cayley trees, Dyck paths, permutations, strings or integer partitions, but most of the work in this area concentrates in efficient algorithms for particular classes, whereas we aim at generic algorithms that apply to a broad family of combinatorial classes. The problem of unranking is intimately related with its converse, the *ranking* problem, as well as with the problems of random generation and exhaustive generation of all combinatorial objects of a given size. The interest of this whole subject is witnessed by the vast number of research papers and books that has appeared in over five decades (see, for instance, [24], [12], [9], [8], [11], [25], [10], [20], [19], [21], [3]).

[14], [13] designed *generic* unranking algorithms for a large family of combinatorial classes, namely, those which can be inductively built from the basic ϵ-class (a class which contains only one object of size 0), atomic classes (classes that contain only one object of size 1 or *atom*) and a collection of admissible combinatorial operators: disjoint unions, labeled and unlabeled products, sequence, set, etc. Now we use such techniques to implement those algorithms in MuPAD [2], [18]. In the open-source algebraic combinatorics package MuPAD-Combinat [1] for the computer algebra system MuPAD there are implemented the unranking for some admissible combinatorial operators, but now we improve such implementation for unlabeled unions and products (and sequences). Moreover, we have also implemented other operators as unlabeled sets and powersets (with and without restrictions).

The paper just considers unlabeled combinatorial classes and it is organized as follows. In Section II we briefly review basic definitions and concepts, the unranking algorithms and the theoretical analysis of their performance. Afterwards, from the computer algebra system MuPAD, we compare the required CPU time of our implementation with the required CPU time of the current implementation in the package MuPAD-Combinat. Moreover, we also explain our current and future work in this subject.

II. Preliminaries

As it will become apparent, all the unranking algorithms in this paper require an efficient algorithm for counting, that is, given a specification of a class and a size, they need to compute the number of objects with the given size. Hence, we will only deal with (some of) the so-called *admissible combinatorial classes* [6], [7]. Those are constructed from *admissible operators*, operations over classes that yield new classes, and such that the number of objects of a given size in the new class can be computed from the number of objects of that size or smaller sizes in the constituent classes. In this paper we just consider unlabeled objects (those whose atoms are indistinguishable[1]) built from these admissible combinatorial operators.

For unlabeled classes, the finite specifications are generated from the ϵ-class, atomic classes, and combinatorial operators including disjoint union ('+'), Cartesian product ('$\times$'), sequence ('Seq'), powerset ('PowerSet'), set ('Set')[2], and sequence, powerset and set (or multiset) with restricted cardinality. Figure 1 gives a few examples of unlabeled admissible classes.

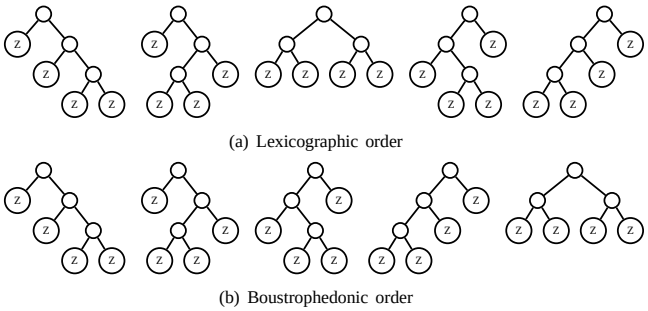
X. Molinero is with the Department of Applied Mathematics III, Universitat Politècnica de Catalunya, E-08240 Manresa, SPAIN. E-mail: xavier.molinero@upc.edu. X. Molinero was partially funded by grant MTM2012-34426/FEDER of the "Spanish Economy and Competitiveness Ministry".

J. Vives is with the Department of Design and Programming of Electronic Systems, Universitat Politècnica de Catalunya, E-08240 Manresa, SPAIN. E-mail: jvives@epsem.upc.edu.

[1] On the contrary, each of the n atoms of a *labeled* object of size n bears a distinct label drawn from the numbers 1 to n.

[2] Also denoted by 'multisets' (MultiSet) to emphasize that repetition is allowed.

ISBN: 978-1-61804-292-7



(a) Lexicographic order

(b) Boustrophedonic order

Fig. 2. Binary trees of size 4.

For the rest of this paper, we will use calligraphic uppercase letters to denote classes: $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, ... Given a class $\mathcal{A}$ and a size n, $\mathcal{A}_n$ will denote the subset of objects of size n in $\mathcal{A}$.

The order $\prec_{\mathcal{C}_n}$ among the objects of size n for a class $\mathcal{C} = \mathcal{A} + \mathcal{B}$ is naturally defined by $\gamma \prec_{\mathcal{C}_n} \gamma'$ if both γ and γ' belong to the same class (either $\mathcal{A}_n$ or $\mathcal{B}_n$) and $\gamma \prec \gamma'$ within their class, or if $\gamma \in \mathcal{A}_n$ and $\gamma' \in \mathcal{B}_n$. It is then clear that although $\mathcal{A} + \mathcal{B}$ and $\mathcal{B} + \mathcal{A}$ are isomorphic ("the same class"), these two specifications induce quite different orders. The unranking algorithm for disjoint unions compares the given rank with the cardinality of $\mathcal{A}_n$ to decide if the sought object belongs to $\mathcal{A}$ or to $\mathcal{B}$ and then solves the problem by recursively calling the unranking on whatever class ($\mathcal{A}$ or $\mathcal{B}$) is appropriate.

For Cartesian products the order in $\mathcal{C}_n = (\mathcal{A} \times \mathcal{B})_n$ depends on whether $\gamma = (\alpha, \beta)$ and $\gamma' = (\alpha', \beta')$ have first components of the same size. If $|\alpha| = |\alpha'| = j$ then we have $\gamma \prec_{\mathcal{C}_n} \gamma'$ if $\alpha \prec_{\mathcal{A}_j} \alpha'$ or $\alpha = \alpha'$ and $\beta \prec_{\mathcal{B}_{n-j}} \beta'$. But when $|\alpha| \neq |\alpha'|$, we must provide a criterion to order γ and γ'. The *lexicographic* order stems from the specification

$$\mathcal{C}_n = \mathcal{A}_0 \times \mathcal{B}_n + \mathcal{A}_1 \times \mathcal{B}_{n-1} + \ldots + \mathcal{A}_n \times \mathcal{B}_0,$$

in other words, the smaller object is that with smaller first component. On the other hand, the *boustrophedonic* order is induced by the specification

$$\mathcal{C}_n = \mathcal{A}_0 \times \mathcal{B}_n + \mathcal{A}_n \times \mathcal{B}_0 + \mathcal{A}_1 \times \mathcal{B}_{n-1} + \mathcal{A}_{n-1} \times \mathcal{B}_1 + \mathcal{A}_2 \times \mathcal{B}_{n-2} + \mathcal{A}_{n-2} \times \mathcal{B}_2 + \ldots,$$

in other words, we consider that the smaller pairs of total size n are those whose $\mathcal{A}$-component has size 0, then those with $\mathcal{A}$-component of size n, then those with $\mathcal{A}$-component of size 1, and so on. Figure II shows the lists of unlabeled binary trees of size 4 in lexicographic (a) and boustrophedonic order (b).

Of course, other orders are also possible, but they either do not help improving the performance of unranking or they are too complex to be useful or of general applicability.

For powersets, among some natural orders (see [14], [16]) we can choose

$$\mathsf{PowerSet}(\mathcal{A}) = \epsilon + \bigcup_{n>0} \bigcup_{j=1}^{n} \bigcup_{k=n\div j}^{1} \Big(\mathsf{PowerSet}(\mathcal{A}_j, \text{card}= k) \times \mathsf{PowerSet}_{n-kj}(\mathcal{A}_{>j})\Big)$$

where

$$\mathsf{PowerSet}(\mathcal{A}_j, \text{card}= k) = \bigcup_{\alpha \in \mathcal{A}_j} \Big(\alpha \times \mathsf{PowerSet}(\mathcal{A}_j^{(\succ \alpha)}, \text{card}= k-1)\Big),$$

being $\mathcal{A}^{(\succ \alpha)} = \{\alpha' \in \mathcal{A} : \alpha' \succ \alpha\}$, and $\mathsf{PowerSet}(\mathcal{A}_{>j})$ is a powerset with $\mathcal{A}$-components of size at least equal to $j+1$. Other orders described in [14], [16] do not change the complexity and they could also be easily adapted to our implementation.

For sets we have analogous isomorphisms but allowing repetitions.

The theoretical performance of these unranking algorithms is summarized in [16], [13].

Theorem 1: The worst-case time complexity of unranking for objects of size n in any admissible labeled class $\mathcal{A}$ using lexicographic ordering is of $\mathcal{O}(n^2)$ arithmetic operations.

Theorem 2: The worst-case time complexity of unranking for objects of size n in any admissible labeled class $\mathcal{A}$ using boustrophedonic ordering is of $\mathcal{O}(n \log n)$ arithmetic operations.

III. Our implementation v.s. MuPAD-Combinat implementation

In this section we compare our implementation[3] for unranking in MuPAD with the current implementation of the package MuPAD-Combinat(using MuPAD Pro 4.0). All our experiments run under Linux in a AMD64X2 4400 at 2.2 GHz with 4 Gb of RAM, and they use the basic facilities for counting already provided by the package MuPAD-Combinat.

For instance, the interface for binary trees has the following inputs:

```
spec := {B = Union(Z, BB), BB = Prod(B, B)};
p1 := combinat::
      decomposableObjects(spec, Lexi/Bous);
p1::unrank(rank, size);
```

where `spec` is the specification[4], `Lexi` or `Bous` forces the lexicographic or boustrophedonic order, respectively, and `rank` and `size` are the considered rank and size, respectively. Thus, the following commands provide all binary trees of size 8 in lexicographic order:

```
spec := {B = Union(Z, BB), BB = Prod(B, B)};
p1 := combinat::
      decomposableObjects(spec, Lexi);
for i from 0 to p1::count(8) - 1 do
  p1::unrank(i, 8);
end_for
```

[3]It is available on request from the first author; send an E-mail to xavier.molinero@upc.edu.

[4]The first class defined in the specification is the considered class (`B` in this case).

Lexicographic order				Boustrophedonic order			
Size	τ_T	τ'_T	ρ	Size	τ_T	τ'_T	ρ
25	1.24	4.55	0.27	25	0.90	2.95	0.30
50	3.63	11.95	0.30	50	2.50	6.52	0.38
75	6.08	20.17	0.30	75	4.27	10.24	0.41
100	9.27	29.65	0.31	100	6.20	13.97	0.44
125	12.46	39.98	0.31	125	6.99	17.91	0.39
150	16.54	51.73	0.31	150	8.39	20.80	0.40
175	21.37	66.16	0.32	175	9.67	24.90	0.38
200	26.85	79.93	0.33	200	11.63	28.87	0.40

TABLE I
AVERAGE CPU TIME (IN MILLISECONDS) FOR BINARY TREES, $\mathcal{B}$=Z+$\mathcal{B}\times\mathcal{B}$. THE CPU TIME REQUIRED TO CALCULATE THE PRE-COMPUTED TABLES IN OUR UNRANKING IS 380 MILLISECONDS.

Lexicographic order				Boustrophedonic order			
Size	τ_T	τ'_T	ρ	Size	τ_T	τ'_T	ρ
25	1.12	3.24	0.34	25	1.06	2.57	0.41
50	2.60	9.33	0.27	50	2.71	5.75	0.47
75	4.81	16.61	0.28	75	4.19	9.74	0.43
100	7.18	25.72	0.27	100	5.95	13.76	0.43
125	9.65	34.27	0.28	125	8.01	17.66	0.45
150	12.28	42.57	0.28	150	9.80	22.43	0.43
175	15.42	54.44	0.28	175	11.49	27.76	0.41
200	18.83	67.08	0.28	200	13.56	32.67	0.41

TABLE II
AVERAGE CPU TIME (IN MILLISECONDS) FOR MOTZKIN TREES, $\mathcal{M}$=Z+Z$\times\mathcal{M}$+Z$\times\mathcal{M}\times\mathcal{M}$. THE CPU TIME REQUIRED TO CALCULATE THE PRE-COMPUTED TABLES IN OUR UNRANKING IS 690 MILLISECONDS.

Notice that, in general, `p1::count(size)` returns the number of objects of `p1` with size `size`.

The selected collection for our experiments are two classical classes: binary trees ($\mathcal{B} = Z + \mathcal{B} \times \mathcal{B}$) and, unary-binary trees or Motzkin trees ($\mathcal{M} = Z + Z \times \mathcal{M} + Z \times \mathcal{M} \times \mathcal{M}$).

Essentially, we have used two techniques in our implementation. First, we have appropriately used the command `option remember`. Second, we have also used some pre–computed tables to store the counting of each considered class and size. The access to the indices of such tables is notably faster than the access to the command `count`. Tables I and II show the improvement of the average CPU time (in milliseconds). We have pre–computed the counting tables and, afterwards, we have generated 10000 random objects of the considered class and size. τ_T is our average time required to unrank a random rank of the considered class and size, τ'_T is MUPAD-COMBINAT average time required to unrank a random rank of the considered class and size, and ρ is the ratio τ_T/τ'_T. For any case, it looks as the improvements tend to be stable when the size n increase. For lexicographic binary trees it approaches to $\rho = 0.33$, for Boustrophedonic binary trees it approaches to $\rho = 0.40$, for lexicographic Motzkin trees it approaches to $\rho = 0.28$, and for Boustrophedonic Motzkin trees it approaches to $\rho = 0.41$. Thus, all results are satisfactorily better. Note that even the pre–computed tables require some CPU time, the average CPU time (when the number of generated objects increase) of our implementation substantially improves the previous one. We have meaningfully improved the average CPU time required to generate a random unranking: In lexicographic order, our implementation spends about 30% of the CPU time of the previous version; and, in boustrophedonic order, it spends about 40% of the CPU time of the previous version.

Sequences are done from unions and products, thus the timing improvements have similar advantages.

On the other hand, we have also done some experiments with classes that involve sets or cycles, for instance, we have considered the so-called *functional graphs* defined by $\mathcal{F}$= Set(Cycle($\mathcal{T}$)) with $\mathcal{T}$=$\times$(Z, Set($\mathcal{T}$)). In such cases, our implementation spends between 60% and 80% of the CPU time of the previous version.

A. Future implementation

The current implementation in MUPAD-COMBINAT does not consider all admissible combinatorial operators as well as restricted cardinalities in sets or powersets. We have added some of these operators in our implementation. In particular, we have considered admissible operators like

$$\varphi(\mathcal{B}, \operatorname{card} \tau\, k)$$

where $\varphi \in \{\mathsf{Seq}, \mathsf{Set}, \mathsf{PowerSet}\}$, $\tau \in \{\leq\ /\ =\geq\}$ and $k \in \mathbb{N}$.

By the way, the required average CPU time for the implemented operators is clearly competitive. Now one of the following open problems is to develop the corresponding pre–computed tables of counting for powersets and sets (see the described isomorphisms for powersets and sets in Section II).

IV. CONCLUSIONS AND FUTURE WORK

We have implemented in MUPAD the unranking applied to some basic unlabeled admissible combinatorial operators: disjoint unions, Cartesian products, and sequences. We are now working on the implementation for (unlabeled) powersets and sets (with and without restricted cardinalities).

Our implementation is making two main improvements for the unranking of unlabeled admissible classes in front of the implementation in the package MUPAD-COMBINAT. First, we have significantly reduced the average CPU time required to generate a random unranking. Second, we are programming more unlabeled admissible combinatorial operators (powersets and sets with and without restricted cardinalities).

Future work is to implement even more unlabeled admissible combinatorial operators (substitution, the open problem for unlabeled cycles, the union among non-disjoint classes, the intersection among classes, etc.).

Another line of research is to study similar operators but from the labeled point of view, that is, to consider that the nodes of the combinatorial structures of size n can be distinguished by labels from 1 to n.

