



International Journal
on Optimization and Applications

International Journal on Optimization and Applications

**VOL 01 - ISSUE 01
2021**

Editor in chief
Prof. Dr. Hanaa HACHIMI

ISSN : 2737-8314



International Journal On Optimization and Applications

**Vol 01 – Issue 01
2021**

**Editor in Chief
Prof. Dr. Hanaa HACHIMI**

ISSN : 2737-8314

FOREWORD

The International Journal on Optimization and Applications (IJOA) is an open access, double blind peer-reviewed online journal aiming at publishing high-quality research in all areas of : Applied mathematics, Engineering science, Artificial intelligence, Numerical Methods, Embedded Systems, Electric, Electronic engineering, Telecommunication Engineering... the IJOA begins its publication from 2021. This journal is enriched by very important special manuscripts that deal with problems using the latest methods of optimization. It aims to develop new ideas and collaborations, to be aware of the latest search trends in the optimization techniques and their applications in the various fields..

Finally, I would like to thank all participants who have contributed to the achievement of this journal and in particular the authors who have greatly enriched it with their performing articles.

Prof. Dr. Hanaa HACHIMI

Editor in chief

Associate Professor in Applied Mathematics & Computer Science

Systems Engineering Laboratory LGS Director, BOSS Team


Sultan Moulay Slimane University

TABLE OF CONTENTS

ARTICLE 1 -- IDENTIFICATION OF AN INDIVIDUAL IN A BIOMETRIC SYSTEM FINGERPRINT TECHNIQUE.....	1
ARTICLE 2 - CLASSIFICATION OF BREAST CANCER USING MACHINE LEARNING ALGORITHMS.....	6
ARTICLE 3 – CLASSIFICATION OF CREDITS IN THE GERMAN MARKET VIA SVM: ATTEMPT AT MODELING.....	11
ARTICLE 4 – FEATURE EXTRACTION AND FACE RECOGNITION USING AUTO ASSOCIATIVE MEMORIES	14
ARTICLE 5 -- HOW SCIENTIFIC RESEARCH IN MOROCCAN UNIVERSITY CAN USE TECHNOLOGY TO ENSURE AN EFFICIENT GOVERNANCE	19
ARTICLE 6 -- AN IMPROVED MODEL FOR COMPLEX OPTIMIZATION PROBLEMS BASED ON THE PSO ALGORITHM	23
ARTICLE 7 -- PROPOSED SOLUTIONS FOR SECURITY OF SMART TRAFFIC LIGHTS USING IOT AND MACHINE LEARNING.....	30
ARTICLE 8 – EFFECT OF CEV ON OPTION PRICES UNDER JUMP-DIFFUSION DYNAMICS WITH STOCHASTIC VOLATILITY: A FINITE ELEMENT METHOD APPROACH	36
ARTICLE 9 – WEAK SOLUTION TO A CLASS OF QUASILINEAR ELLIPTIC SYSTEM IN ORLICZ-SOBOLEV SPACES.....	41
ARTICLE 10 – ENTROPY SOLUTIONS FOR A NONLINEAR PARABOLIC PROBLEM WITH LOWER ORDER TERMS IN MUSIELAK-ORLICZ SPACES.....	49
ARTICLE 11 – OPTIMIZATION OF THE SUBSTATION LOCATION WITHIN AN OFFSHORE WIND FARM USING PARTICLE SWARM AND GENETIC ALGORITHMS.....	64
ARTICLE 12 – COMPARISON BETWEEN THE RING FLANGE AND THE NEW QUADRA-SECTOR FLANGE- BY MODELING AND ANALYZING OF THE MECHANICAL STRESS QUALIFICATION	69
ARTICLE 13 – SOLUTION OF FUZZY DIFFERENTIAL EQUATION BY APPROXIMATION OF FUZZY NUMBER.....	75
ARTICLE 14 – A NEW MODEL BASED ON GLOBAL HYBRIDATION OF MACHINE LEARNING TECHNIQUES FOR “CUSTOMER CHURN PREDICTION IN BANKING SECTOR”	81
ARTICLE 15 – SPREAD OF THE CORONA VIRUS EPIDEMIC IN MOROCCO, USA, FRANCE, SPAIN, ITALY, EGYPT, LEBANON AND TUNIS BY THE HETEROSCEDASTIC MODEL DURING THE FIRST WAVE	85

Identification of an individual in a biometric system Fingerprint Technique

Mohammed Boutalline

Systems Engineering Laboratory
Sultan Moulay Sliman University
Beni Mellal, Morocco
boutalline@gmail.com 

Hassan Faouzi

Sultan Moulay Sliman University
Beni Mellal, Morocco
faouzi.hassan.mi@gmail.com

Hamzaoui Radoine

Innovation Laboratory in Mathematics,
Applications & Information
Technologies
Sultan Moulay Sliman University
Beni Mellal, Morocco
info.hamzaoui@gmail.com

Belaïd Bouikhalene

Innovation Laboratory in Mathematics,
Applications & Information
Technologies
Sultan Moulay Sliman University
Beni Mellal, Morocco
b.bouikhalene@usms.ma

Mohamed Fakir

Information Processing and Decision
Support Laboratory
Sultan Moulay Sliman University
Beni Mellal, Morocco
m.fakir@usms.ma

Abstract—The fingerprints based about minutiae correspond essentially to the terminations and bifurcation of fingerprint patterns. Since the quality of fingerprint images is often low, automatic minutiae detection is a very difficult task and the extraction algorithms produce a large number of false alarms. A complete minutiae extraction scheme for automatic fingerprint recognition systems is presented. A great deal of work has been dedicated to fingerprint enhancement and to minutiae extraction. A new approach to combine different extraction algorithms is presented. The methods proposed, as confirmed by our simulations, allows an acceptable overall performance to be achieved.

Keywords— *Biometric system, fingerprints, minutiae, extraction, minutia matching.*

I. INTRODUCTION

A. General

Biometrics is a set of mathematical analysis of biological characteristics of a person (biometric technologies) that use human characteristics such as physical or behavioral fingerprint, signature, iris, voice, face, approach, and a hand gesture at the end to determine his identity conclusively. It also proves as a powerful tool for identification / verification to scenes such as ATMs, doors, scene of the crime in the legal sector.

The fingerprint recognition is a mature biometric application for any identification or verification of individuals. In this document describes the design and development of an automatic identity authentication by

fingerprint. This system of fingerprint recognition is based on a series of complex algorithms related to the fields of image processing and / or recognition of patterns.

The systems are increasingly found in applications related to security such as access control. There are always at least two modules in a biometric system: the learning module and the recognition (figure.1).

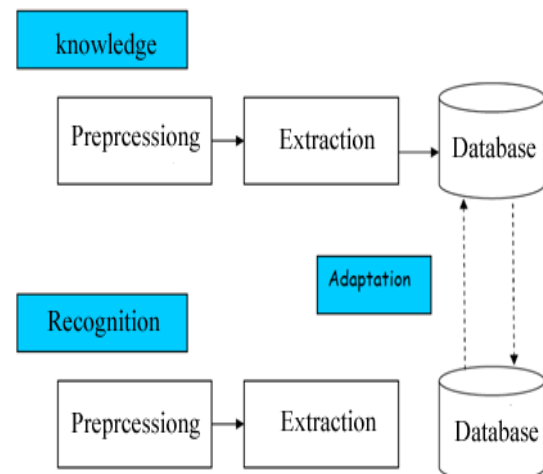


FIGURE 1. WORKING OF A BIOMETRIC SYSTEM

B. Fingerprint

Regarding the fingerprint, the British F. Galton [2] showed that in 1888 the first permanent design papillary birth to death, and unaltered. This particular arrangement of the papillary lines forms characteristic points, called minutiae, which are the cause of the individuality of digital designs.

A fingerprint consists of a lot of wrinkles and valleys. These wrinkles are characterized by minutiae. The figure.2 illustrates the types of minutiae.

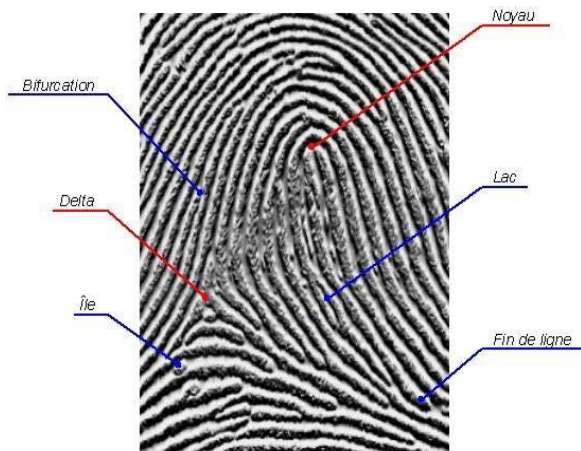


FIGURE.2. 6 MAIN TYPES OF MINUTIAE

The types of minutiae reported in the literature, two types are the most used are stopping ridge, the end of a ridge, and the bifurcation, and the point on the ridge in which two branches are derived (figure.3)



FIGURE.3. TWO TYPES OF MINUTIAES USED IN THE LITERATURE

C. Overview of different approaches

In general there are classes of algorithms for fingerprint recognition: The first category includes algorithms rather "conventional" that are based on the relative position of each minutiae, while the second contains algorithms to extract other features of the fingerprint such as the local management of ridges [3] and [4], or the frequency components of the local texture in the heart of the image [5]. The following diagram (figure.4) illustrates the steps of recognition of the fingerprint.

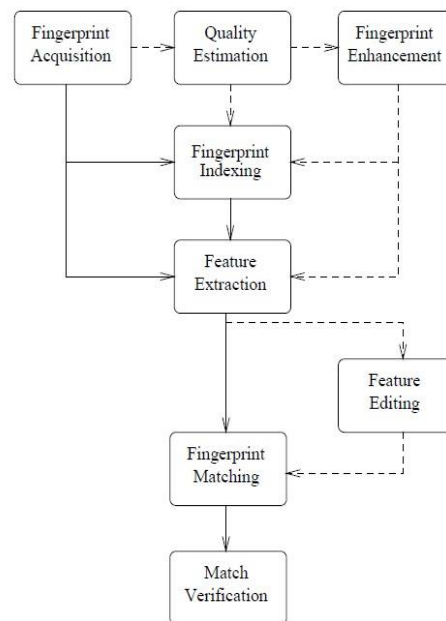


FIGURE.4. DESIGN OF A BIOMETRIC SYSTEM BASED ON FINGERPRINTS.

The approach, belonging to the first category is that proposed by A.K Jain [6] which is probably the best known. It is performed successively directional filtering and binarization of the image, the thinning (or skeletonization) of the ridges (figure.5), then determines the position of minutiae within the image to quantify the characteristics of similarity between two templates by « point pattern matching ».

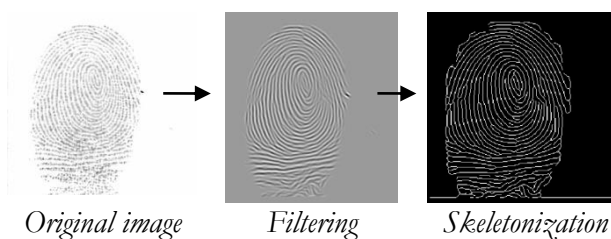


FIGURE.5. PREPROCESSING OF FINGERPRINT IMAGES TO MINUTIAE EXTRACTION

II. RECOGNITION ALGORITHM

A. Preprocessing

The first step toward improving the image of the fingerprint using the Gabor filter [11], the binarization and skeletonization. The aim of this treatment is to obtain an image of the orientation of the grooves of the fingerprint. In other words, is assigned to each pixel of the image orientation of the ridge which it belongs. The calculation of the orientation takes place the following proposed by Jain Anil et Al [10]:

$$V_x(i, j) = \sum_{u=i-\frac{w}{2}}^{i+\frac{w}{2}} \sum_{v=j-\frac{w}{2}}^{j+\frac{w}{2}} (2G_x(u, v)G_y(u, v)) \quad (1)$$

$$V_y(i, j) = \sum_{u=i-\frac{w}{2}}^{i+\frac{w}{2}} \sum_{v=j-\frac{w}{2}}^{j+\frac{w}{2}} (G_x^2(u, v) - G_y^2(u, v)) \quad (2)$$







$$\theta(i, j) = \frac{1}{2} \tan^{-1} \left(\frac{V_x(i, j)}{V_y(i, j)} \right) \quad (3)$$

With

- W is the size of local window;
- G_x and G_y are the gradient magnitudes in directions of x and y, respectively
- θ(i, j) is the local orientation of the pixel (i, j)

The second step is to share the directional image into sub blocks (W*W pixels) and characterize them by an average orientation (among the eight directions defined) and only one, as determined by histogram calculation what is the direction that occurs most frequently within each block. Then filtered the original image at each point by applying a directional convolution [11] (2D Gabor filters) based on the average direction of lines (power steering) previously calculated. After filtering, the different ridges and valleys by binarization. This is a local binarization (threshold in a block of pixels W * W) to overcome the non-uniformity of intensity over the entire image. Finally pre-treatment of the fingerprint image, a skeletonization is achieved, comparable to a thinning operation. The effectiveness of such an algorithm for thinning is difficult to quantify, our choice was guided by the experiments summarized in Table 1: they compare two similar algorithms (based on morphological iterative methods) skeletonization and highlight the many benefits of the approach « Zhang » [8] compared to that of « Shapori » [9].

TABLE I. EXPERIMENTAL COMPARISON OF ALGORITHMS FOR SKELETONIZATION OF ZHANG AND SHAPORI

	Shapori	Zhang
		
		

B. Minutiae extraction

The representation of attributes of the fingerprint is the most crucial phase in the design of a verification system. Representation clearly determines the accuracy. Existing systems are based on the minutiae of details.