The ranking, exhaustive and random generation should also be implemented [24], [12], [16].

On the other hand, these algorithms are also able to develop some structures useful to generate molecules [4], [5] applied to chemistry and influence graphs [17] applied to game theory and social networks, among other topics [22].

Finally, to what we know, it is still open to study the unranking, ranking and exhaustive generation of combinatorial structures from the viewpoint of genetic algorithms [15], [23]. Thus, it should be very interesting to establish some genetic algorithms to solve these problems.

ISBN: 978-1-61804-292-7

ACKNOWLEDGMENTS

We thank anonymous referees for their useful comments and suggestions that helped us to improve the contents of the paper.

REFERENCES

[1] MuPAD-Combinat – open-source algebraic combinatorics package for the computer algebra system MUPAD. URL: http://mupad-combinat.sourceforge.net/.

[2] C. Creutzig and W. Oevel. MUPAD *Tutorial*. SciFace Software (SciFace), Paderborn, 2004.

[3] S. Even. *Combinatorial Algorithms*. MacMillan, New York, 1973.

[4] P. Flajolet and B. Salvy. Computer algebra libraries for combinatorial structures. *J. Symbolic Computation*, 20:653–671, 1995.

[5] P. Flajolet, B. Salvy, and P. Zimmermann. Lambda-upsilon-omega: The 1989 cookbook. Technical Report 1073, INRIA, 1989.

[6] P. Flajolet and R. Sedgewick. The average case analysis of algorithms: Counting and generating functions. Technical Report 1888, INRIA, 1993.

[7] P. Flajolet and J.S. Vitter. Average-case Analysis of Algorithms and Data Structures. In J. Van Leeuwen, editor, *Handbook of Theoretical Computer Science*, chapter 9. North-Holland, 1990.

[8] D.L. Kreher and D.R. Stinson. *Combinatorial Algorithms: Generation, Enumeration and Search*. CRC Press LLC, 1999.

[9] Greg Kuperberg, Shachar Lovett, and Ron Peled. Probabilistic existence of regular combinatorial structures. *CoRR*, abs/1302.4295, 2013.

[10] J. Liebehenschel. Ranking and unranking of lexicographically ordered words: An average-case analysis. *J. of Automata, Languages and Combinatorics*, 2(4):227–268, 1997.

[11] J. Liebehenschel. Ranking and unranking of a generalized dyck language and the application to the generation of random trees. In *The Fifth International Seminar on the Mathematical Analysis of Algorithms*, Bellaterra (Spain), 1999.

[12] A. Lorenz and Y. Ponty. Non-redundant random generation algorithms for weighted context-free languages. *Theoretical Computer Science, Elsevier, 2013, Generation of Combinatorial Structures*, 502:177–194, 2013.

[13] C. Martínez and X. Molinero. A generic approach for the unranking of labeled combinatorial classes. *Random Structures & Algorithms*, 19(3-4):472–497, 2001.

[14] C. Martínez and X. Molinero. Efficient iteration in admissible combinatorial classes. *Theoretical Computer Science*, 346(2–3):388–417, November 2005.

[15] M. Mitchell. *An Introduction to Genetic Algorithms (Complex Adaptive Systems)*. The MIT Press.

[16] X. Molinero. *Ordered Generation of Classes of Combinatorial Structures*. PhD thesis, Universitat Politècnica de Catalunya, November 2005.

[17] X. Molinero, F. Riquelme, and M. J. Serna. Cooperation through social influence. *European Journal of Operation Research*, 242(3):960–974, May 2015.

[18] MUPAD: The computer algebra system. URL: http://es.mathworks.com/discovery/mupad.html.

[19] A. Nijenhuis and H.S. Wilf. *Combinatorial Algorithms: For Computers and Calculators*. Academic Press, Inc., 1978.

[20] J.M. Pallo. Enumerating, ranking and unranking binary trees. *The Computer Journal*, 29(2):171–175, 1986.

[21] E.M. Reingold, J. Nievergelt, and N.Deo. *Combinatorial Algorithms: Theory and Practice*. Prentice-Hall, Englewood Cliffs, NJ, 1977.

[22] R. Sedgewick and P. Flajolet. *An Introduction to the Analysis of Algorithms*. Addison-Wesley, Reading, MA, 1996.

[23] R. Keller W. Banzhaf, P. Nordin and F. Francone. *Genetic Programming An Introduction*. San Francisco, CA: Morgan Kaufmann, 1998.

[24] Y. Wei. The grouping combinaton generating algorithm. In *Proceedings of the International Conference on Computer, Network Security and Communication Engineering (CNSCE 2014)*, pages 670–674, 2014.

[25] H.S. Wilf. East side, west side ... an introduction to combinatorial families- with MAPLE programming. Technical report, 1999. URL: http://www.cis.upenn.edu/~wilf/lecnotes.html.

Defect Detection Research of Laser Ultrasonic Based on the Improved BP Network

Hongjia Chen, Hui Liu, Xiaoyan Wang, Yanping Bai

***Abstract*—**In the laser ultrasonic surface wave defect detection experiments, we got reflected ultrasonic and the transmitted wave signals in different width and depth cases. According to non-linear, non-stationary characteristics of ultrasonic detection signal, we can use the Mel Frequency Cepstral Coefficient method to extract characteristic coefficients of two waveform signals, and turn the high-dimensional data into low-dimensional signal data, to achieve effective detection of different defect categories. We compare the training and classification of BP neural network to the improved BP neural network with additional momentum. For the test set of reflected wave, the classification accuracy rate of BP network is 84%, and the improved BP network arrives 96%; then for the test set of transmittance wave, correct classification rate of BP network is 82.67%, and the improved BP network is 85.33%. The correct recognition rates of two types of network test are more than 80%, and the improved BP network is much more precise. Then we respectively select a simple in five signals, using the trained networks to identify them. The result is that all correct rates arrive 100%.

***Keywords*—**additional momentum, BP neural network, defect detection, MFCC, SAW.

I. Introduction

LASER ultrasonic technique is an important non-destructive testing technology [1], which stimulate SAW due to the high sensitivity of surface and sub-surface tiny cracks, very suitable for the detection of tiny cracks. In industrial tests, the feature extraction of ultrasonic echo identify defect primarily extracts temporal characteristics, frequency domain characteristics and time - frequency domain features, by extracting the defect features information of relevantfield to achieve the purpose of identifying defects. There are many feature extraction methods, , in general, the method based on time-domain feature extraction is time-series model (AR model, ARMA model, etc.); the method based on frequency domain feature extraction is Fast Fourier Transform (FFT); the methods based on time – frequency domain feature extraction are short-time Fourier transform (STFT), time-frequency distribution (Wigner-Ville distribution, Choi-William distribution), wavelet transform, Hilbert-Huang Transform.

Hongjia Chen is with North University of China, School of Science (e-mail:c1224hj@126.com).

Hui Liu is with North University of China, School of Information and Communication Engineering (e-mail:1159894600@qq.com).

Xiaoyan Wang is with North University of China, School of Science (e-mail:xywang00@126.com).

Yanping Bai is with North University of China, School of Science (corresponding author to provide phone: 0351-3942729; e-mail:baiyp666@163.com).

Because of ultrasonic signals of defectsrecognitionextraction, selection of evaluation methods and theeigenvalues of the law is still in the exploratory stage, with uncertainty. Mel frequency is based on the human auditory characteristic features put forward, and a nonlinear corresponding relationship with it in Hz frequency. Mel Frequency Cepstral Coefficient (MFCC) [2] uses the relationship between them to calculate spectral characteristics of Hz. We will use MFCC to extract feature information of the defect, to achieve the data dimension reduction and easily deal with.

Artificial neural network [3], [4] is an intelligent information processing system built to mimic the human brain, and has a highly nonlinear global mapping. It has a very strong adaptive, self-learning ability and high fault tolerance and robustness on the environment. Then we extract laser ultrasonic wave signal features, with BP neural network, using the network of its powerful massively parallel, distributed processing, self-organizing, self-learning ability, and improve BP network, study ultrasound wave signals, looking for its inherent characteristics to detect different types of defects, and get a good defect category prediction.

II. TheExperimentalPrincipleAndTheUltrasonic FlawSignalsToBeIdentified

Laser ultrasonic flaw detection technique is based on the theory of sound and light effects. It generates thermal elastic effect when pulsed laser irradiates at the sample surface, and generates an ultrasonic signal containing information of the measured surface. By detecting the defects of the ultrasonic signal after modulate to extract information about the defect for defect detection. If the specimen has discontinuous region such as defects, it will occur scattering, reflection and transmission phenomena when the ultrasonic wave propagates to the region, and then leading to the characteristics of ultrasonic signal will change significantly.

In our experiment, the sample to be tested is a 270*70*40mm aluminum plate, we use 2M ultrasound probe to detect ultrasonic reflected wave and transmitted wave, and the bandwidth of probe is 2M, i.e., the detection frequency range of 1-3M, and the defect specifications respectively are as follows: (1) width 0.1mm depth 0.3mm; (2) width 0.1mm depth 0.5mm; (3) width 0.1mm depth 0.7mm; (4) wide and 0.1mm deep 0.9mm; (5) non-destructive. We repeat each experiment measurements for five times, and the sampling rate is 200MHz, the sampling points are 44,000, trigger position 10%. We put probe at 20mm from the laser spot to detect reflected waves, and transmitted wave at 40mm. We show the schematic diagrams for the experiment of reflected

wave and transmitted wave measurement respectively in Figure 1 and Figure 2. And the five groups of signal of two kinds of waves in our experiments are as Figure 3.

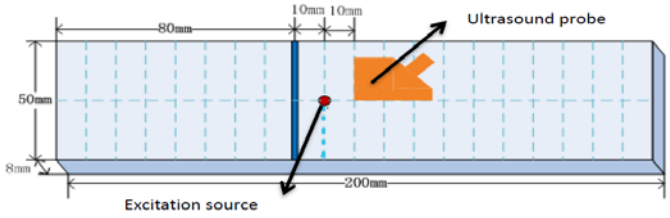


Fig. 1The schematic diagram for the experiment of reflected wave measurement

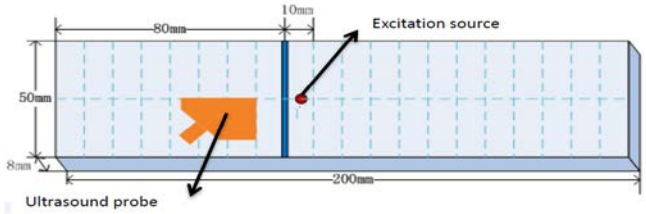
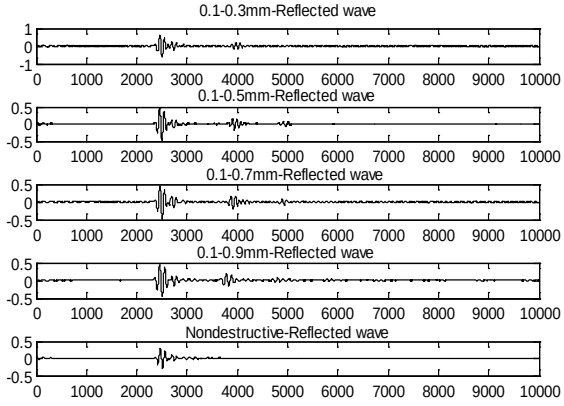
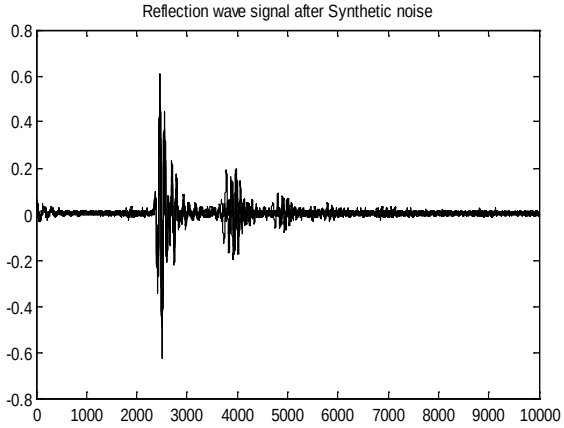


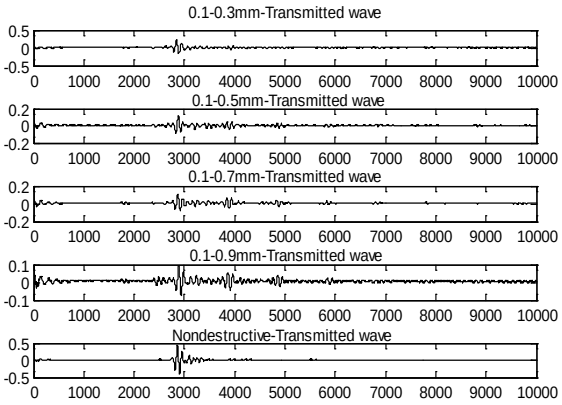
Fig.2 The schematic diagram for the experiment of transmitted wave measurement



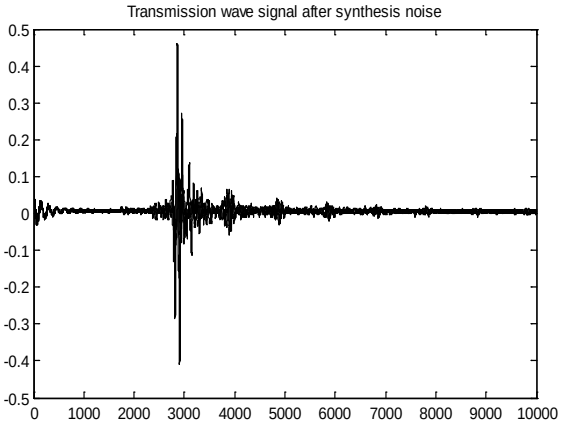
(a)



(b)



(c)



(d)

Fig. 3 Five groups of signal of two kinds of waves (a) Five types of signals of reflected wave. (b) Synthetic reflected wave signals of reflected wave. (c) Five types of signals of transmitted wave. (d) Synthetic reflected wave signals of transmitted wave.

III. MFCCPRINCIPLE

Cochlea is substantially equivalent to a filter set, the filtering effect of the cochlea is used on a logarithmic frequency scale .Below 1,000HZ, there is a linear relationship between the human ear's perception and the frequency; In more than 1,000HZ, the human ear's perception does not constitute a linear relationship with frequency, but more inclined to logarithmic relationship, which makes the human ear to low frequency signal is more sensitive than the high-frequency signal. Mel frequency and frequency conversion formula is:

$$F_{mel} = 2595 * \lg(1 + f_{HZ} / 700) \quad (1)$$

A. Pre-emphasis Processing

Pre-emphasis is actually a high-pass filter, the transfer function of the high pass filter is:

$$H(Z) = 1 - \alpha Z^{-1} \quad (2)$$

ISBN: 978-1-61804-292-7

Where the value of α is 0.97, the role of high-pass filter is filtering low frequency, high frequency characteristics of the ultrasonic signal is more emergent.

B. Frame and Window Treatments

Due to the ultrasonic signal exhibits only present stability in a relatively short period of time (generally considered 10-30ms), thus the ultrasonic signal is divided into a short period that a frame. To avoid loss of dynamic information of ultrasonic signals, there must be some overlapping area between adjacent frames, and the frame length of the overlapping area is a period of 1/2 or 1/3. And then multiply each frame by a window function, in order to increase the continuity of the left and right ends of each frame [5].

C. Each Frame Signal FFT Transformation

We make the FFT transform of Sub-frame for each windowed frame signal to get the spectrum of each frame. And get the square norm of the frequency spectrum of ultrasonic signal to get the power of the ultrasonic signals.

D. Calculate Triangular Filter Coefficients

Defines a plurality of band-pass filter $(k), 0 <= m <= M$, M is the number of triangular filter whose center frequency is $f(m)$, the frequency response for each band pass triangular filter is:

$$H_m(k) = 0 \quad k < f(m-1) \tag{3}$$

$$\frac{k - f(m-1)}{f(m) - f(m-1)} \quad f(m-1) \le k \le f(m) \tag{4}$$

$$\frac{f(m+1) - k}{f(m+1) - f(m)} \quad f(m) \le k \le f(m+1) \tag{5}$$

$$0 \quad k > f(m+1) \tag{6}$$

And satisfies

$$Mel(f(m)) - Mel(f(m-1)) = Mel(f(m+1)) - Mel(f(m))$$

Calculated filter coefficients is $m(i), i = 1, ...p,$ p is the filter order.

E. Triangular Filtering and Discrete Cosine Transform DCT

$$C_i = \sum_{k=1}^{p} \log(m_k) \cos[l(k - \frac{1}{2})\frac{\pi}{p}] \tag{7}$$

C_i is the desired extracted feature parameters.

MFCC has been widely used in speech recognition. Its extraction process is as Figure 4:

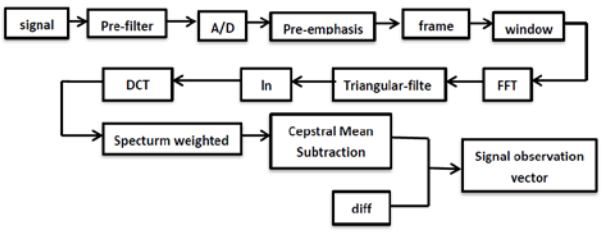


Fig.4The processing of MFCC extracts signal characteristics

IV. THEPRINCIPLEOFBPNETWORK

A. BP Network Structure

Back-propagation (BP) neural network idea was first proposed in 1969 by Bryson etc., it was not until 1986 that Rumelhart [5, 6] and his team published their findings in the journal Nature, the BP network to get the attention of people. BP network is actually a multi-layer perception and a supervised learning algorithm, The network consists of a large number of processing units constructed through an extensive interconnected network system, with massively parallel, distributed processing, self-organizing, self-learning, etc. advantage, is widely used in function approximation, pattern recognition, classification, data compression, and many other fields.

BP neural network [7] is a multilayer feedforward neural network, its main features isto transmit before the signal, the error back-propagation. In the forward pass, an input signal from the input layer through the hidden layer processing layer by layer, until the output layer. The states of neurons of each layer only affect the next layer neuron state. If the output layer is not expected, then transferred back propagation, adjust the network weights and thresholds based on prediction error, allowing BP neural network to predict the output constantly approaching the desired. And BP neural network topology is shown in Figure 5.

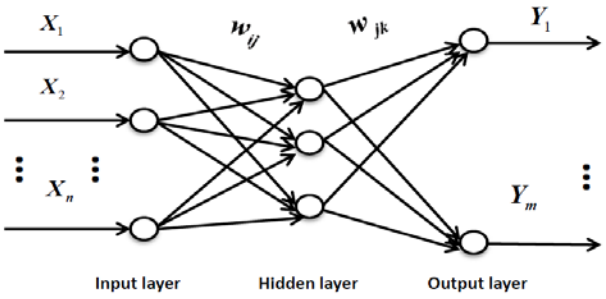


Fig.5 BP neural network topology

Where $X_1, X_2, \cdots, X_n$ are the input values of BP neural network, $Y_1, Y_2, \cdots, Y_m$ are the predicted values of BP neural network, w_{ij} and w_{jk} is the weights of BP neural network. As we can be seen from the figure, BP neural network can be viewed as a nonlinear function, network input is independent variable of the function, and predicted values is the dependent variable of the function. When the input nodes are n, the

output nodes is m, BP neural network expresses a function mapping relationship of n independent variables with m dependent variables.BP algorithm's learning rule is based on the gradient descent method, due to the gradient descent method is an effective nonlinear data fitting method, and it's the direct and effective method needs to calculate the derivative of unconstrained optimization algorithm, this is good maneuverability and overall convergence. Here, the gradient is a vector that we call it derivative.

B. The Working Principle of BP Network

Before BP neural network forecasting, first to train network, through training the network has associative memory and forecasting capabilities.The algorithm's basic idea, that is turning the input and output problems of a set of sample into a nonlinear optimization problem, using gradient descent method that commonly used in optimization to realize mean square error (mse) minimum of the network between actual output and desired output, and finishing BP network training task.Training process of BP network is consist of forward transmission of work signals and back propagation of error signals:

1) Forward transmission of work signals. In the process of the spreading, the input samples income from input layer, finally to output layer afterhidden layer processing step by step. We compare the actual output value to the desired output value of output layer, if there is a deviation, immediately go to the back propagation of error signals process;
2) Back propagation of error signals. The deviation between network's actual output and desired output is the error signal.This process is that the output error propagates along the original connection path step by step, and according to the way of minimizing error to adjust the weight matrix of the network.We adjust the network weights constantly and make the real output value of the network gradually approaching the designer's expectations.

Each layer weights adjustment process of forward transmission of work signals and back propagation of error signals is iterative, and the constantly adjust process of network weights is the training process of the network.

BP neural network training process includes the following steps.

1) Network initialization.According the system input and output sequence (X,Y) to determine the network input layer nodes n, hidden layer nodes l, output layer nodes m. Initialize the connection weights of neurons w_{ij}, w_{jk} between input layer and hidden layer and output layer, initialize the hidden layer threshold a, the output layer threshold b, given the learning rate and neuronal excitation functions.
2) Calculate the hidden layer output. According to the input vector X, the connection weights w_{ij} between the inputlayer and the hidden layer and the hidden layer threshold a, calculate the output of the hidden layer H.

$$H_j = f(\sum_{i=1}^{n} w_{ij}x_i - a_j) \qquad j = 1,2,...,l \tag{8}$$

Where l is the hidden layer node; f representatives hidden layer excitation function, which has a variety of forms, our paper selected the function is:

$$f(x) = \frac{1}{1+e^{x}} \tag{9}$$

3) Output layer output calculation.According to the output of the hidden layer H, the connection weights w_{jk} and thresholds b, calculate predicted output T of BP neural network.

$$T_k = \sum_{j=1}^{i} H_j w_{jk} - b_k \qquad k = 1,2,...,m \tag{10}$$

4) Error calculation. According to the network predicted output T and the desired output Y, computing network prediction error e.

$$e_k = Y_k - T_k \qquad k = 1,2,...,m \tag{11}$$

5) Update weights. According to the predicted error e update the network weights w_{ij}, w_{jk}.

$$w_{ij} = w_{ij} + \eta H_j(1-H_j)x(i)\sum_{k=1}^{m} w_{jk}e_k \qquad i = 1,2,...,n;\ j = 1,2,...,l \tag{12}$$

$$w_{jk} = w_{jk} + \eta H_j e_k \qquad j = 1,2,...,l;\ k = 1,2,...,m \tag{13}$$

Where η is the learning rate.

6) Threshold update. According to the network predicted error e update the network node threshold a, b.

$$a_j = a_j + \eta H_j(1-H_j)\sum_{k=1}^{m} w_{jk}e_k \qquad j = 1,2,...,l \tag{14}$$

$$b_k = b_k + e_k \qquad k = 1,2,...,m \tag{15}$$

7) Judge iterative algorithm to determine whether the end, if not end, return (2).

C. Improved BP network

At present, in the field of the application of neural network, BP algorithm is the most widely used. Although BP network can approximate any nonlinear function in theory, because there are many parameters in network training learning choice without theoretical basis, in practice, the algorithm itself has some limitations and shortcomings, mainly includes the following aspects:

1) The learning rate of BP network is fixed, its slow convergence speed and long-time training, it often requires thousands of times even more iterative training.

2) The gradient descent method of BP network makes the network easy to fall into local minimum and can't get the global optimal, and when the training patterns learn one by one, the network connection weights will be readjusted, this makes training process time longer.
3) The number of nodes in the hidden layer of network is usually determined by experiences, there is no exact theory instruction, which makes the design of the network model become more complex, the network training time can be longer.

Aiming at the limitations and disadvantages of BP network, the researchers spent a lot of energy in improving the performance of its research work, and put forward many improvements.

In the ultrasonic wave signal recognition and classification, using the steepest descent algorithm of BP network can make weights and threshold vector to get a stable solution, but the learning process is slow convergence, network easily trapped in a local minimum. At the same time due to the BP network is sensitive to learning rate, simply increase the learning rate to accelerate the convergence, the algorithm may be unstable and oscillating. Therefore, to solve these problems is very important. Thus we adopt additional momentum to solve [8, 9].

On the basis of the back-propagation, each weight change pluses a value proportional to the previous weight change value, and in accordance with back propagation method to generate a new weight value conversion. With additional momentum factor weights value adjustment formula is:

$$w_{ij}(k+1) = w_{ij}(k) + \Delta w_{ij}(k+1) + \delta[w_{ij}(k) - w_{ij}(k-1)] \quad (16)$$

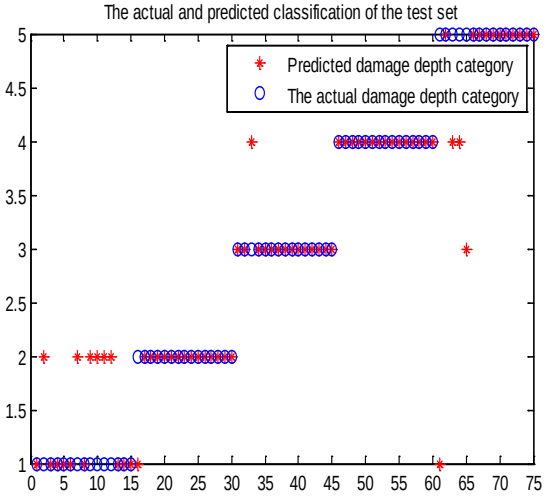
$$b_j(k+1) = b_j(k) + \Delta b_j(k+1) + \delta[b_j(k) - b_j(k-1)] \quad (17)$$

Where δ is the momentum factor, generally take around 0.95.

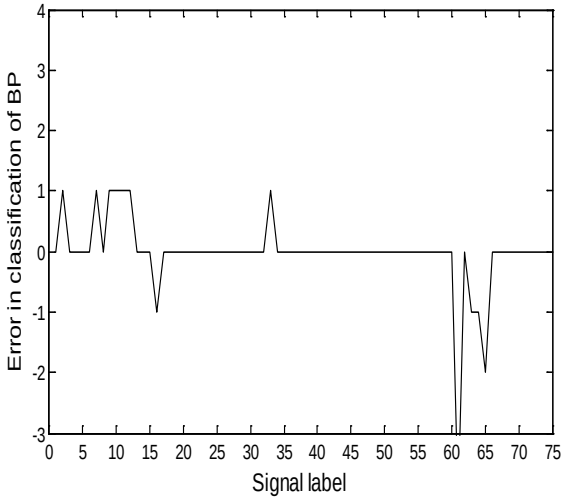
V. Laser Ultrasonic Surface Wave Defect Detection Experiment Contrast And The Analysis Of The Results

A. Laser Ultrasonic Surface Wave Flaw Detection Experiments Contrast

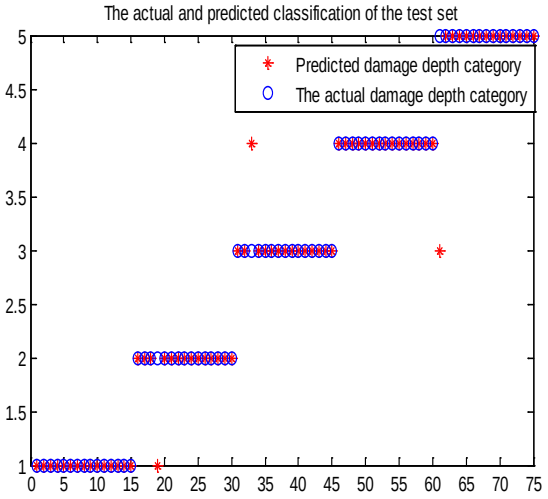
We sample five kinds of defect signals, and get data of 10,000*5. Then we extract feature with Mel Cepstral method. Each defect signal extract 24 features, sample length is 75, then the five kinds of defect signals turn into the sample characteristics 375 * 24.We randomly selected 300 samples as the training set of the network, each type of defect signals selected 15 samples for testing set, that the number of samples tested set was 75. The number of neurons in the input layer is 24, one hidden layer and the number of the hidden layer neurons is 9, number of neurons in the output layer is 5. After the network is fully trained using a test set for testing, BP network and improved BP network results in the Figure 6 and Figure 7.



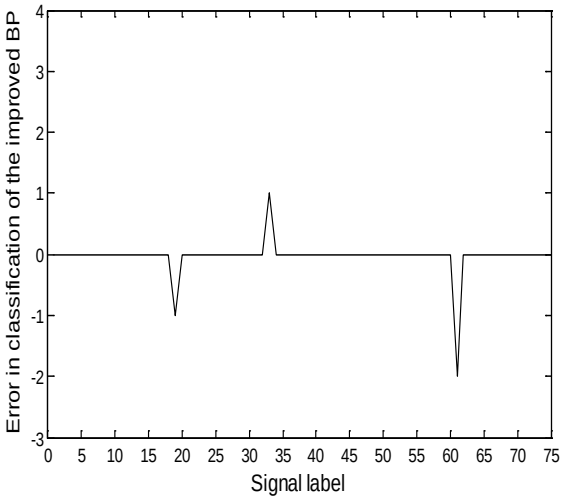
(a)



(b)



(c)



(d)

Fig. 6 Reflected wave experiment contrast. (a) BP network classification results of reflected wave. (b) Error of BP network classification. (c) The improved BP network classification results of reflected wave. (d) Error of the improved BP network classification.

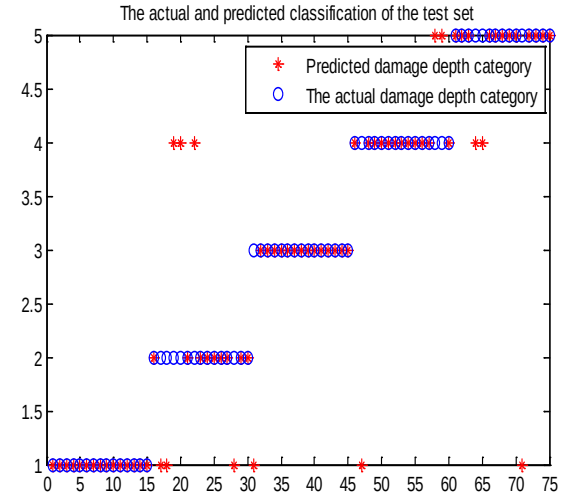
In the Figure 6, (a) represents BP network classification results of reflected wave, and the red '*' is the predicted damage depth category, the blue 'O' is the actual damage depth category; (b) represents error in classification of BP network; (c) represents the improved BP network classification results of reflected wave, and the red '*' is the predicted damage depth category,the blue 'O' is the actual damage depth category; (d) represents error in classification of the improved BP network.

TABLE 1: THE RESULT OF TWO KINDS OF NETWORKS TEST REFLECTED WAVE SIGNAL.

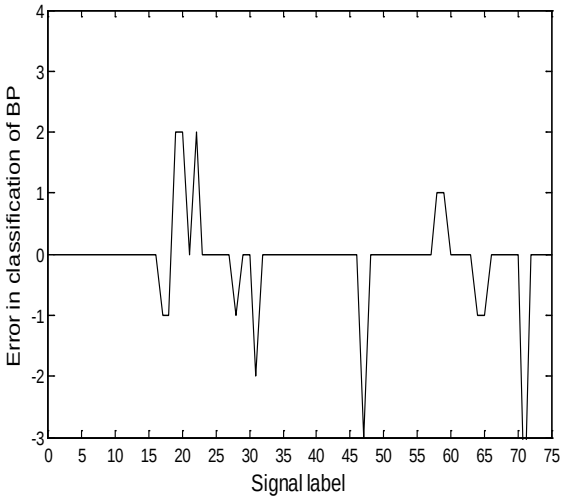
Identification method	Number of training samples	Test samples	Identify the correct number	The correct rate
BP network	300	75	63	0.8400
Improved BP network	300	75	72	0.9600

By the TABLE 1, we can see that for the reflected wave, the number of test set samples is 75, the correct identification of BP network number is 63, the correct identification rate is 84%, while the correct identification of the improved BP neural network number is 72, right recognition rate is 96%.Both methods recognition rate are in excess of 80%.