The minutiae of the fingerprint are extracted from its skeleton by computing the "connectivity" CN at each point of the image P [1] as follows:

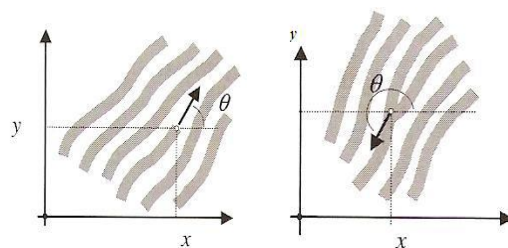
$$cn(p) = \frac{1}{2} \sum_{i=1 \dots 8} \text{val}(P_{i \bmod 8}) - \text{val}(P_{i-1}) \quad (4)$$

P₉ = P₁, P_i is the value of pixels in the neighborhood 3*3 of P. Indeed the coefficient CN has characteristics (Table 2) that identify the nature of a detail according to the result obtained in the calculation of CN.

TABLE II. IDENTIFICATION OF A MINUTIAE FROM THE NC CALCULATION

CN	Nature of the minutiae in P
0	Error => Point isolé
1	Termination
2	Error => Through the ridge
3	Bifurcation
4	Error => Defining a more complex minutiae.

Bifurcations and endings are two types of minutiae used the most because they are easily detectable (Figure.6). The information in each detail of a fingerprint are then included in a vector form (type of minutiae, x, y, θ), which is a convenient way to save the information of a fingerprint on a computer.



a) Representation of a termination. (b) Representation of a bifurcation.

FIGURE.6. DETERMINING THE DIRECTION OF THE MINUTIAE

Upon execution of this stage all the minutiae will be found but the presence of false minutiae due to the acquisition that is not flawed then the next step is

necessary to eliminate these minutiae. These will be eliminated in the post-processing step following simple rules of structures can be used to detect many false minutiae that generally affect thinned binary fingerprint images. Xiao and Raafat (1991b) [7] identified the common false minutiae structures and introduces an approach to the neighbor to remove (Figure.7).

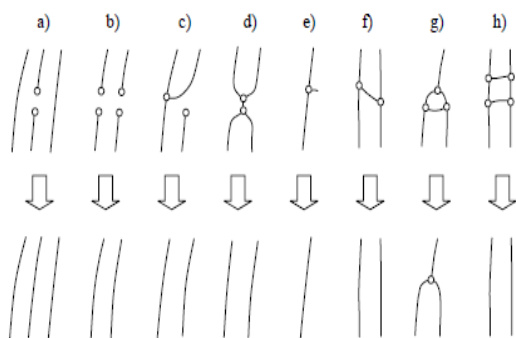


FIGURE.7. STRUCTURE OF THE MOST COMMON FALSE MINUTIAE

C. Comparison of minutiae

The comparison of a fingerprint with a database is to achieve agreement between a fingerprint image from a recording sheet (sheet 10 where the police are inked finger prints) and a latent, using minutiae.

The identity authentication system is based on comparing two sets of minutes, corresponding respectively to two fingers to compare. To determine whether two sets of minutiae extracted from two images correspond to the same finger prints, it is necessary to adopt a benchmarking system which is insensitive to possible translations, rotations and deformations that affect systematically fingerprints (Figure.8). From the same fingerprint, the calculation will never give 100% similarity due to the acquisition of two fingers (small deformations or displacements or rotations that affect systematically fingerprints).

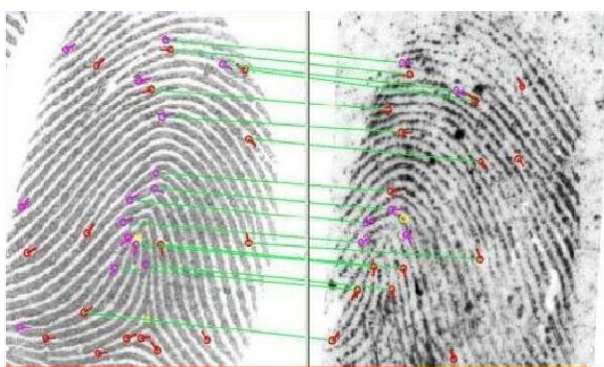


FIGURE.8. COMPARISON TEST WITH A MARKED TEMPLATE DATABASE

In the classification stage we used the method based on Euclidean distance, so to get it set up a test database containing the features extracted by the same method of learning and after it traverses the database of learning by calculating the Euclidean distance of the polar coordinates of the test image and the images of the database of learning, and the minimum distance will be that corresponding to the desired fingerprint.

$$d(x, y) = \sqrt{\sum_i^N (x_i - y_i)^2} \quad (5)$$

III. RESULTS AND CONCLUSION

Biometric work presented in this paper have led to the development of a system of identifying individuals by fingerprint recognition which the authentication phase itself (compare the clouds of minutiae).

To validate the system we worked on 50 tracks from five different people, 50% of prints were used for the learning phase, and 50% were used for tests to evaluate system performance.

Encouraging results appear. Indeed, the 50 tracks played, 44 were recognized, a recognition rate of 88%.

IV. BIBLIOGRAPHY

- [1] Anil Jain et al.,2000, Automated Fingerprint Identification and imaging Davide Maltoni, Dario Maio, Anil K. Salil Prabhakar , 2003, Fingerprint Handbook
- [2] Francis Galton, Fingerprint, McMillan, London, 1892.
- [3] U. Halici et G. Onguin. *Fingerprint classification through self-organizing feature maps modified to treat uncertainties*. Proceedings of the IEEE, Vol. 84, n° 10, Octobre 1996.
- [4] R. Capelli, A. Lumini, D. Maio et D. Maltoni. *Fingerprint classification by directional image partitioning*. IEEE Transactions on pattern analysis and machine intelligence, Vol. 21, n° 5, Mai 1999.

- [5] A.K. Jain et S. Pankanti. *FingerCode: a filterbank for fingerprint representation and matching*. IEEE, 1999.
- [6] A.K. Jain, L. Hong, S. Pankanti et R. Bolle. *An identity authentication system using fingerprints*. Proceedings of the IEEE, Vol. 85, n° 9, Septembre 1997.
- [7] Xiao and Raafat (1991b). Xiao Q. and Raafat H., “Fingerprint image post-processing: A combined statistical and structural approach,” *Pattern Recognition*, vol. 24, no. 10, pp. 985–992, 1991b.
- [8] J.R. Parker. *Algorithms for image processing and computer vision*. Wiley & Sons, Novembre 1996.
- [9] Haralick, Robert et Shapiro. *Computer and robot vision*. Vol. 1, Addison-Wesley, 1992.
- [10] Anil Jain et al., 2000, Automated Fingerprint Identification and imaging
- [11] <http://fr.wikipedia.org/>

Classification of breast cancer using Machine Learning Algorithms

Houssam Benbrahim

Engineering Sciences Laboratory,
National School of Applied Sciences,
Ibn Tofail University, Kenitra,
Morocco

houssam.benbrahim@uit.ac.ma

Hanaâ HACHIMI

Systems Engineering Laboratory,
Sultan Moulay Slimane University,
Beni Mellal, Morocco

hanaa.hachimi@usms.ac.ma

Aouatif AMINE

Engineering Sciences Laboratory,
National School of Applied Sciences,
Ibn Tofail University, Kenitra,
Morocco

aouatif.amine@uit.ac.ma

Abstract—In Morocco, the number of cases of breast cancer is increasing, and this can become dangerous due to a delay in the diagnosis phase or failure in the prediction of the disease. We have built a solution capable of classifying breast cancer in order to help doctors, especially in Morocco, to better diagnose, detect, and quickly identify patients attacked by cancer to speed up the workflow. This work consists to produce a comparative study between the Logistic Regression algorithm and 10 machine learning algorithms using a breast cancer dataset. The results of the experimentation show that the Logistic Regression was achieved 94.74% of accuracy, which proves the capacity, performance, and efficiency of this algorithm.