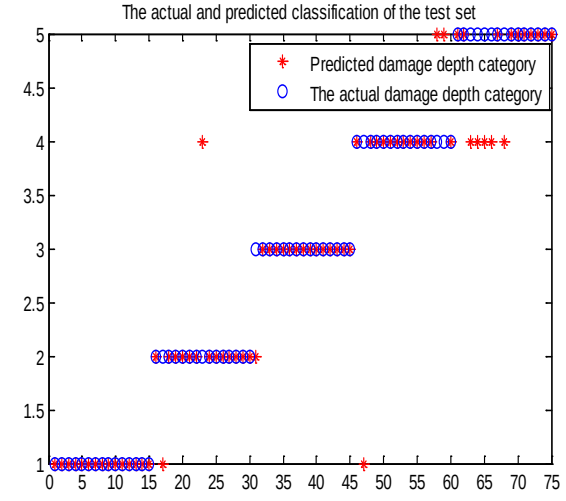
In the Figure 7, (a) represents BP network classification results of transmitted wave, and the red '*' is the predicted damage depth category, the blue 'O' is the actual damage depth category; (b) represents error in classification of BP network; (c) represents the improved BP network classification results of transmitted wave, and the red '*' is the predicted damage depth category, the blue 'O' is the actual damage depth category; (d) represents error in classification of the improved BP network.



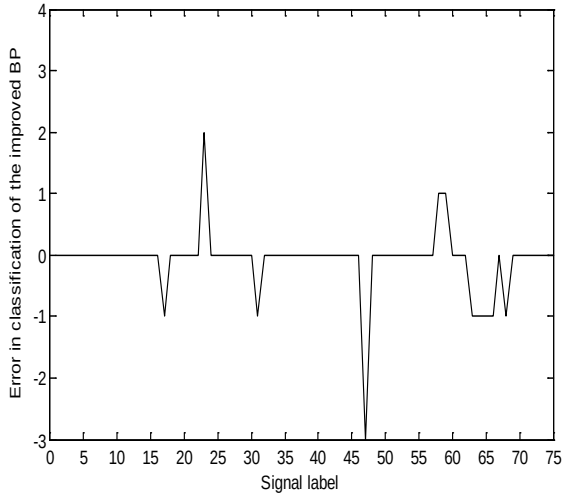
(a)



(b)



(c)



(d)

Fig. 7 Transmitted wave experiment contrast. (a) BP network classification results of transmitted wave. (b) Error of BP network classification. (c) The improved BP network classification results of transmitted wave. (d) Error of the improved BP network classification

TABLE 2: THE RESULT OF TWO KINDS OF NETWORKS TEST TRANSMITTED WAVE SIGNAL.

Identification method	Number of training samples	Test samples	Identify the correct number	The correct rate
BP network	300	75	62	0.8267
Improved BP network	300	75	64	0.8533

As can be seen from the TABLE 2, for the transmitted wave, the number of test set samples is 75, the correct identification of BP network number is 62, the correct identification rate is 82.67%, while the correct identification of the improved BP neural network number is 64, right recognition rate is 85.33%. The correct recognition rate of two networks for the transmission wave is lower than reflected waves, but the recognition rate of both over 80%.

From the above two kinds of networks for reflected wave and transmission wave flaw detection experimental results, it can be seen that the two methods both can be effectively used for defect detection, and the improved BP network performance is better.

B. Analysis of Experimental Results

Now, in order to cancel the order of magnitude difference between the various dimensions of data, we convert sample data into [0,1] by the maximum and minimum normalization method, to avoid big network prediction errors caused by the magnitude of the difference between the input and output data. Maximum and minimum normalized form follows function:

$$x_n = \frac{x_n - x_{\min}}{x_{\max} - x_{\min}} \tag{18}$$

Where $x_{\min}$ represents the minimum value of the input data sequence, $x_{\max}$ is the maximum value of the sequence. According to the desired type of the signal identifies, we encode each category as the goal desired output vector, and the defect types code as shown in TABLE 3.

After the network is fully trained, using the test set for testing, the test results are shown in TABLE 4. As can be seen, the actual output of the network is very consistent with the expected output, and the diagnostic accuracy is 100%. This proofs that the neural network model is reliable and can accurately defect types of ultrasonic signal for effective identification and classification.

VI. CONCLUSION

We extract the characteristic coefficients of ultrasonic reflected and transmitted wave signals with Mel Cepstral method, and successfully reduce the sample data dimension from 10,000*5 to 24*375. Training and identifying the data with neural network, our experiments show, BP and improved BP network can achieve effective detection of different types of defects. However, in general, the improved BP network is much more precise.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests regarding the publication of this paper.

ACKNOWLEDGEMENTS

This work was financially supported by the National Natural Science Foundation ofChina(61275120).

TABLE 3: DEFECT CATEGORY CODE DESIGN.

Parameters	Fault type（width(mm)*depth(mm)）				
	0.1*0.3	0.1*0.5	0.1*0.7	0.1*0.9	Non-destructive
Defect code	1 0 0 0 0	0 1 0 0 0	0 0 1 0 0	0 0 0 1 0	0 0 0 0 1

TABLE 4: SAMPLE TEST RESULTS.

Experiment Type	Defect types	Target	Prediction	Correct classify-cation rate
BP Reflected wave	0.1*0.3	1 0 0 0 0	0.9801 0.00500.0129 0.0020 4.141e-07	100%
	0.1*0.5	0 1 0 0 0	0.01010.8871 0.0573 3.336e-04 0.0451	
	0.1*0.7	0 0 1 0 0	0.1444 0.2436 0.4685 0.0014 0.1421	
	0.1*0.9	0 0 0 1 0	0.0283 0.0318 0.0554 0.8603 0.0242	
	Non-destructive	0 0 0 0 1	0.0931 0.0767 0.0082 0.0191 0.8029	
Improved BP Reflected wave	0.1*0.3	1 0 0 0 0	0.8025 0.0131 0.0699 0.0728 0.0417	100%
	0.1*0.5	0 1 0 0 0	0.2539 0.7109 0.0026 0.0269 0.0057	
	0.1*0.7	0 0 1 0 0	0.0244 0.0401 0.8348 0.0881 0.0126	
	0.1*0.9	0 0 0 1 0	0.0571 0.1036 0.0162 0.7494 0.0737	
	Non-destructive	0 0 0 0 1	0.0115 0.1114 0.1263 0.0709 0.6799	
BP Transmitted wave	0.1*0.3	1 0 0 0 0	0.9801 0.0050 0.0129 0.0020 4.141e-07	100%
	0.1*0.5	0 1 0 0 0	0.0023 0.9915 0.0016 0.0015 0.0031	
	0.1*0.7	0 0 1 0 0	0.0074 0.0121 0.8427 0.1379 2.941e-10	
	0.1*0.9	0 0 0 1 0	5.999e-14 0.0002 3.406e-09 0.9885 0.0114	
	Non- destructive	0 0 0 0 1	2.419e-08 0.0003 4.073e-12 0.0001 0.9995	
Improved BP Transmitted wave	0.1*0.3	1 0 0 0 0	0.4033 0.0478 0.2090 0.0598 0.2801	100%
	0.1*0.5	0 1 0 0 0	0.17340.4095 0.1947 0.1759 0.0465	
	0.1*0.7	0 0 1 0 0	0.0263 0.1533 0.44290.12090.2566	
	0.1*0.9	0 0 0 1 0	9.490e-04 0.0306 0.0961 0.5959 0.2764	
	Non- destructive	0 0 0 0 1	0.1714 0.00476.074e-04 0.0974 0.8973	

REFERENCES

[1] Gang Li. Laser ultrasonic technology and its application in non-destructive testing of metal [D]. Central China Normal University, 2004.

[2]Yichuan Wang, Zhizhong Li. Ship target classification based on Mel Cepstral and BP neural network[J]. Sensors and Microsystems, 2011,06: 55-57 + 67.

[3] Jiaohong Yi, Weihong Xu, Yuantao Chen. Novel Back Propagation Optimization by Cuckoo Search Algorithm [J].The Scientific World Journal, 2014.

[4]Jianyong Liu, Huaixiao Wang, Yangyang Sun, Chengqun Fu, JieGuo. Real-Coded Quantum-Inspired Genetic Algorithm-Based BP Neural Network Algorithm [J]. Mathematical Problems in Engineering , 2015

[5]Xiaochuan Wang, Feng Shi, Yang Li. The analysis of 43 cases in MATLAB Neural Network[M]. Beijing: Beijing University of Aeronautics and Astronautics Press,2013.8.

[6] Feng Cui, Fenghe Qi, HuiyunLiu.A BP neural network sound characteristic signal identification method [J].Daqing Normal University, 2012,06: 23-26.

[7] Chaoqin Peng, Chun Cao, Jiaoying Huang, Qiusheng Liu. Seismic signal recognition using improved BP neural network and combined feature extraction method[J]. Journal of Central South University,2014,05:1898-1906.

[8]ZhiguoMeng. The analysis of BP network's application in land using classification[D].Jilin University, 2004.

[9]Junqin He, Ting Jiang,Zhihao Xing.A method of target detection and identification based on RPROP and UWB channel characteristic parameters, *Globecom Workshops (GC Wkshps), 2012 IEEE*, On page(s): 1460 – 1463.

Optimizing complex problems solving with a memory based isomorphism

Alberto Arteta, Juan Castellanos and Luis Fernando Mingo

Abstract— During years proper memory utilization has been the differential factor for algorithms that try to solve complex problems in optimal time. Computational complexity is currently measured in time and space. The way that algorithms are built can make a huge difference when obtaining desired solutions to problems. Memory engineering based algorithms reveal themselves as essential when fast results are needed and offer unlimited options to improve whatever bioinspired algorithm that based its performance in terms of time and space. This work focuses on creating a mathematical generic strategy by building structures through the use of an isomorphism that optimize the memory utilization and the resolution time of bioinspired models when dealing with high computational problems. Therefore offers a great help when solving complex and known issues.

Keywords— Bioinspired models, General Optimization , Complex problems resolution

I. Introduction

This section presents a software technique for improve the functionality of a bioinspired system [1] [2] by optimizing the memory resources. During years new algorithms are trying to get better performance [3][4] when solving complex problems such as the knapsack problem [5] or the travel salesman problem [6]. Some techniques try to establish general procedures for optimizing the performance in bioinspired systems such as the ones in [7] and [8]. Others techniques focus in the rules election phase that takes place in membrane models [9]. This work introduces and compares several auxiliary structures [10] and establishes a way to ensure the best possible performance of a given bio-inspired system.

By using linear structures and allocating them in the physical memory, response timing gets reduced considerably. In scenarios where traditional algorithms do not obtain good performance, the technique explained in this paper guarantees an excellent performance.

This paper is structured as follows: New Definitions for building the structures in memory [10]. These are:

- Concept of patterns:
- Two sets of functions: The composition of the functions of each set relates application of evolution rules to set of patterns.
- Creation of a main n-dimensional structure in RAM and a virtual auxiliary n-dimensional structure. That main structure is an application that establishes a link between the initial multisets, and the number of times that each evolution rule should be applied in order to obtain an extinguished multiset [9].

Once the structures are created, they will be proven to be consistent.

A Comparative study between the methods that optimize the memory resources. Advantages and disadvantages of using this technique will be explained. A proposal describing the best scenario to be implemented in will be provided too. The comparison shows that a memory strategy is essential regardless is virtual [11] [12] or physical [8][13].

II. Definitions

Patterns $N \otimes N \rightarrow A \qquad A \subseteq P(N)$

$[i, j] = \{ k \in N \mid i \leq k \leq j \quad i, j \in N \}$

Set of patterns

$S \subseteq P(N)$ is a *Set of patterns is defined as the set:*

$\{ [\![a_i, b_j]\!] \mid \exists n, m \in N, i \leq n \; j \leq m \; a_i, b_j, i, j \in N \}$

Set of set of patterns

$SS = \{ S_i \; i \in N \; S_i \; is \; a \; set \; of \; patterns \mid \exists n \in N \; i \leq n \}$.

Observation

Given a region R and alphabet of objects U, and R (U, T) set of evolution rules over U and targets in T re`presented as follows:

$r_1 : a_1^{u_{11}} a_2^{u_{12}} . a_n^{u_{1n}} \rightarrow C_1$

$r_2 : a_1^{u_{21}} a_2^{u_{22}} . a_n^{u_{2n}} \rightarrow C_2$

$\ldots \qquad \rightarrow \ldots$

$r_m : a_1^{u_{m1}} a_2^{u_{m2}} . a_n^{u_{mn}} \rightarrow C_m$

There is always a set of set of patterns $SS_{R(UT)}$ associated to it. This set of set of patterns contains all the possible extinguished multisets and it is obtained by expanding the formula included in the definition of extinguished multiset:

ISBN: 978-1-61804-292-7

$$\bigcap_{l=1}^{m}\left[\bigcup_{i=1}^{n}\left(u_i - \sum_{j=1}^{m}(k_j \cdot u_{ji}) \le u_{li}\right)\right]$$

$$\begin{pmatrix}
[[0,u_{11}],[0,u_{12}]...[0,u_{1n}]],[[0,u_{11}],[0,u_{12}]...[0,u_{2n}]],..,[[0,u_{11}],[0,u_{12}]...[0,u_{mn}]],\\
...\\
[[0,u_{11}],[0,u_{22}]...[0,u_{1n}]],[[0,u_{11}],[0,u_{22}]...[0,u_{2n}]],..,[[0,u_{11}],[0,u_{22}]...[0,u_{mn}]],\\
...\\
[[0,u_{11}],[0,u_{m2}]...[0,u_{1n}]],[[0,u_{11}],[0,u_{m2}]...[0,u_{2n}]],..,[[0,u_{11}],[0,u_{n2}]...[0,u_{mn}]],\\
...\\
[[0,u_{21}],[0,u_{12}]...[0,u_{1n}]],[[0,u_{21}],[0,u_{12}]...[0,u_{2n}]],..,[[0,u_{21}],[0,u_{12}]...[0,u_{mn}]],\\
...\\
[[0,u_{21}],[0,u_{22}]...[0,u_{1n}]],[[0,u_{21}],[0,u_{22}]...[0,u_{2n}]],..,[[0,u_{21}],[0,u_{22}]...[0,u_{mn}]],\\
...\\
[[0,u_{21}],[0,u_{m2}]...[0,u_{1n}]],[[0,u_{21}],[0,u_{m2}]...[0,u_{2n}]],..,[[0,u_{21}],[0,u_{n2}]...[0,u_{mn}]],\\
...\\
[[0,u_{m1}],[0,u_{12}]...[0,u_{1n}]],[[0,u_{1n}],[0,u_{12}]...[0,u_{2n}]],..,[[0,u_{m1}],[0,u_{21}]...[0,u_{mn}]],\\
...\\
[[0,u_{m1}],[0,u_{22}]...[0,u_{1n}]],[[0,u_{1n}],[0,u_{22}]...[0,u_{2n}]],..,[[0,u_{m1}],[0,u_{22}]...[0,u_{mn}]],\\
...\\
[[0,u_{m1}],[0,u_{m2}]...[0,u_{1n}]],[[0,u_{1n}],[0,u_{m2}]...[0,u_{2n}]],..,[[0,u_{m1}],[0,u_{2n}]...[0,u_{mn}]],
\end{pmatrix}_{Lin} \text{ is}$$

###

Linear Multisets isomorphism Φ_1

Building Φ_1, as a part of the physical evolution rule isomorphism is created as a function which has an m-dimensional set of natural numbers as input and an n-dimensional set of natural numbers as output, where m is the number of evolution rules and n is the number of symbols included in a given multiset of objects.

Definition: Linear Multisets function φ_1

Let U= $\{a_i \mid i = 1,..n\}$ be a set of objects. Let T beset of targets. Let $\omega = a_1 a_{2..} a_n$ be a multiset of objects and let x_i be the multiplicity of a_i. Let R (U, T) be a multiset of evolution rules with objects in U and targets in T. Let m be the number of evolution rules within R (U, T). Let k_i be the number of times that the rule $r_i \in R(U,T)$ is applied. Then φ_1 is defined as:

$$\varphi_1 : N^m \rightarrow N^n$$
$$(k_1, k_2, .., k_m) \rightarrow (x_1, x_2, .., x_n)$$

$$\varphi_1(k_1,k_2,..k_m) \equiv \begin{pmatrix} u_{11} & u_{21} & ..u_{m1} \\ u_{12} & u_{22} & ..u_{m2} \\ ... & ... & ... \\ u_{1n} & u_{2n} & ..u_{mn} \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ ... \\ k_m \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{pmatrix}$$

Once the function φ_1 is created, the isomorphism Φ_1 is defined as follows:

Definition: Linear Multisets isomorphism Φ_1

Let φ_1 be the multiset linear function related to a given evolution rules set $\Phi_1 : N^n \rightarrow N^n$

$$\Phi_1(x) = \begin{cases} \varphi_1(x) & \neg\exists y \neq x \mid \varphi_1(y) = \varphi_1(x) \\ random\{\varphi_1(x), \varphi_1(y)\} & \exists y \neq x \mid \varphi_1(y) = \varphi_1(x) \end{cases}$$

Definition

Physical evolution rules linear Isomorphism is then defined as follows:

$$\Phi = \Phi_1 \circ \Phi_2 : N^m \rightarrow P(N^n)$$
$$(k_1, k_2, ..., k_m) \xrightarrow{\Phi_1} (x_1, x_2, .., x_n) \xrightarrow{\Phi_2} (x_1, x_2, .., x_n) - SS_{R(U,T)}$$

Based on the previous definitions, the following one is established and created to take full advantage of the virtual memory.

Definition: Virtual linear Multisets function

$$\Phi_{1vir} : N^n \rightarrow P(N^m)$$
$$[x_1, x_2, .., x_n] \rightarrow \begin{bmatrix} k_1, k_2, .., k_m \\ k_1', k_2', .., k_m' \\ ... \\ k_1'^p, k_2'^p, .., k_m'^p \end{bmatrix}$$

Given an input $x = (x_1, x_2, .., x_n) \in N^n$, it returns the set of numbers

$k = (k_1, k_2, .., k_m) \in N^m / \varphi_1(k) = x$

Definition: Virtual linear Pattern function

$$\Phi_{2virt} : P(N^m) \rightarrow N^n$$
$$\begin{pmatrix} [[0,u_{11}],..[0,u_{m1}]],..,[[0,u_{11}],..[0,u_{mn}]], \\ ... \\ [[0,u_{12}],..[0,u_{m1}]],..,[[0,u_{12}],..[0,u_{mn}]], \end{pmatrix} \rightarrow \begin{bmatrix} x_1, x_2, .., x_n \\ x_1', x_2', .., x_n' \\ ... \\ x_1'^p, x_2'^p, .., x_n'^p \end{bmatrix}$$

Given a *set of set of patterns* as the input it returns a set of numbers $x = (x_1, x_2, .., x_n) \in N^n$. The elements of this resulting set are all the combinations of all the possible

$x^j = (x_1, x_2, .., x_n) \in N^n$ where $x_i^j \in pattern\ (i) \in SP\ (j)$ of a set of patterns contained in the matrix of set of patterns.Now it is possible to build the physical and virtual linear structures from the multisets isomorphism.

III. ISOMORPHISM BASED STRUCTURES

The structure is created as follows.

$$L[\Phi(k_1, k_2, .., k_m)] = \begin{cases} (k_1, k_2, ..., k_m) & L[\Phi(k_1, k_2, ..., k_m)] = \phi \\ random\{(k_1, k_2, ..., k_m), no\,action\} & L[\Phi(k_1, k_2, ..., k_m)] \neq \phi \end{cases}$$

When L (i, j) already has a value, then, a random election must be done. This election can be either overwriting the old value with the

new one, or to leave the old value (No action).This linear structure has to comply with having all possible numbers, up to a combination of benchmarks, reasonably high. Each symbol $\{ X_i \mid i = 1,.. n \}$ will have a benchmark. The combination of all the benchmarks will define the number of entries that the linear structure has. Each entry stores the values $\{ k_1,.. k_m \}$. These values indicate the number of times that an evolution rule should be applied to an initial multiset in order to obtain an extinguished multiset.

The linear structure L_Φ must guarantee that it contains all the entries corresponding to the combination number of all benchmarks associated to the symbols $\{ X_i \mid i = 1,.. n \}$, i.e. it can not have non functioning entries. That could damage response timing and it would increase the computational complexity in terms of time. It's necessary to prove that this structure has not null values.

Let L_Φ be a linear structure built from the *evolution rules isomorphism* and a certain *V* multiset of objects and *R (U, T)* multiset of evolution rules. $\rightleftharpoons$

$L_\Phi [X_1, X_2,.., X_n] \in N^m \ \forall (X_1, X_2,.., X_n) \in N^n \ X_i \le X_i'$

$\exists \forall i \le n \ \ benchmark \ \ (X_i) = X_i' \quad |R(U,T)| = m \ and \ \ n = |V|$

Proof

Let $\varphi_1 \equiv \begin{pmatrix} A_{11} & .. A_{m1} \\ .. & .. \\ A_{1n} & .. A_{mn} \end{pmatrix}$ be the *function* associated to *R (U, T)*

$$\begin{pmatrix} A_{11} & .. A_{m1} \\ .. & .. \\ A_{1n} & .. A_{mn} \end{pmatrix} \begin{pmatrix} k_1 \\ .. \\ k_m \end{pmatrix} \Rightarrow$$

$$\begin{array}{lll} A_{11} k_1 + & + .. & A_{m1} k_m \quad (1) \\ .. & \ldots & (i) \\ A_{1n} k_1 + & + .. & A_{mn} k_m \quad (n) \end{array}$$

.

A resulting set of n-sums is obtained. Let A_i be = {min A_{ij} of the sums (i)} .Moreover, let $P(A_{ij})$ be the Matrix of set of patterns resulting from *R (U, T).*

$\forall i \in N \ i \le n, \exists S_{R(U,T)} \ set \ of \ \ patterns / S_{R(U,T)}[i] = [0, A_i]$

. Proof is trivial

Thus,

$\forall A_{ij} \neq A_i \ A_{ij} = 0 \Rightarrow \exists j \in N \ j \le m / A_i \cdot k_j + [0, A_i] = x \ \forall x \in N$

.

Observation:

This proves that when following this method, any natural number from $\{k_1,.., k_m\} \in N^m$ can be generated. Thus, the structure does not have null values from any entry. As the process is an isomorphism, any natural number between 0 and a given benchmark will be uniquely related to a combination of $\{k_1,.., k_m\} \in N^m$ where *m* is the number of evolution rules included in *R (U, T).* Moreover the calculation of *extinguished multisets* will be immediate. When $\{k_1,.., k_m\} \in N^m$ are found, applying the evolution rule r_i a number of k_i times to the initial multiset, will calculate them.

Given *V*= $\{\{ X_i \mid i = 1,.. n\}$ be a *multiset of symbols and given R (U, T),* a multiset of evolution rules. *Given the set* $\{k_i \in N$ *the number of times that the evolution rule* r_i *is applied over the initial multiset}* the following evolution rules function is defined:

$$\Phi_{virt} = (\Phi_{1\,virt} \circ \Phi_{2\,virt}) (P_{R(U,T)}(A_{ij})$$

IV. ALGORITHM

Following is the code that returns the multisets as outputs

(1) $X, Y \leftarrow Multiplicity(R(U,T))$

(2) $BEGIN$

(3) $output(L_\Phi(X,Y))$

(4) END

(5) $L_\Phi(X,Y) \leftarrow L_{\Phi virt}(X,Y)$

The algorithm search in the physical structure the position(X, Y) which are the input values corresponding to the multiplicities of the initial multiset. When the value is returned, the algorithm finishes and a new value coming from the virtual structure overwrite the value stored in the position (X, Y), keeping the non deterministic nature of the system.

V. CONCLUSIONS AND FURTHER WORK

Although it is clear that proper memory utilization is a great help to speed up the information processing, it is still necessary to build the right structures to optimize the performance; and that is not always easy. This paper contributes with a general technique that consists of building an isomorphism based structure that improves the performance of traditional algorithms when dealing with complex problems. The isomorphism matches the initial data sets multiplicities with the number of times that each rule should be applied in order to obtain solutions in optimal time; In that way, this method clearly reduces the execution times of the algorithms when finding maximal or extinguished multi data sets.

The nature of the isomorphism ensures that the main bioinspired system features are preserved and the overall system's functionality is not modified, the only difference between the first model and the second one is the performance. Although the idea of using memory resources to increase performance when solving complex problems is not new, defining the right structure based on isomorphism can offer a promising way to generalize and to create a standard methods for optimizing the memory resources of traditional algorithms. Improving the design and the implementation of new isomorphisms reveal themselves as keys factors in the strategy to deal with complex problems by utilizing memory engineering.

ACKNOWLEDGMENT

The authors would like to thank the Natural Computing group of Madrid for providing the entire infrastructure to perform the tests.

REFERENCES

[1] Gh. Păun, "Computing with Membranes", Journal of Computer and System Sciences, 61(2000), and Turku Center of Computer Science-TUCS Report nº 208, 1998.

[2] Gh. Păun, "Membrane computing. Basic ideas results, applications", Pre-Proceedings of First International Workshop on Theory and Application of P Systems, Timisoara (Romania), pp. 1-8, September , 2005.

[3] Kenneth Price, Rainer M. Storn, and Jouni A. Lampinen. Differential Evolution: A Practical Approach to Global Optimization (Natural Computing Series). Springer- Verlag New York, Inc., 2005.ISBN 3540209506..

[4] K.V. Price. Differential evolution: a fast and simple numerical optimizer. In Fuzzy Information Processing Society, 1996. NAFIPS. 1996 Biennial Conference of the North American, pages 524– 527, 1996. doi: {10.1109/NAFIPS.1996.534790}.

[5] Lingian Pan, Carlos Martin-Vide "Solving multidimensional 0-1 knapsack problem by P systems with input and active membranes", Journal of Parallel and Distributed Computing Volume 65 , Issue 12 (December 2005)

[6] Lin, Shen, Kernigham BW An Effective Heuristic Algorithm for the Traveling-Salesman Problem". Operations Research **21** (2): 498–516. doi:10.1287/opre.21.2.498. Volume 21, Issue 2, March- April 1973,

[7] A. Arteta, L.Fernandez, J.Gil "Algorithm for Application of Evolution Rules based on linear diophantine equations" Synasc 2008.Timisoara Romania September 2008

[8] Arroyo, A. Arteta,, A. Goñi. Calculating maximal multisets using RAM as support, Artificial life and Robotics 2010, Beppu Japan

[9] Alberto Arteta, Nuria Gomez Luis Fernando Mingo. ,Solving complex problems with a bioinspired model. Engineering Applications of Artificial Intelligence,Volume 24, Issue 6, September 2011, Pages 919–927

[10] Alberto Arteta ,, Angel Castellanos, Ana Martinez: Membrane computing: non deterministic technique to calculate extinguished multisets of objects. International Journal " Information Technologie s and Knowledge", Vol. 4, Number 1, 2010

[11] Chiang, Jui-Hao. Optimization Techniques for Memory Virtualization-based Resource Management Publisher: The Graduate School, Stony Brook University: Stony Brook, NY. Date: 1-Dec-12

[12] Ganon, Jalby Strategies for cache and local memory management by global program transformation Journal of Parallel and Distributed Computing, Volume 5, Issue 5, October 1988, Pages 587–616

[13] L. Fern´andez, J.Castellanos, F. Arroyo, J. tejedor, I Garc´ıa "New algorithm for application of evolution rules", Proceedings of the 2006 International Conference on Bioinformatics and Computational Biology, BIOCOMP'06, Las Vegas, Nevada, USA, 2006.

Prof. Alberto Arteta Dr. Arteta is currently an Associate professor of the technical University of Madrid since 2007. He works in the Department 'Applied mathematics'. His main area of interest includes the study of mathematical models and their applications to Biology and Medical Science, and IT. He is also editor and reviewer of several prestigious international Journals.

Prof. Luis Fernando de Mingo López
Dr. Mingo is currently an Associate professor of the technical University of Madrid since 2005. He works in the Department 'Organization of data structures'. His main area of interest includes the study of patterns detections tfor data mining. He holds his Ph.D in this field and has published more then 30 papers. He is also a editor of the "International scientific society journal"

Prof. Juan Castellanos
Prof.Juan Castellanos is an Associate professor of the technical University of Madrid since 2005. He works in the Department 'Artificial Intelligent'. His main area of interest includes the study of new computational models such us neural networks and Particle Swarm Optimization, and the innovation of educational method Optimization. He is the head of the Natural Computing research group.

On the partition of vertex's neighborhood in a graph

Hayat Issaadi, Hacene Ait Haddadene, and Safia Zenia‡

Abstract

A (k,l) partition graph (also called (k,l) graph) G is a partition of its vertices into k stable sets and l cliques. A vertex v in a graph G is called a (k,l) split vertex if its neighborhood in G admits a (k,l) partition. We say that G is (k,l) split neighborhood graph (denoted $SN(k,l)$ graph) if every induced subgraph of G contains a (k,l) split vertex. In this paper, we investigate perfect and non perfect (k,l) split neighborhood graph for some values of k and l.

(k,l) Split Neighborhood Graph, Recognizing algorithm, Maximum clique, Perfect graph, Optimal coloring.

1 Introduction

The problem of partitioning the vertex set of a graph has been studied by many researchers [9] [13] [19] , in particular when the vertex set is partitionned into stable sets and cliques. A (k,l) partition graphs were first studied by Brandstädt [7]. He reported that this class generalizes bipartite and split graphs and remarked that deciding whether a graph is a (k,l) graph is NP-complete whenever $k \geq 3$ or $l \geq 3$. He also showed that finding a $(1,2)$ or $(2,1)$ partition (respectively $(2,2)$ partition) in a graph can be performed in $O\left(n^2m\right)$ (respectively $O\left(n^{10}m\right)$) [6] [7] [17]. Later other researchers gave a new algorithm to find a $(2,2)$ partition graph [13] [17] .Special families of (k,l) partition graphs include: (k,l) chordal graphs [16], (k,l) cographs [12] , (k,l) perfect graphs [13] and (k,l) P_4-sparse graphs [8].

In this paper we studied $SN\,(k,l)$ graph which is a generalisation of (k,l) partition graph for some values of k and l. This class also contains the class of split neighborhood graph [3],[2] [21] and the quasi adjoint graph [1].

The purpose of this paper is to propose polynomials algorithms for the recognition problem and the size of the maximum clique problem of $SN\,(k,l)$ for $1 \leqslant l,k \leqslant 2$ graphs. We also present a polynomial algorithm for the optimal coloring of perfect $SN\,(1,2)$ graphs. The latter algorithm is interresting because it is a direct consequence of the validity of the SPGC. Note that a strong perfect graph conjecture (SPGC) due to Berge [5] ≪states a graph G is perfect (G perfect if the chromatic number is equal to maximum clique size for every induced subgraph H of G) if and only G is Berge graph (G Berge graph if it is contains neither odd hole, nor odd anti hole)≫. This conjecture is resolved by Chudnovsky and al [11][10], who called it the strong perfect graph theorem (denoted SPGT). However this proof does not provide any algorithm for coloring perfect graphs.