Keywords—Breast Cancer, Logistic Regression, Machine Learning, Accuracy

I. INTRODUCTION

Cancer disease is a major health problem requiring a comprehensive [1] care policy. Breast cancer is the number one cancer in women in the world, affecting 2.1 million women each year, and specifically targeting women between the ages of 40 and 70 [2]. This disease is considered to be an important subject of public health, however, most cases are preventable if detected early with better prediction [3]. The situation in Morocco is very worrying, breast cancer is not only the most frequent cancer in Moroccan women (36.1% for breast cancer) but causes a significant number of deaths due to error or diagnostic delay [4]. For a better medical examination of cancer, there are several techniques and tools for forecasting and decision support, among them we find Machine Learning (ML) algorithms.

ML is a way to model phenomena to make strategic decisions. It uses a variety of algorithms that iteratively learn from data to improve, describe inputs, and predict outcomes. ML problems can be differentiated into two categories, supervised and unsupervised [5]. It is a technology that can revolutionize the healthcare industry. ML algorithms can comprehensively assess cancer disease, achieving very high performance. In Morocco, the use of new technologies is very low in the health sector [6], which negatively affects the

detection of cancer in general and breast cancer in particular. Morocco needs a national electronic health system that will allow doctors to better analyze and diagnose all diseases [7].

In this article, we create a comparative study between Logistic Regression and 10 ML algorithms using a database containing information on patients affected by breast cancer and examining the index of accuracy.

II. LOGISTIC REGRESSION

Logistic regression, logit model or binomial regression is a predictive technique [8]. It is a supervised classification algorithm popular in ML. LR a statistical approach that can be used to evaluate and characterize the relationships between a response variable of binary type, Y , and one, or more, explanatory variables, which can be categorical or continuous numeric type, X [9]. The formulation of the model is:

Either the dataset D made up of n pairs (x, y) , with the description of an individual according to d descriptors, in the form of a real vector of size d , and y the membership class of this individual among 2 possible classes:

$$D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\} \quad \forall (x_i, y_i) \quad \leq i \leq n : x_i \in \mathbb{R}^d, y_i \in \{0; 1\} \quad (1)$$

Logistic regression modeling the conditional probability $p(Y = 1|X = x)$ as follows:

$$p(Y = 1|X = x) = \frac{1}{1 + e^{-f(x)}} \quad (2)$$

We recognize the sigmoid function, S , also called logistic function:

$$S: \mathbb{R} \rightarrow [0; 1] \quad (3)$$
$$\alpha \mapsto \frac{1}{1 + e^{-\alpha}}$$

With $\alpha = f(x)$, which is a linear function of x :

$$(4)$$

$$f: \mathbb{R}^d \rightarrow \mathbb{R}$$

$$x \mapsto \omega_0 + \omega x + \omega_2 x_2 + \dots + \omega_d x_d$$

Logistic regression is a powerful multivariate analysis method. It makes it possible to control for possible confusion bias. Its use is made easy by the use of statistical software [10].

III. MATERIALS AND METHODS

In this article, first of all, we developed a comparative study between the Logistic regression algorithm and 10 ML algorithms using an open access Data Base "Breast Cancer Wisconsin (Diagnostic) Data Set" [11]. An ML algorithm is a method by which models appropriate for the application will have learned from the example data [12] There are many algorithms, we will choose a particular type of algorithm depending on the type of task we want to perform and the type of data available. Basically, in this study we have to use 30 different columns and 569 rows of a 144.5KB CSV file. We will predict the health status of patients, cases affected by breast cancer or not. To do this, we used the following list of algorithms: Logistic Regression (LR), Gaussian Naive Bayes (GNB), K Nearest Neighbors (KNN), Random Forest (RF), Decision Tree (DT), Linear Support Vector Classifier (SVC(linear)), Stochastic Gradient Descent (SGD), Quadratic Discriminant Analysis (QDA), Linear Discriminant Analysis (LDA), Neural Network (NN), and Extra Tree (TE). These algorithms are the most frequently used for this kind of problem. The choice of the latter comes after a long study and a very precise filtering. To analyze the data, we used the Python 3.7 programming language. The overall process of the experiment is illustrated in Figure 1.

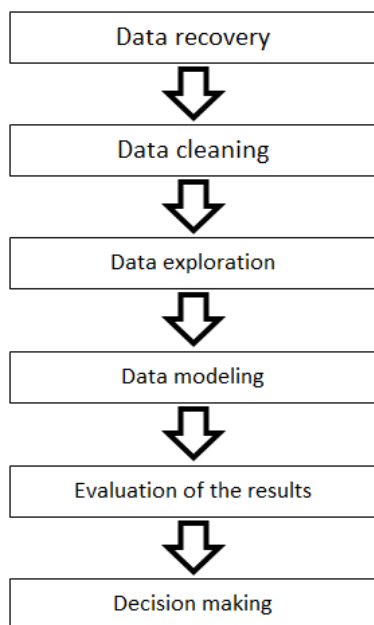


Fig. 1. The overall test process.

IV. RESULTS

A. Data Explore

Regarding the database of our study, after importing the data, we noticed that the first column is the ID of type int64, the second is the diagnosis which means the health status of the patient after a medical examination: M (a malignant breast) and B (a benign breast) of object type, the last is NaN and the rest are different characteristics (30 columns) of type float64 which are: radius_mean, texture_mean, perimeter_mean, area_mean, smoothness_mean, compactness_mean, concavity_mean, concave points_mean, symmetry_mean, fractal_dimension_mean, radius_se, texture_se, perimeter_se, area_se, smoothness_se, compactness_se, concavity_se, concave points_se, symmetry_se, fractal_dimension_se, radius_worst, texture_worst, perimeter_worst, area_worst, smoothness_worst, compactness_worst, concavity_worst, concave points_worst, symmetry_worst, and fractal_dimension_worst. Then we cleaned up the data, we removed the ID which represents the patient identifier, and Unnamed (NaN) which represents nothing.

First, we want to know the number of patients who are benign and malignant in the database. According to the results obtained, there are 357 benign cases with a percentage of 62.8% and 212 malignant cases with a percentage of 37.2%. The figure 2 corresponds to the results of the distribution of cases B/M.

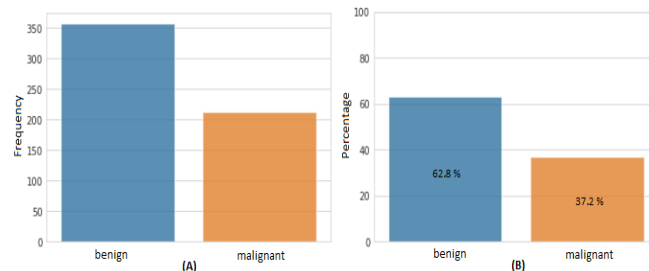


Fig. 2. Histogram of B / M cases. (A): frequency of diagnostic results; (B): percentage of diagnostic results.