On the other hand, the problem to determine an optimal coloring (ie the chromatic number) of a graph is NP-complete in the general case but polynomial for perfect graphs due to Grötschell and al who developed a polynomial algorithm to solve combinatorial optimization problems the class of perfect graphs using an alternative of the ellipsoids method for the resolution of linear programs. Their algorithms are not practically efficient, undoubtedly because they do not take the combinatorial structure of perfect graphs into account [14] .Given a real situation may be mod-

*Hayat Issaadi, and Hacene Ait Haddadeneis with USTHB-University, Mathematics Faculty, Operations Research Departement, LaROMaD-Laboratory, BP 32 El-Alia, Bab-Ezzouar 16111, Algiers, Algeria. e-mail: (issaadi-hayat@yahoo.fr and aithaddadenehacene@yahoo.fr).

†Safia Zenia are with Ecole Nationale des Veterinaires,El Harrach, Alger,Algerie. e-mail:(e-mail:safiazenia@yahoo.fr)

‡Manuscript received April 19, 2005; revised January 11, 2007.

eled as a graph coloring problem of a particular class of graph such as a class of perfect $SN(1,2)$ graphs where we have a coloring algorithm this can be a very interesting tool for solving such practical problems.

2 Preliminary and Background

All graphs considered here are finite, undirected without loops or multiple edges. Let G = (V, E) be a graph, $|V| = n$ and $|E| = m$, $V' \subset V$ is stable (denoted S) set iff for all $u, v \in V'$ u is not adjacent to v. $V' \subset V$ is a clique(denoted K) iff for all $u, v \in V'$, u is adjacent to v. A k-coloring is a mapping C : $V \to \{1, 2, ..., k\}$ such that $C(u) \neq C(v)$ for every edge uv. Note that each color class is a stable set , hence a k-coloring can be thought of as a partition of the vertices of a graph into stable sets $S_1, ..., S_k$. The chromatic number $\chi(G)$ is the smallest k such that G admits a k-coloring. $\omega(G)$ is the size of largest clique in G. In a graph $G = (V, E)$,a subgraph induced by $X \subseteq V$ is denoted $G[X]$. For $v \in V$, we denote by $N_G(v)$ the neighborhood of a vertex v in a graph G and by $N_G[v] = N_G(v) \cup \{v\}$. In section 3 we will propose an algorithm which determines an ordering of the vertices $v_1, ..., v_n$ of G such that $N_G(v_i)$ is a (k, l) partition graph for $1 \leq k, l \leq 2$ using Brandstädt's results [7]:

Theorem 1. *It can be recognized in $O(n^2m)$ steps whether a graph G is (2,1) partition graph.*

Corollary 1. *It can be recognized in $O(n^2m)$ steps whether a graph G is (1,2) partition graph.*

Theorem 2. *It can be recognized in $O(n^{10}m)$ steps whether a graph G is (2,2) partition graph.*

In section 4 we shall show that largest clique of a $SN(k, 2)$ graph for $1 \leq k \leq 2$ can be found in polynomial time. As there is a combinatorial polynomial algorithm for finding largest stable set and a minimum clique cover of a bipartite graph, Hoàng and Lê [17] used the following function (noted: Find-Omega-Bip) to compute largest clique of a $(2, 2)$ partition graph (called also 2-split graphs) see also [17].

Function Find-Omega-Bip(G)

Input: a graph G that is the complement of a bipartite graph.
Output: a number Find-Omega-Bip(G) = $\omega(G)$.

It is clear that if G is a (1,2)graph, the step 3 in the algorithm is not necessary.
In Section 5 we propose a method for coloring any perfect $SN(1, 2)$ graph G, using the following two methods of coloring: The trichromatic exchange method which was proposed by Ait Haddadene and Maffray in [4]. Let v be a vertex of a graph G with $\omega(G) \geq 4$ and assume $G - v$ has been $\omega(G)$-colored. Suppose that the following property holds: there exist a triple of distinct colors $i, j, k \in \{1, 2, ..., \omega(G)\}$ such that $G[S_i \bigcup S_j \bigcup S_k \bigcup \{v\}]$ is a $K_4 - free$ Berge subgraph. We can apply Tucker's algorithm [23] to 3-color the component of this subgraph containing v, and get in this way an $\omega(G)$-coloring of G in $O(n^3)$ time. We say that v is a Tucker vertex if the previous property holds for every $\omega(G)$-coloring of $G-v$.

Theorem 3. *[23] There exists a polynomial algorithm to color any K_4-free graph in $O(n^3)$.*

The second method was proposed by Hayward [15]. Two non adjacent vertices u and v in a graph G form a 2-pair if every chordless path between them has length two. For a given pair u, v in a graph G, we denote by G_{uv} the graph obtained by deleting u and v and adding a new vertex uv adjacent to precisely those vertices of $G - u - v$ which were adjacent to at least one of u or v in G. We say that G_{uv} is obtained by contracting on u, v. The importance of this contraction operation is that u, v is a 2-pair in G then the chromatic number of G is equal to the chromatic number of G_{uv} , this fact yields a simple procedure which given a k-coloring of G_{uv} yields a k-coloring of G. As we shall see, these procedures can be used to develop fast algorithms for finding an optimal coloring [2][20][22].

ISBN: 978-1-61804-292-7

3 Recognizing $SN\,(k,l))$ graphs for some values of k and l

Let G $SN\,(k,l))$ be a graph. We call an ordering a $SN\,(k,l)$ elimination ordering in G if we can order the vertex set of G as follow: $v_1, v_2, ..., v_n$, where $N(v_i)$ is a (k,l) partition graph in the induced subgraph $G_i = G[v_i, v_{i+1}, ...v_n]$.

Theorem 4. *A graph G is a $SN\,(k,l)$ graph if and only if G admits a $SN\,(k,l)$ elimination ordering of its vertices.*

Proof. If G is a $SN\,(k,l)$ graph, then for each $H \subseteq G$ in particular, G has a vertex v_1 whose neighborhood can be partitioned into k cliques and l stable sets. As a result, we can always find a vertex v_i in $G_i = G[v_i; v_{i+1}; ...; v_n]$ such that $G[N(v_i)]$ can be partitioned into k cliques and l stable sets. So, we can find a $SN\,(k,l)$ elimination ordering $v_1, v_2, ..., v_n$ of the vertices of G. Let $v_1, v_2, ..., v_n$ be an elimination ordering of G's vertices and suppose that G is not a $SN\,(k,l)$. Let $H \subset G$, H is induced by the vertices $v_1, v_2, ..., v_{|V(H)|}$ we can find a vertex v_i among the vertices $v_1, v_2, ..., v_{|V(H)|}$ such that $G[N(v_i)]$ can be partitioned into k cliques and l stable sets in G. As a result, $N(v_i)$ can be partitioned into k cliques and l stable sets in H because $H \subset G$: This implies that G is a $SN\,(k,l)$ graph which is a contradiction. □

Algorithm 3.1

- **Input:** G = (V, E) a graph, $|V(G)| = n$, $|E(G)| = m$.
- **Output:**
  - G is not $SN\,(1,2)$ (respectively $SN\,(2,1)$) graph.
  - $SN\,(1,2)$ (respectively $SN\,(2,1)$)is split vertex elimination numbering $v_1, v_2, ..., v_n$.

Begin $G_1 := G$;
For $i := 1$ for n **Do**

- Find a vertex v_i of G_i such as $G[N(v_i)]$ is a $(1,2)$ or $(2,1)$ (respectively $(2,2)$ partition graph;
- If v_i does not exist G is neither $SN\,(1,2)$ or $SN\,(2,1)$ graph (respectively $(2,2)$; if not $G_{i+1} := G_i - v_i$;

End
We can determine this $SN\,(1,2)$ elimination ordering in polynomial time. Effectively, find a vertex v_i of G_i such as $G[N_{G_i}(v_i)]$ is a $(1,2)$ (respectively $(2,1)$) partition, this is done in $O(n^2m)$ by Brandsätdt's algorithm. So the determination of $SN\,(1,2)$ (respectively $SN\,(2,1)$) elimination ordering can be performed in $O(n^3m)$ and the recognition of $SN(2,2)$ graphs will be of complexity $O(n^{11}m)$.

4 Maximum clique of $SN\,(k,l)$ graphs for some values of k and l

Using Algorithm 1 [17](section 2), we shall show that largest clique of $SN\,(1,2)$ graph (respectively $SN\,(2,2)$) can be found in $O(n^{4.5})$ (respectively in $O(n^{4.5}m)$). Let $v_1, v_2, ..., v_n$ a $SN\,(k,l)$ elimination ordering. To compute the size of the maximum clique of G denoted by $\omega(G)$, we can used: $\omega(G) = Max\,\{\omega(G[N[v_i]]), \omega(G - v_i)\}$.
Algorithm 4.1

- **Input:** G = (V, E) is a graph with $SN\,(1,2)$ elimination ordering $v_1, v_2, ..., v_n$ and a partition of $N_G(v_i)$ K_1, K_2, S where K_i's are cliques and S stable set .
- **Result:** Find $\omega(G)$

Begin $G_1 = G$

1. **for** $i = 1, ..., n$ **do**
   Find $T = Omega - Bip(K_1 \bigcup K_2 \bigcup \{v_i\})$
2. **For** each vertex $s \in S$ **Do**
   $k = \omega(K_1 \bigcup K_2 \bigcup \{v_i\} \bigcap N(s))$
   If $k + 1 > T$ **Then** $T = k + 1$
   Else return **(1)**
3. Find $\omega(G) = Max\,\{\omega(G[N[v_i]]), \omega(G - v_i)\}$

ISBN: 978-1-61804-292-7

End
Algorithm 4.2

- **Input:** G = (V, E)is a graph with $SN(2,2)$ elimination ordering $v_1, v_2, ..., v_n$ and a partition of $N_G(v_i)$ K_1, K_2, S_1, S_2 where K_i's are cliques and S_i's are stable sets .
- **Result:** Find $\omega(G)$

Begin $G_1 = G$

1. **For** $i = 1, ..., n$ **Do**
   Find $T = Omega - Bip(K_1 \bigcup K_2 \bigcup \{v_i\})$

2. **For** each vertex $s \in S$ **Do**
   $k = \omega(K_1 \bigcup K_2 \bigcup \{v_i\} \bigcap N(s))$
   If $k + 1 > T$ **Then** $T = k + 1$

3. **ElseFor** each edge ab with $s_1, s_2 \in S_1 \cup S_2$, **Do**
   k := Find-Omega-Bip $(N(s_1) \cap N(s_2) \cap (\{v_i\} \cup K_1 \cup K_2))$;
   If k + 2 > max **Then** $T := k + 2$, return **(1)**

4) Find $\omega(G) = Max\{\omega(G[N[v_i]]), \omega(G - v_i)\}$
End

The largest clique in $SN(1,2)$ or in $SN(2,2)$ graphs can be found in polynomial time because the value of $\omega(K_1 \bigcup K_2 \bigcup \{v_i\})$ can be calculated efficiently in $O(n^{2.5})$ time using the algorithm of Hopcroft and Karp (1973) [18] since this operation is equivalent to determining a maximum stable set in the complement graph of $K_1 \bigcup K_2 \bigcup \{v_i\}$ which is a bipartite graph. Then for each $s \in S$; there is at most n; (respectively $s_1, s_2 \in S_1 \cup S_2$; there is at most m) find largest clique in $K_1 \bigcup K_2 \bigcup \{v_i\} \bigcap N(s)$ (respectively $N(s_1) \cap N(s_2) \cap (\{v_i\} \cup K_1 \cup K_2)$. As the number of the vertices of G is n then the overall complexity of the algorithm 4.1 is $O(n^{4.5})$ and of algorithm 4.2 is $O(n^{4.5}m)$.

5 Optimal coloring of perfect $SN(1,2)$ graph

Let G be a $SN(1,2)$, v a $(1,2)$ split vertex then $\forall H \subseteq G$, $N_H(v)$ is induced by a stable set S and two cliques K_1 and K_2. We denote $K_1 \bigcup K_2 = A$. Note that Either $\omega(K_1) = \omega(K_2) = \omega(G)$ or $\omega(K_i) < \omega(G) - 1$ for $i = 1, 2$. Without lost of generality $\omega(K_1) < \omega - 1$ and $\omega(K_2) = \omega - 1$. Our principal result is the following:

Theorem 5. *Let G=(V, E) be a perfect $SN(1,2)$ graph and v a $(1,2)$ split vertex. Then from any $\omega(G)$-coloring of G-v one can obtain an $\omega(G)$-coloring of G in polynomial time.*

For the proof of theorem 5.1, we will need a result of Tucker [23] and proposition 5.2.

Proposition 1. *Let G be a $SN(k,2)$ graphs $(k \geq 1)$ without hole of length ≥ 5 and v a $(k,2)$ split vertex*

1. *$\forall H \subseteq G$ there is always a 2-pair in $\{v\} \cup N_H(v)$; otherwise A is a clique and $A \cap N_H(s) = \emptyset$, $\forall s \in S$.*
2. *The contraction of all 2-pair in $\{v\} \cup N_H(v)$ induces a $SN(k,2)$ graphs for $k \geq 1$.*

Proof. Let G be a $SN(k,2)$ graphs, v a $(k,2)$ split vertex and H a subgraph of G .

1. Let us suppose that there is not a 2-pair in $\{v\} \cup N_H(v)$ i.e. $\forall (a,b) \in \{v\} \cup N_H(v)$, a is connected to b by a chain P of length $\geqslant 3$; $P = \{a, v_1, ..., v_k, b\}$ for $k \geqslant 2$; but $P \cup \{v\}$ is a hole of lenght $\geqslant 5$,contradiction. So if $\forall (a,b) \in \{v\} \cup N_H(v)$ a is not connected to b in $N_H(v)$ then A is a clique, $\bigcup_{i=1}^{k} S_i$ is a stable set and $\forall s \in \bigcup_{i=1}^{k} S_i, N_{N_H(v)}(s) \cap A = \phi$

2. Let (a,b) be a 2-pair in $N_H(v)$ the contracted vertex denoted ab is adjacent to every neighbord of a and b:
   a) Either $(a,b) \in \bigcup_{i=1}^{k} S_i$ so the contraction yields $\bigcup_{i=1}^{k} S_i^*$ with size $|\bigcup_{i=1}^{k} S_i| - 1$.
   b)Or (a,b) are in the cliques K_1, K_2 so the contraction yields cliques K_1^*, K_2^*. If there exist a clique which all vertices are contracted then $N_H^*(v)$ will partition in $(k,1)$.
   c) Or (a,b) are in $\bigcup_{i=1}^{k} S_i$ and in the clique $K_j, j = 1, 2$ respectivly, so the contraction yields $\bigcup_{i=1}^{k} S_i^*$ with size $|\bigcup_{i=1}^{k} S_i| - 1$.

ISBN: 978-1-61804-292-7

□

Method of coloring: Lets G=(V, E) be a perfect $SN\,(1,2)$ graph, v a $(1,2)$ split vertex, $\{v_1, v_2, ..., v_n\}$ a $SN\,(k,l)$ elimination ordering of G and M its adjacency matrix. As G[N(v)] can be partitioned into two cliques A and a stable set S, assume that $|K_k| = \omega - p$ for $p \geq 1$ for $k = 1, 2$. There are at least one color which is misse in A. Let the coloring of A with colors:$1, 2, ..., \omega - p, p \geq 1$ such as we will index the vertices of K_k: $v_1, v_2, ..., v_{\omega-p}$ so that $C(v_i) = i$ for $i = 1, ..., \omega - p$, $p \geq 1$. Let us denote by S_i a stable of color i, and let us index the vertex of the stable set S by $s_1, s_2, ..., s_{|S|}$ so that any vertex of the stable set is denoted s_k for $k = 1, ..., |S|$.
$\forall i = 1, ..., n$, we contract all 2-pair existing in $N_{G_i}(v_i)$. We will show (proof theorem 5.1) that the contracted graph denoted G^* can have an optimal coloring in polynomial time. Then to relax the graph G^*, it is sufficient to relax $G_i^*, \forall i = 1, ..n$ and give the same color at the relaxed vertex. In an iterative way we manage to have an optimal coloring of G with $\omega(G)$-colors. Let G be a graph such as $\omega(G) \geq 5$ (if not $G[v \bigcup N_G(v)]$ is K_4- free whose coloring is done in polynomial time by the algorithm of Tucker [23]).

Proof. (Theorem 5.1)
Let G be a perfect $SN\,(1,2)$ graph and let v be a (1,2) split vertex , there is a vertex $v_i \in A = K_1 \bigcup K_2, i = 1, 2(\omega - p)$, $s_k \in S, k = 1, ..., |S|$ such as the couple (s_k, v_i) form a 2- pair in $N_G(v)$ (proposition 5.2). Let G^* be the graph obtained from G after a sequence of contraction of the existing 2- pairs in $N_G(v)$. Lets $S_{2-pair} = \{s_k \in S$ for $k = 1, ..., |S|/v_i \in A(i = \overline{1, 2(\omega - p)}, p \geq 1)$ the couple of vertices (s_k, v_i) form a 2-pair in $N_G(v)\}$, $N^*(v)$ the neighborhood of v in G^*. We will distinguish two cases:

1. Either $|S| = |S_{2-pair}|$. In this case after the sequence of contraction, $N^*(v)$ is a partition of two cliques of size $(\omega - p), p \geq 1$. Therefore, the coloring of $N^*(v)$ G with ω(G)-colors is the coloring of the clique K with at most $\omega - p$ colors and the remaining color will be assigned to v.
2. Or $\exists$ at least a vertex $s \in S/S_{2-pair}$ (it is clear that s is not adjacent to A in $N_G(v)$). In this case after the sequence of contraction, $N^*(v)$ is $(1,2)$ split vertex (proposition 5.2). Moreover s is not adjacent to A in $N^*(v)$, so let us show that v is a Tucker's vertex. Let us suppose the opposite. Then for each triple of distinct colors $i, \alpha, j \in 1, ..., \omega - p$ $p \geq 1$ the subgraph $G[S_i \bigcup S_\alpha \bigcup S_j \bigcup v]$ contains a K_4. Obviously, this implies that for each $i, \alpha, j \in 1, ..., \omega - p$, $p \geq 1$ there is a triangle whose vertices are colored i,α, j but the stable set S is not adjacent to A then we can use $\omega - p, p \geq 1$ colors for the coloring of A and the remaining color will to assign to v, which is contradiction. Finally relaxed the graph G by assigning the same color to the relaxed vertices.

□

Let G be a perfect $SN\,(1,2)$ graph and v a $(1,2)$ split vertex. For a given a $SN\,(k,l)$ elimination ordering of G and each subgraph G_i induced by $\{v_i, ..., v_n\}$ (complexity $O(n^3m)$) ; $i \in \{1, ..., n\}$; compute $\omega(G_i)$ for $i = 1, ..., n-1$, (complexity $O(n^{4.5})$). We determine all 2- pair in $N_{G_{i+1}}(v_{i+1})$ (complexity $O(n^2m)$ [15]) then start from the trivial coloring of G_n^* , and iteratively find an optimal coloring of G_i^* from an optimal coloring of G_{i+1}^* using proof of theorem 5.1, complexity at most $O(n^3)$. Finally we relax the graph G_i^* for $i = 1..., , n$ and start from the trivial coloring of G_n, and iteratively find an optimal coloring of G_i from an optimal coloring of G_{i+1} (complexity at most O(n)). This yields an $O(n^5)$ algorithm for finding an $\omega(G)$-coloring of G.
Algorithm 5.1

- **Input:** A perfect $SN\,(1,2)$ graph G and its adjacency matrix M.
- **Result:** Optimal coloring of G

1. Determine a $SN\,(k,l)$ elimination ordering $v_1, ..., v_n$;
2. **For** i decreasing from n-1 to 1 **Do**
   Determine $\omega(G)$;

ISBN: 978-1-61804-292-7

3. Determine the adjacency matrix of G;

4. **For** i decreasing from n-1 to 1 **Do**
   Determine all 2-pairs in $N_{G_i}(v_i)$;

5. **For** i decreasing from n-1 to 1,**Do**
   From an optimally coloring of G^*_{i+1},find an optimally coloring of G^*_i by applying Theorem 5.1;

6. **For** i decreasing from n-1 to 1 **Do**
   Relax the graph by giving the same color at the vertices relaxed.

Conclusion: In this paper, we focus in a generalization of (k, l) partition graph , where we proposed polynomial algorithms for recognition, maximum clique of this generalization and optimal coloring for any perfect graphs of this class for some values of k and l.This work can lead to other interesting views of research.

References

[1] H Ait Haddadene and A Hamadi. Coloring perfect quasi-locally quasi-adjoint graphs. In *Proceedings of the Modelling, Computation, Optimization in information system and Management*, volume 1. Metz, Jul 2004.

[2] H Ait Haddadene and H Issaadi. Perfect graphs and vertex coloring problem. *IAENG International Journal of Applied Mathematics*, 39(2):128–133, 2009.

[3] H Ait Haddadene and H Issaadi. Coloring of split neighborhood graph. Ciro 2005.

[4] H Ait Haddadene and F Maffray. Coloring perfect degenerate graphs. *Discrete Mathematics*, 163:211–215, 1997.

[5] C Berge. les problemes de coloration en theorie des graphes. *Publications de l'Institut de Statistique Universit Paris 9*, pages 123–160, 1960.

[6] A Brandstadt, B Le, V, and T Szymczak. The complexity of some problems related to graph 3-colorability. *Discrete Appl. Math*, 89:59–73, 1998.

[7] A Brandstädt. Partition of graphs into one or two independent sets and cliques. *Discrete Maths*, pages 47–54, 1998.

[8] S F Bravo, R, S Klein, T Nogueira, L, and F. Protti. Characterization and recognition of p_4-sparse graphs partitionable into k independent sets and cliques. *Discrete Applied Mathematics*, (159):165–173, 2011.

[9] K Cameron, M Eschen, E, T Hoàng, C, and R. Sritharan. The complexity of the list partition problem for graphs. *SIAM J. Discrete Math*, 21(4):900–929, 2007.

[10] M Chudnovsky, G Cornujols, P Seymour, X Liut, and K Vuskovic. Recognizing berge graphs. *Combinatorica*, 25(2):143–186, 2004.

[11] M Chudnovsky, N Robertson, P Seymour, and R Thomas. Progress on perfect graph. *Math.Programming, Ser B*, (97):405–422, 2003.

[12] M Demange, T Ekim, and D. de Werra. Partitioning cographs into cliques and stable sets. *Discrete Optimization 2*, pages 145–153, 2005.

[13] T Feder, P Hell, R Klein, and S Motwani. Complexity of graph partition problems. pages 464–472. 31st Annual ACM Symposium on Theory of Computing, Plenum Press,New York, 1999.

[14] M Grotschel, L Lovasz, and A Schrijver. Polynomial algorithms for perfect graphs. *Ann Disc Math*, (21):325–356, 1984.

[15] R Hayward. Weakly triangulated graph. *Journal of Theory*, B39:200–208, 1985.

[16] P Hell, S Klein, T Nogueira, L, and F Protti. Partitioning chordal graphs into independent sets and cliques. *Discrete Applied Mathematics*, (141):185–194, 2004.

[17] T Hoang, C and B Le, V. Recognizing perfect 2-split graphs. *SIAM J, Discrete Math*, (13):48–55, 2000.

ISBN: 978-1-61804-292-7

[18] E Hopcroft, J and M. Karp, R. An $o(n^{2.5})$ algorithm for maximum matchings in bipartite graphs. *SIAM. J. Comput*, 2:225–231, 1973.

[19] Y Huanga and Y Chub. note on the computational complexity of graph vertex partition. *Discrete Applied Mathematics*, (155):405–409, 2007.

[20] F Maffray. On the coloration of perfect graphs. In *CMS Books In Mathemathics*. Springer editors ba-reed and c.l.sales edition, 2003.

[21] F Maffray and M Preissmann. Split - neighbourhood graphs and the strong perfect graph conjecture. *Combinatorial Theory Serie B*, pages 1–21, 1995.

[22] A Malaguti and P.Toth. A survey on vertex coloring problems. *Intl.Transin OP. Res*, (17):1–34, 2010.

[23] A Tucker. A reduction procedure for coloring perfect k_4 - free graphs. *Combinatorial Theory*, (B43):151 – 172, 1987.

ISBN: 978-1-61804-292-7

Computing Non-Hydrostatic Pressure on Flip Buckets by Processing *NASIR* Finite Volume Solver Results

Saeed-Reza Sabbagh-Yazdi,

Professor of Civil Engineering Department KNToosi University of Technology, Tehran, Iran

(SYazdi@KNTU.ac.ir),

VahidKermani,

MSc Graduate of Civil Engineering Department KNToosi University of Technology, Tehran, Iran

and

Nikos Mastorakis

Professor of Technical University of Sofia

Abstract

Due to the centrifugal force, pressure distribution is non-hydrostatic in flip bucket spillways,. In this work, two dimensional modeling of supercritical flow parameters which is computed by ***NASIR***software and local curvature are used forpost-processing and calculation of dynamic pressure on every grid points of the bucket bed. The module of the utilized software solves Shallow Water Flow-solver adopted for steep slopes on the unstructured triangular meshes using finite volume method and computes flow depth and depth average velocities. The curvature is modeled using geometrical features of the unstructured triangular mesh which is utilized for modeling the three-dimensional bed surface. The geometrical modeling of the bed surface curvatures is utilized for estimating the vertical curvatures on the nodal points of the computational mesh. An analytical relation which utilizes bed curvature value is applied for calculating the vertical distribution of local dynamic pressure at the mesh nodes. This simulation strategy is verified by evaluating non-hydrostatic pressure for ending curves of the spillway buckets at the end of supercritical flow water were modeled. The non-hydrostatic pressure results of present modeling strategy are compared with the reported experimental measurements.

Keywords: Flip Bucket, Non-Hydrostatic Pressure, ***NASIR*** Solver, Shallow Water Equations

1. Introduction

For design of the flip bucket spillways, pressure distributions at bucket bed as one of the vulnerable parts of a spillway have to be evaluated. The bucket of the spillways is one of the points that sever pressure changes may accrue due to the vertical curvature of the super critical flow bed and the presence of high velocity flow in the locations with a vertical curve. This condition cause significant centrifugal forces, and consequently, excess positive or negative non-hydrostatic pressure would be developed in these places.

Developments in powerful computer hard-wares and capable software's have made the numerical simulation as a suitable means for modeling the real world engineering cases. For the flow problems in which variation parameters in current depth is negligible, the horizontal two-dimensional (depth averaged) numerical flow solvers are attractive alternatives [1]. Such a model is practically applicable for supercritical flow in which the depth average value covers most of the profile normal to the bed surface [2].

However, the shallow water equations which are commonly used as the most appropriate mathematical model for horizontal two-dimensional simulation, suffers from two major restriction due to considering mild slope and hydrostatic assumptions.

The mild slope application restriction of horizontal two-dimensional flow solvers is recently relaxed by modification of the shallow water equations for steep slope (in the main flow direction) [3]. Such a model is successfully applied for modeling supercritical flow on some types of chute spillways with variable slope [4]. This modeling strategy, paved the way for post-processing of the air concentrations (entered from the free surface and bottom aerators) [5].

In present work, a post-processing on computed depth averaged flow parameter are proposed for treatment of the non-hydrostatic values of the pressure field in the vicinity of the vertically curved supercritical flow bed. This treatment is performed by utilizing the analytical relations [6] which calculates local centrifugal pressure using the computed depth averaged flow parameters (by a version of ***NASIR***[1] software which solves modified Shallow Water Equations for variable steep slope surface [4]) and nodal values of the curvature (on the mesh of the three-dimensional bed surface).

2. Depth Average Mathematical Model for Varying Steep Slopes

Mathematical model used in the depth averaged flow solver module of ***NASIR*** software includes shallow water equations corrected for varying steep slope as [4]:

$$\frac{\partial h'}{\partial t}+\frac{\partial (h'u')}{\partial x'}+\frac{\partial (h'v)}{\partial y}=0$$

$$\frac{\partial (h'u')}{\partial t}+\frac{\partial (u'h'u')}{\partial x'}+\frac{\partial (vh'u')}{\partial y}+\frac{\partial}{\partial x'}\left[h'\frac{gh'/\cos\alpha}{2}\right]+\frac{gh'^2}{2\rho_w\cos\alpha}\frac{\partial \rho_m}{\partial x'}=gh'Sin\alpha-gh'S_{fx'}$$

$$\frac{\partial (h'v)}{\partial t}+\frac{\partial (u'h'v)}{\partial x'}+\frac{\partial (vh'v)}{\partial y}+\frac{\partial}{\partial y}\left[h'\frac{gh'/\cos\alpha}{2}\right]+\frac{gh'^2}{2\rho_w\cos\alpha}\frac{\partial \rho_m}{\partial y}=-gh'S_{fy}$$

here

$$S_{fx'}=\frac{n^2u'\sqrt{u'^2+v^2}}{h'^{4/3}} \quad ; \quad S_{fy}=\frac{n^2v\sqrt{u'^2+v^2}}{h'^{4/3}}$$

and x' tangential axis to the bed in slope direction and horizontal y axis are general coordinates. u' and v are velocity components in

[1] ***N***umerical ***A***nalyzer for ***S***cientific and ***I***ndustrial ***R***equirements (***NASIR***)

ISBN: 978-1-61804-292-7

direction of x' and y, h' is flow depth, and g is acceleration due to gravity. α is bed slope angle fallowing x', and $S_{fx'}$ and S_{fy} are friction slopes in direction of x' and y. n is Manning roughness coefficient. ρ_m is water and air mixed density, ρ_w is the pure water density, and C_{mean} is air averaged density. For the cases that water and air mixed density is not considered, term $\frac{gh'^2}{2\rho_w \cos\alpha}\frac{\partial \rho_m}{\partial x}$ and $\frac{gh'^2}{2\rho_w \cos\alpha}\frac{\partial \rho_m}{\partial y}$ will be omitted from the above equations.

3. Finite Volume Formulation of the Depth Average Flow Equations

Unstructured triangular grid and finite volume numerical solution method have been used in this modeling.In order to do this flow equations given in previous part could be written in fallowing vector form:

$$\frac{\partial Q}{\partial t}+\frac{\partial E}{\partial x'}+\frac{\partial F}{\partial y}=S,$$

Where

$$E=\begin{bmatrix} h'u' \\ hu'^2+\frac{gh'^2}{2\cos\alpha}+\frac{\rho_m gh'^2}{2\rho_w\cos\alpha} \\ h'u'v \end{bmatrix},\ Q=\begin{bmatrix} h' \\ h'u' \\ h'v \end{bmatrix},$$

$$S=\begin{bmatrix} 0 \\ gh'(Sin\alpha-S_{fx}) \\ -gh'S_{fy} \end{bmatrix}, \text{ and}$$

$$F=\begin{bmatrix} h'v \\ h'u'v \\ h'v^2+\frac{gh'^2}{2\cos\alpha}+\frac{\rho_m gh'^2}{2\rho_w\cos\alpha} \end{bmatrix}.$$

If the above set of the equations be placed in general integrated over the control volume area Ω, it can be written as:

$$\int_\Omega\left(\frac{\partial Q}{\partial t}+\frac{\partial E}{\partial x'}+\frac{\partial F}{\partial y}\right)dx'dy=\int_\Omega S dx'dy.$$

The discrete finite volume formulation form of the above equation is written as:

$$Q^{n+1}=Q^n-\frac{\Delta t}{\Omega}\sum_{k=1}^{N}\left(\overline{E}\Delta x-\overline{F}\Delta y\right)_k+S\Delta t.$$

In which, $\overline{E}$ and $\overline{F}$ are average flux components at the boundary edge of k of the control volume with Δx and Δy Cartesian coordinate components, and N is up to the method of solution (cell vertices, cell center and Galerkin methods). It has to be mentioned that Q^n and Q^{n+1} refered to the vectors of the unknown variables at the centre of the control volume in two sequential iterations or time steps of explicit solution procedure.