Second, we have represented the different properties of the database. For this reason, we used multiple visualization using the density plot, we deployed both the data distribution and the degree of separation of the two sets of malignant and mild cases, in each entity direction. Figures 3, 4, and 5 visualize the characteristics of the tumor for positive and negative diagnosis.

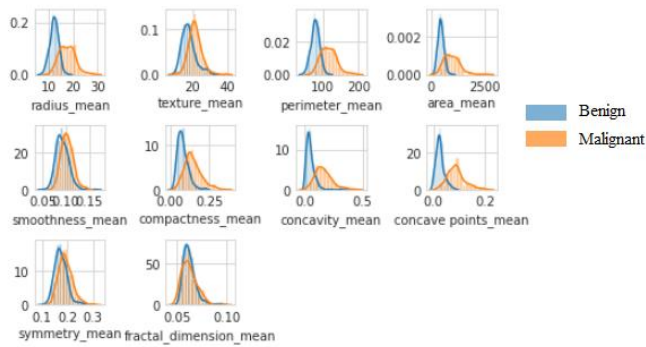


Fig. 3. Distribution of benign and malignant tumors of the \"_mean\" data group.

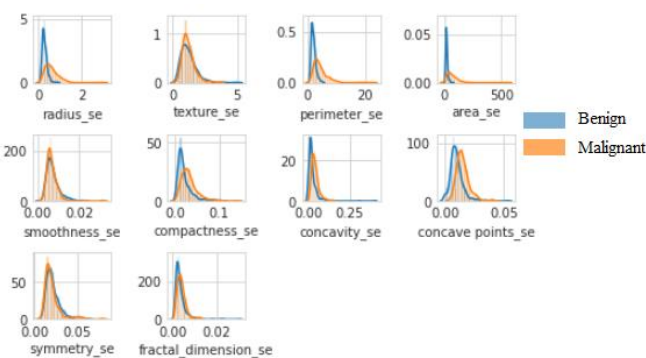


Fig. 4. Distribution of benign and malignant tumors of the \"_se\" data group.

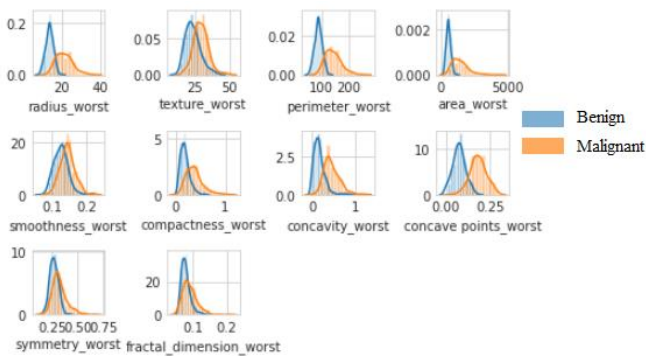


Fig. 5. Distribution of benign and malignant tumors of the \"_worst\" data group.

The values of radius_, perimeter_, and area_ can be used in the classification of cancer. Higher values of these parameters tend to show a correlation with malignant tumors. Values of smoothness_, symmetry_, or factual_ do not show particular preference of one diagnosis over the other. In any of the density plots there are no noticeable large outliers that warrant additional cleaning. Third, to perform a correct and very successful analysis, it is necessary to determine the correlation matrix between the different instances of the database. The latter designates the proximity between two variables and establishes a linear relationship between them. This correlation is represented in the figure 6.

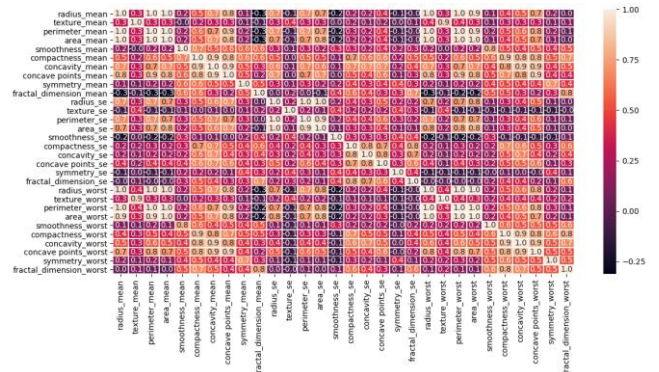


Fig. 6. The correlation between all features of the database.

From the correlation matrix shown in figure 6, we can deduce that there is a linear relationship between several variables. In addition, there are several parameters which are negatively correlated and others are not correlated. Figures 7, 8, and 9 show an example for each correlation case.

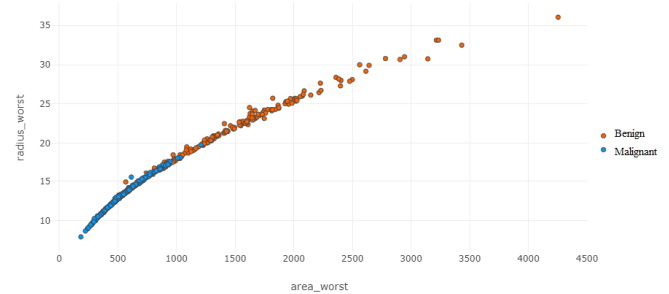


Fig. 7. Example of positive correlation.

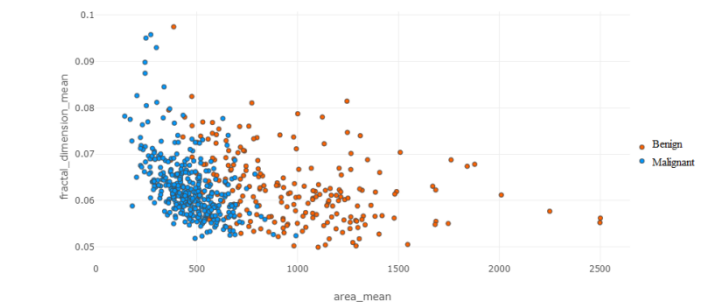


Fig. 8. Example of negative correlation.

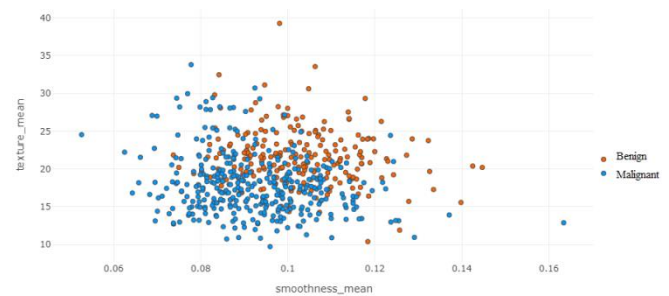


Fig. 9. Example uncorrelated attributes.