The above mentioned finite volume formulation can be used for numerical computation of depth-averaged parameter (ie*u'*, *v* and *h*) at every nodal point of the utilized mesh. Therefore, the hydrostatic pressure can be calculated at any position of the flow depth of each nodal point.

Considering that the pressure has two components of hydrostatic and non-hydrostatic components, the non-hydrostatic component of the pressure should be calculated using computed local velocity at the mesh nodes with curvature in vertical plane.

4. Dynamic Pressure on Curved Beds

Over the flip bucket spillways, hydrostatic pressure assumption is not valid. In such parts, non-hydrostatic pressure (less or more than hydrostatic pressure) may form due to the curvatures in the flow bed. Hence, considering flow's curvature affect corrections have to be imposed on the hydrostatic pressure. In order to do that in the present research, analytic and

ISBN: 978-1-61804-292-7

experimental formulas have been used for calculation of local non-hydrostatic pressure.

An analytical relation which is suitable for evaluating the pressure distribution in flows over beds with vertical curves was derived by Chow, as [6]:

$$h_t = h_s + \frac{V^2 h_s}{gR},$$

Where, h_t is total local pressure head, h_s is local hydrostatic pressure head, V is local velocity (in mail flow direction), R is bed curvature in vertical plane and g is the gravity acceleration. The above analytical relation is derived considering basic assumption of uniform distribution of velocity and constant curve radii R along the segment associated with the location [6].

5. Curvature at nodal points of mesh

The relation reviewed in the former section calculates local non-hydrostatic pressure over the vertically carved beds using flow parameters and the bed curvature value.

Local velocity and hydrostatic pressure head which are required in this analytical relation can be obtained from the numerically computed results of depth-averaged flow equations, While bed curvature as a key parameter must be known by a reliable method.

Therefore, the total pressure including hydrostatic pressure and non-hydrostatic pressure head components, can be calculated by post-processing on the computed flow parameters and the curvature on the nodal points of the 3D surface mesh (representing the channel's bed geometric characteristics).

Local curvature in vertical plane ($\kappa = 1/R$) at every point has to be specified prior to the start of the computations. In order to calculate curvature should be computed at all points of a triangular unstructured mesh (which is converted as a 3D surface by assigning the z coordinate for every nodal point to model flow bed geometry).

Although in spillway design, there are common mathematical functions such as specific polynomials or conic curves which are used to define spillway bed geometry at every part, a general geometry modeling of nodal points with arbitrary generation, will be more comprehensive.

In order to consider the local curvature, the 3D surface mesh which is used to solve the depth-average flow equations can be utilized. The grid points of the mesh provide required geometric data in terms of Cartesian coordinate $X_k(x_k, y_k, z_k), k = 1,2,3,.....,n$. The curvature value at any nodal point of the mesh can be estimated by evaluating curvature of the curve fitted through the points.

Since the acceleration component normal to the direction of curvilinear flow is the major cause of non-hydrostatic pressure [6], that is to be calculated is the curvature values of the fitted curve along the main flow direction.

Therefore, application of an appropriate curve fitting method [7] in geometric modeling, to model the longitudinal profiles of the bed surface would be an admissible solution. The curvature obtained by curve fitting on the boundaries of the spillway parallel to the stream can be used for finding the curvature values at the grid points between two boundary lines by interpolation.

It should be noted that, in order to accurate geometry modeling (curve fitting) of the cases with multiple bed slopes and complex geometric features, there may require considering several independent segments. However, the mesh partition facilities of the flow solver for dividing

ISBN: 978-1-61804-292-7

the mesh zones according to various flow regimes can help resolving this requirement.

6. Evaluation of Modeling Results

In this study, hydro-static pressure at the bucket are modified by developing a post processing for systematic calculations of non-hydrostatic pressure as and adding to the hydrostatic pressure obtained from the results of depth-average flow solver. This post processing considers the supercritical flow parameters and vertical curvature of bed surface, and then, calculates consequent excessive non-hydrostatic pressure using analytical relations. Finally, the static pressure values resulted by flow solver are sum up algebraically, and the outcome would be total pressure.

6.1 Flip bucket at the end of steep slope chute

Inthis part, reported measurements of an experimental model whichwas introduced by*AAKhan&PM.Steffler (1996)*[8], are used for accuracy evaluation of the dynamic pressure computed by the present numerical simulation and post-processing.

The model boundary conditions for these cases are specified upstream depth (h_0) and vanishing derivatives of extra pressure and velocity variables. Flow conditions on inlet boundary have been entered within discharge per unit width. As downstream flow is supercritical, no conditions are applied at downstream end.

1.1 Comparison of Modeling Results with Experimental Results

The geometrical dimensions of the spillway are digitally modeled according to the laboratory model. The bottom surface of the spillway is modeled by an unstructured triangular mesh (Figure 1-a).

The solve domain is divided to five sections in this case based on the flow regime type as well as the bed slope and curvature. In table (1) domain partitions for systematic calculations of pressure over flip bucket model are presented.

In supercritical region without curvature the flow solver pressure results are acceptable as well as subcritical section. The section which has been chosen to distinguish the error is supercritical part that has curved bed. In figure 14 water surface profiles obtained from the flow solver has been compared with the experimental results. As it is noticeable, numerical results compared to the experimental data have a negligible error.

It is evident that the effect of the bucket curvature on pressure is increasing as it is a concave curve. As can be seen from Figure (2) and Table (2) results obtained from post-processing using experimental equations given by Heller et.al has much less error than using chow formula. As it is shown in Figure (3), using these experiential equations leads to achieve a smooth distribution for pressure.

According to Table (2) Averaged relative error of bed pressure values obtained by the present modeling (relations of Heller et.al.) for discharges of $q = 0.0187 m^2/s$ and $q = 0.0292 m^2/s$ respectively have been decreased 50% and 58% comparing with the hydrostatic pressure values.

7. Conclusion

In this paper, the hydrostatic pressure computed from a depth average flow solver (which is adopted for steep slopes) is modified by applying an analytic relation for calculation of dynamic pressure at spillway buckets. The analytical relations for calculation of centrifugal pressure force at bucket bed uses local curvature of the flow bed and flow velocity magnitude.

ISBN: 978-1-61804-292-7

Here, the required local curvature is calculated from the geometric features of the three dimensional unstructured triangular surface mesh which is used for numerical computation of depth averaged flow parameters (ie flow depth and depth averaged velocity components). The required local velocity magnitudes are calculated from the depth averaged velocity components. Finally, the total pressure at every point is calculated by sum of hydrostatic and non-hydrostatic pressures at the nodal points of the computational mesh. The comparison of the computed results with the reported experimental measurements shows promising agreements.

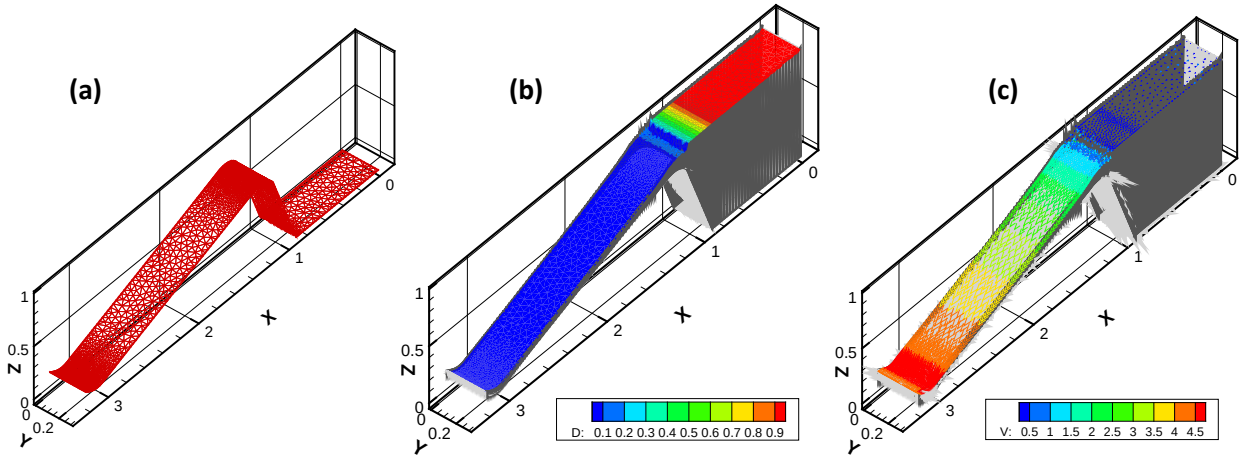


Figure 1 - Bucket profile at the end of slopping chute a) 3D view of unstructured triangular grid b) 3D view of flow depth map obtained from flow solver c) velocity vectors obtained from flow solver.

Table (1) - Domain partitions for systematic calculations of pressure over flip bucket

Section No.	Start Coordinates (m)	End Coordinates (m)	Flow Considerations	curvature	Fitted curve
1	0.0	1.15	Subcritical	No Curvature	-
2	1.15	1.3	Supercritical	Convex Curvature	Conic
3	1.3	3.00	Supercritical	No Curvature	-
4	3.00	3.25	Supercritical	Concave Curvature	Circle

ISBN: 978-1-61804-292-7

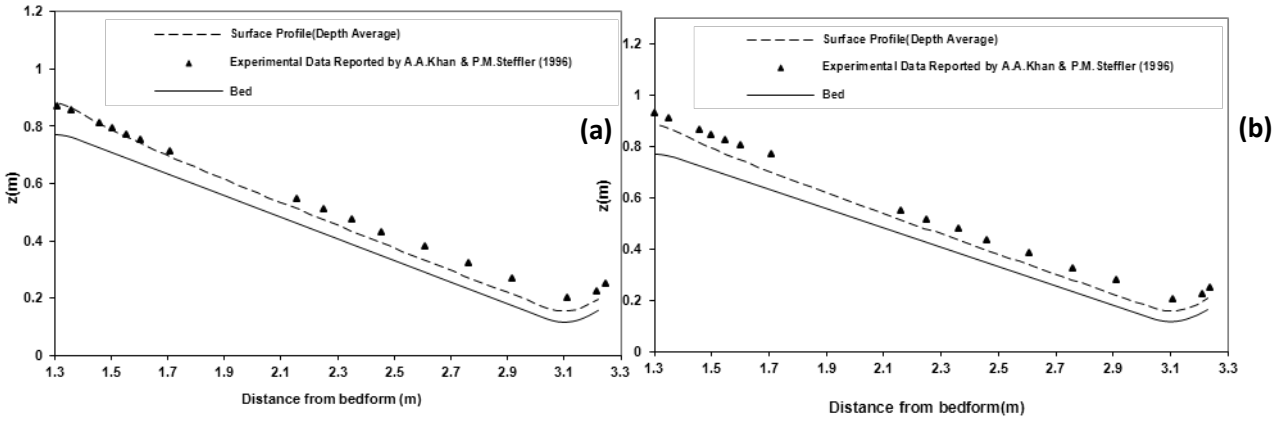


Figure 2-Comparisonofwater surfaceprofile obtained from the numerical model with the experimentalmeasurements for the flipbucketat the end ofaslopping chute a) $q = 0.0187m^2/s$ b) $q = 0.0292m^2/s$

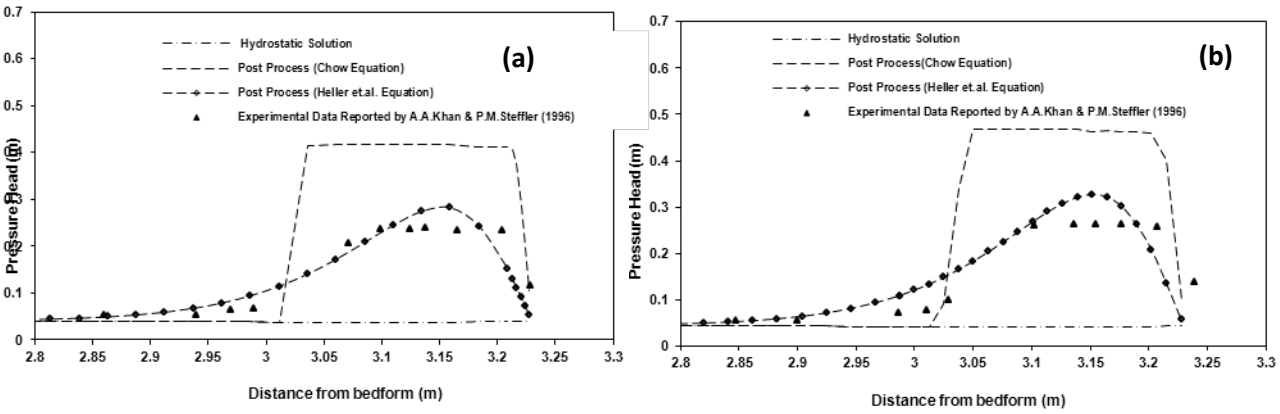


Figure 3 - Comparison of pressure obtained from the numerical model with the experimental results for the flip bucket at the end of slopping chute a) $q = 0.0187m^2/s$ b) $q = 0.0292m^2/s$

Table (2) – Comparison of relative error values between numerical models for the bucket at the end of slopping chute

Depth-averaged numerical model	**Averaged Relative Error for Pressure (%)**	
	$q = 0.0187m^2/s$	$q = 0.0292m^2/s$
Hydrostatic pressure	64.08	78.21
Corrected hydrostatic pressure (Chow formula)	69.35	80.11
Corrected hydrostatic pressure (Heller et al formula)	14.27	20.25

ISBN: 978-1-61804-292-7

8. References

[1] Bhallamudi, S.M. &Chaudhry, M.H. (1992) "Computation of Flows in Open Channel Transitions", J. of Hydraulic Engineering, Vol. 30, No. 1, pp.77-93.

[2] Sabbagh-Yazdi, S.R. &Mohammadzadeh, M. (2004) "Finite Volume Solution of mixed Sub & Super Critical 2D Free Surface Flow Using Unstructured Meshes", 9th Int. Con. On Hydroinformatics (IAHR),Singapore.

[3] Sabbagh-Yazdi, S.R. (2006) Spillway Flow Modeling by Finite Volume Solution of Slopping Depth Averaged Equations on Triangular Mesh; Application to KAROUN-4. 10th WSEAS International Conference on Applied Mathematics, Dallas (Texas), USA

[4] Sabbagh-Yazdi S.R. , Mastorakis EN &Zounemat-Kermani M (2007) " Velocity Profile over Spillway by Finite Volume Solution of Slopping Depth Averaged Flow", 2nd IASME/WSEAS International Conference on Continuum Mechanics, Protozoa (Porto rose), Slovenia, May 15-17

[5] Sabbagh-Yazdi S.R., Mastorakis N.E., and Safaieh R. (2008)"3 modeling strategies for computing aerated skimming flow parameters over stepped chutes using depth averaged flow solver", Int. J. of Mathematics and Computers in Simulations, Vol. 2, no.2, pp.134-143

[6] Chow, V.T (1973) "Open-Channel Hydraulics." Mc.Grow Hill Book, pp.444-448.

[7] Anand, V.B (1993) "Computer Graphics and Geometric Modeling for Engineers.", John Wile &Sons,Inc.

[8] Khan, A&Steffler, P.M (1996) "Vertically Averaged and Moment Equations Model for Flow over Curved Beds.", J. of Hydraulic Engineering, ASCE, Vol.122, No.1, pp.3-9.

Authors Index

ISBN: 978-1-61804-292-7