The highest correlations are between:

perimeter_mean and radius_worst, perimeter_mean and area_mean, radius_mean and area_mean, radius_mean and perimeter_mean, radius_worst and area_mean, radius_worst and radius_mean, area_worst and radius_worst, area_worst and are _mean, perimeter_worst and radius_mean, perimeter_worst and perimeter_mean, perimeter _worst and area_mean, perimeter_worst and radius_worst, perimeter_worst and area_worst, perimeter_se and radius_se, and area_se and radius_se.

B. Experiment

For a better classification of breast cancer, the database has been divided into two phases: the first is the training phase with a percentage of 80%, and the second is the test phase with a percentage. by 20%. For our experiment we only used the best 6 characteristics, which are: area _mean, perimeter _mean, radius _mean, radius _worst, perimeter _worst, area _worst. We calculated the performance index of the Logistic Regression and 10 ML algorithms for our experiment by measuring the accuracy. Accuracy is the percentage of correct predictions for the test data. The figure 10 visualizes the variation of the score of the accuracy of the LR and 10 algorithms for 6 characteristics.

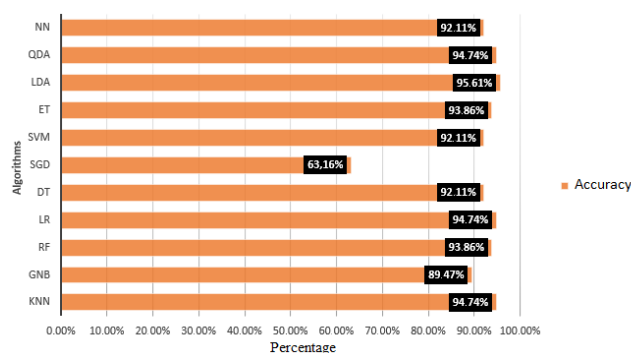


Fig. 10. The index of the accuracy of the Logistic Regression and 10 ML algorithms for 6 attributes.

The results of the first experiment show that most of the algorithms tested achieved high performance over 92% except GNB which scored 89.47% and SGD which achieved only 63.16%. What is important that the LR is scored an important ranking 94.74% of the accuracy for the study experience.

V. DISCUSSION

During this study, we established a classification of breast cancer using ML algorithms. Regarding this classification, we have launched a comparative study between the Logistic Regression and 10 ML algorithms and we can deduce that LR is a reliable, powerful and efficient algorithm. It kept very high results for our experience exercised for this problematic, 94.74% of the accuracy. This method achieves a very advanced capability of predicting breast cancer in the database. in another work a comparison of six ML algorithms: LR, KNN, GRU-SVM, Multilayer Perceptron, Softmax regression, and SVM. The results of this comparison show that all the ML algorithms tested performed very well (all the algorithms exceeded 90% test

precision) on the [13] classification task. In similar work [14], the percentage of accuracy of neural classifiers reaches almost 98% for the diagnosis of breast cancer. Also, in another comparison, the SVM and KNN classification techniques achieved an accuracy of 98.57% and 97.14% [15]. From this comparison of different experiments performed in the same database, it can be deduced that the LR algorithm has proven its prediction capacity and performance, and that the small differences that appear in the percentage of the accuracy between the different work is due to the way of selecting the characteristics which represent a strong and positive relationship.

VI. CONCLUSION

In this paper, we have developed a comparison between the Logistic Regression algorithm and 10 ML algorithms. The implementation of this list of algorithms was performed using the University of Wisconsin Breast Cancer (Diagnosis) Database. The aim is to classify patients who may be carriers of this disease between mild and malignant using diagnostic characteristics. In the experiment we chose 6 attributes. Firstly, the results show that the accuracy of all the algorithms tested was very important except that the SGD algorithm which achieved a low success rate. Secondly we saw that the LR algorithm scored important results for and that it achieved very high performance. We can conclude that the LR algorithm has a very advanced capacity for classifying patients between benign and malignant in relation to cancer disease.

REFERENCES

- [1] Bekkali, R. LUTTE CONTRE LE CANCER AU MAROC L'APPORT DE LA FONDATION LALLA SALMA. International Journal of Medicine and Surgery, vol. 4, no 1, p. 55, 2017.
- [2] World Health Organization: Breast Cancer. <https://www.who.int/cancer/prevention/diagnosis-screening/breast-cancer/en/>
- [3] International Agency for Research on Cancer. World cancer report: World Health Organization. IARC Press, 2003.
- [4] Programme de détection précoce. http://www.contrelecaner.ma/en/detection_precoce_action.
- [5] Hurwitz, J., & Kirsch, D. Machine learning for dummies. IBM Limited Edition, 75, 2018.
- [6] Benbrahim, H., Hachimi, H., & Amine, A. Survey on the Use of Health Information Systems in Morocco: Current Situation, Obstacles and Proposed Solutions. In International Conference on Advanced Intelligent Systems for Sustainable Development, Springer, Cham, 2018. p. 197-204, 2018.
- [7] Benbrahim, H., Hachimi, H., & Amine, A. Moroccan Electronic Health Record System. In Proceedings of the International Conference on Industrial Engineering and Operations Management Paris, France, 2018.
- [8] Peng, C. Y. J., Lee, K. L., & Ingersoll, G. M. An introduction to logistic regression analysis and reporting. The journal of educational research, vol. 96, no 1, p. 3-14, 2002.
- [9] Hoffman, J. I. Biostatistics for medical and biomedical practitioners. Academic press. 2015.
- [10] El Sanharawi, M., & Naudet, F. Comprendre la régression logistique. Journal français d'ophtalmologie, vol. 36, no 8, p. 710-715, 2013.
- [11] Dua, D., & Graff, C. UCI Machine Learning Repository [<http://archive.ics.uci.edu/ml/>]. Irvine, CA: University of California,

- School of Information and Computer Science, zuletzt abgerufen am: 14.09. 2019.
- [12] Lahmiri, S. On simulation performance of feedforward and NARX networks under different numerical training algorithms. In Handbook of research on computational simulation and modeling in engineering, IGI Global, p. 171-183, 2016.
- [13] Agarap, A. F. M.: On Breast Cancer Detection: An Application of Machine Learning Algorithms on the Wisconsin Diagnostic Dataset. In: Proceedings of the 2nd International Conference on Machine Learning and Soft Computing. ACM, Phu Quoc Island, Viet Nam, p. 5-9, 2018. doi: 10.1145/3184066.3184080
- [14] Anagnostopoulos, I., Anagnostopoulos, c., Rouskas, A., Kormentzas, G., Vergados, D.: The Wisconsin Breast Cancer Problem: Diagnosis and DFS time prognosis using probabilistic and generalised regression neural classifiers. Oncology Reports, Special Issue Computational Analysis and Decision Support Systems in Oncology. vol. 15, no 4, p. 975-981, 2006.
- [15] Islam, M. M., Iqbal, H., Haque, M. R., Hasan, M. K.: Prediction of Breast Cancer Using Support Vector Machine and K-Nearest Neighbors. In: IEEE Region 10 Humanitarian Technology Conference (R10-HTC). IEEE Press, Dhaka, Bangladesh, p. 226-229, 2017. doi: 10.1109/R10-HTC.2017.8288944

Classification of credits in the german market via SVM: Attempt at modeling

AZIZ CHIHAB

Department of mathematics/ University
Ibn Tofail, Kenitra, Morocco

aziz.sma.1994@gmail.com

Mohammed Kaicer

Laboratory of Analysis, Geometry and
Applications

mohammed.kaicer@uit.ac.ma

Keywords— Machine Learning, Deep learning, SVM, Credit market, prediction.

I. SUMMARY:

The work is divided into two main parts. The first part is the theoretical part in which we have defined basic notions of artificial intelligence, machine learning and these supervised and unsupervised learning algorithms and then the SVM algorithm, which is based precisely to integrate complexity control into the estimation; that is, the number of parameters that is associated in this case with the number of Supports vectors, and for this we chose to work in the practical part by the support algorithm for machine vectors, and we obtained better results in terms of the accuracy of the prediction which is increased up to 75%.

II. INTRODUCTION ET METHODOLOGY

A. Introduction

In the practical part of this work, we followed the following steps. First, we collected the data from the Learning UCI machine repository, and then in the second stage we did a data pre-processing (Cleaning and Encoding). Then to develop a model of machine Learning the Data Set has been divided into two parts: Train Set and Test Set. with the Train Set data we have developed a transformation function called a transformer which allows us to process us to drive an estimator. Once this step is completed, the transformer and estimator are used to transform the Test Set data to obtain a new prediction. Then we test the three models(SVM model with "linear" kernel, SVM model with "polynomial" kernel and SVM model with "RBF" kernel)

B. Methodology

In our work we used:

- support vecteur machine(SVM) algorithm .
- SVM with "linear" kernel.
- SVM with "polynomial" kernel .
- SVM with "RBF" kernel .

- k-nearest neighbor (k-NN) et k-means .
- Cross-validation technique.
- Supervised learning algorithms.
- Unsupervised learning algorithms.

III. RESULTS AND INTERPRETATIONS:

A. Numerical results

With the cross-validation technique that gives the best parameters that give the right estimator for each model the following results are obtained:

- SVM with "linear" kernel:
 $svm.SVC(kernel='linear', C=0.1)$.
- SVM with "polynomial" kernel:
 $svm.SVC(kernel='poly', degree=1)$.
- SVM with "RBF" kernel :
 $svm.SVC(kernel='rbf', C=1000, gamma=0.000001)$.

Kernel	RÉSULTATS	Training		Testing	
		1 : BONE	2 : MAUVAIS	1 : BONE	2 : MAUVAIS
Linéaire (C=0.1)	Précision	0.75	0.63	0.75	0.77
	Recall	0.93	0.28	0.96	0.31
	F1_score	0.83	0.39	0.84	0.44
	support	632	268	68	32
Polynômial (Degré= 1)	Précision	0.74	0.69	0.73	0.88
	Recall	0.96	0.20	0.99	0.22
	F1_score	0.84	0.31	0.84	0.35
	support	632	268	68	32
RBF (C=1000 ET gamma=0,000001)	Précision	0.73	0.75	0.70	0.75
	Recall	0.98	0.15	0.99	0.09
	F1_score	0.84	0.25	0.82	0.17
	support	632	268	68	32

Table 1: Summary of Results Achieved With Cross-Validation Technique

B. Graphic results:

temps d execution NOYAU linear : 198.1477530002594

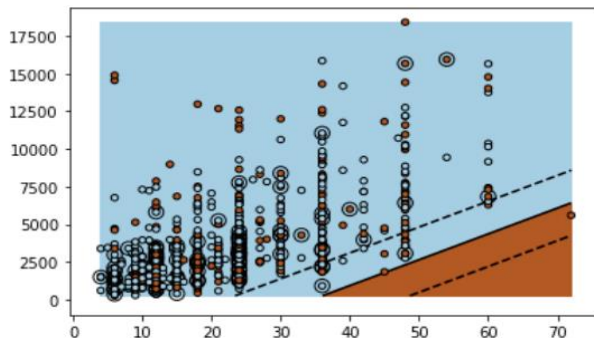


Figure 1:Graph in the case of the "linear" kernel

temps d execution noyau poly : 258.98491287231445

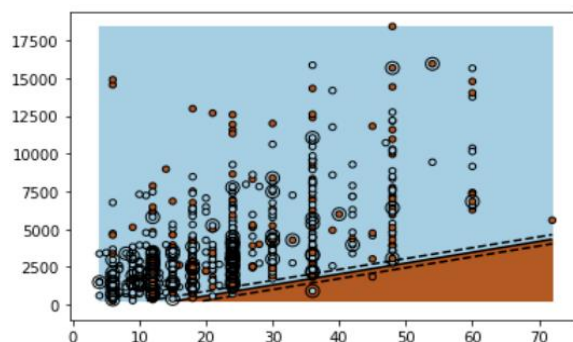


Figure 2:Graph in the case of the "polynomial"kernel

temps d execution noyau RBF : 12.293589353561401

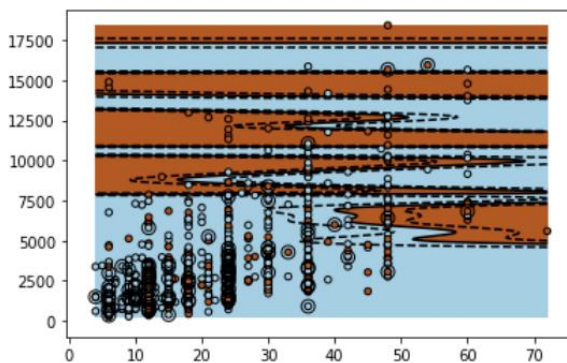


Figure 3:Graph in the case of the "RBF" kernel

This brief is an introduction to scientific research in the field of artificial intelligence and to develop models of machine learning, which being able to classify credit in the market in general one especially in the German market. This work is designed to solve non-linear problems using a non-linear kernel SVM algorithm, which allows a good visualization of the classification and data classes of the best results in terms of production accuracy, In our work we got 75% accuracy.

C. Interpretation

The linear model has greater precision in the classification of unsuccessful treatments ("1", 75%). In particular, the 96% recall suggests that almost none of the unsuccessful treatments are missing from the entire test sample. However, the model is almost unable to identify successful cases ("2"). It captures only about 31% of potential candidates and includes many false positives. Moreover, the proportion of true positive classifications that are truly positive is very, very low (45%).

The polynomial model has a higher precision in the classification of unsuccessful treatments ("1", 73%). In particular, the 99% recall suggests that almost none of the unsuccessful treatments are missing from the entire test sample. However, the model is almost unable to identify successful cases ("2"). It captures only about 22% of potential candidates and includes many false positives. Moreover, the proportion of true positive classifications that are truly positive is very, very low (35%).

The RBF model has a higher precision in the classification of unsuccessful treatments ("1", 70%). In particular, the 99% recall suggests that almost none of the unsuccessful treatments are missing from the entire test sample. However, the model is almost unable to identify successful cases ("2"). It captures only about 9% of potential candidates and includes many false positives. Moreover, the proportion of true positive classifications that are truly positive is very, very low (17%).

By comparing the three classification ratios of these three models it was concluded that the optimal model for classifying our database is the linear kernel which has an accuracy of 77% and f1-score 44% for the second class "bad credit". And among my future work perspective it has improved the performance of this modeled so as to increase the f1-score and precision.

IV. REFERENCE

- [1] Nilsson, Nils J. "John McCarthy." National Academy of Sciences (2012): 1-27.
- [2] Abe, Shigeo. Support vector machines for pattern classification. Vol. 2. London: Springer, 2005.
- [3] McGarry, Thomas J., et al. "Classification system for partial edentulism." Journal of Prosthodontics 11.3 (2002): 181-193.
- [4] Géron, Aurélien. Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts,

- tools, and techniques to build intelligent systems. O'Reilly Media, 2019.
- [5] Berrani, Sid-Ahmed, Laurent Amsaleg, and Patrick Gros. "Recherche par similarités dans les bases de données multidimensionnelles: panorama des techniques d'indexation." *INGENIERIE DES SYSTEMS D INFORMATION 7.5/6* (2002): 9-44.
- [6] Mack, Yue-Pok. "Local properties of k-NN regression estimates." *SIAM Journal on Algebraic Discrete Methods* 2.3 (1981): 311-323.
- [7] Zouhal, Lalla Meriem, and Thierry Denoeux. "An evidence-theoretic k-NN rule with parameter optimization." *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 28.2 (1998): 263-271.
- [8] Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media, 2009.
- [9] Dangeti, Pratap. *Statistics for machine learning*. Packt Publishing Ltd, 2017.
- [10] Theobald, Oliver. *Machine learning for absolute beginners: a plain English introduction*. Scatterplot Press, 2017.
- [11] Burkov, A. "The Hundred-Page Machine Learning Book by Andriy Burkov." *Expert Systems* 5.2 (2019): 132-150.
- [12] Azencott, Chloé-Agathe. *Introduction au machine learning*. Dunod, 2019.
- [13] Hilali, Hassane. *Application de la classification textuelle pour l'extraction des règles d'association maximales*. Diss. Université du Québec à Trois-Rivières, 2009..
- [14] Descôteaux, Steve. *Les règles d'association maximale au service de l'interprétation des résultats de la classification*. Diss. Université du Québec à Trois-Rivières, 2014.

Feature extraction and face recognition using auto associative memories

Mohamed Gouskir

[#] Laboratory of of Sustainable Development, Sultan Moulay Slimane University Beni Mellal, Morocco

m.gouskir@usms.ma

Abstract— Face recognition is an important tool for verification and identification of an individual. It can be of significant value in safety and commerce applications. Currently there are a numerous method to recognize and identify a person in an image. These methods can be divided into two categories: geometric methods and global methods, the performance of these methods depends on the precision with which relevant information is extracted from the face (as some of the face and eyes, nose, mouth ...).

This paper presents a face recognition approach based on auto associative memory which is an advantage of neural networks in the field of pattern recognition.

Keywords— Face detection, Face recognition, Associative memory, Auto associative memory and Legendre moments.

Introduction

The face detection problem has been treated by several different methods and technology with the aim to find the best extract the image face that allows a better recognition rate.

According Hjelmås and Low [1], methods of face detection can be classified into global approach and local approach in which the face is analyzed as a whole, locate and bring together the different elements of the face.

Face recognition is based on our ability to recognize people; she has no great difficulties for a human being why he thinks a computer system that will play this role. There are many approaches to remedy this problem. We can also divide the facial recognition methods in three categories in general [3]: global methods (the full face as a source of information) among these techniques Faces clean, DCT Networks is cited neurons, LDA [2]. Local methods (the Eigen Object (EO), the HMM (Hidden Markov Models)) [4] and hybrid methods this technique is to combine several methods to solve the identification problem.

The most commonly used in object recognition techniques in the last year are: neural networks [2] principal component analysis (PCA) [5], the independent component analysis (ICA), support vector machine (SVM) [6], neural networks Learning Vector Quantization LVQ-RN [4].

One of the Most Important Application of neural networks is the associative memories All which area exploratory tool in our work.

Any process automatic face recognition should take into account several factors that contribute to the complexity of its task, because the face is a dynamic entity that is constantly

changing under the influence of several factors. Fig. 1 illustrates the principle itself.

This work is organized as follows: in Section 1 we will treat the acquisition phase and pretreatment face, Section 2 describes the feature extraction method by Legendre moments and in Section 3 our methods face recognition by auto associative memories.

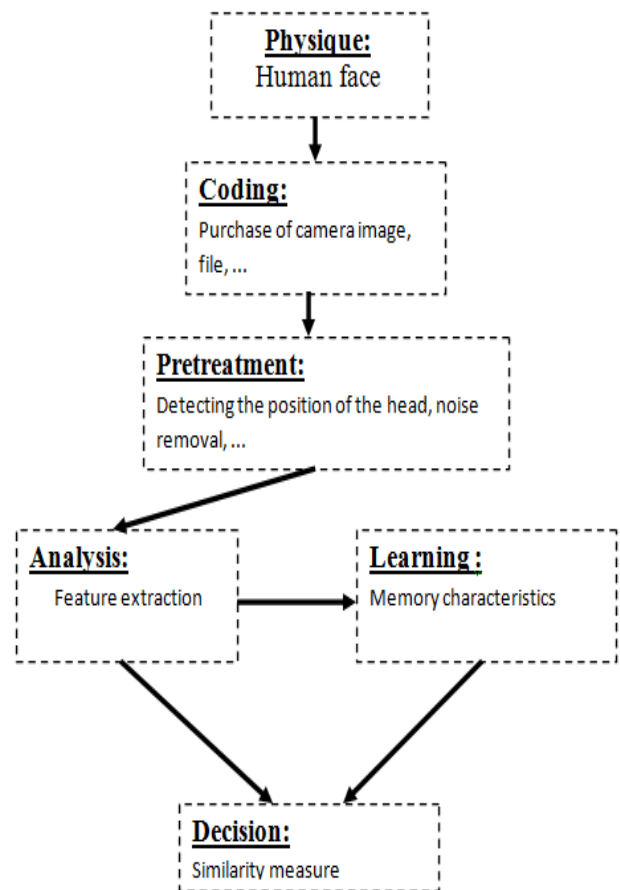


Fig. 1: The process of face recognition.

I. PRETREATMENT

After the image acquisition (extract the real world for a two-dimensional representation of objects in 2D), this can be static (Camera, Scanner, etc...) or dynamic (Camera, Web Cam), in this case we will be a movie. At this level we will have a raw image.

The raw image can be affected by various factors causing its deterioration, it can be noisy, to overcome these problems, there are several treatment methods and image enhancement, such as: standardization, histogram, equalization, etc.

This step is very important because it should influence the next steps in my memory we've used this three pre-treatment methods: normalization, binarization and filtering.

1. NORMALIZATION

The normalization consists of two processes: geometric and photometric. Geometric normalization is necessary because the size of the face within the image acquired may vary depending on the distance between the acquisition module and the person (Fig.2).

The photometric normalization step attempts to eliminate or reduce the effects of the illumination of image. In some cases, the step of photometric normalization can be applied before, or before and after the step of geometric normalization. It can also be applied during the detection phase.

Normalization of the image is scaled to a fixed size for all images in database.

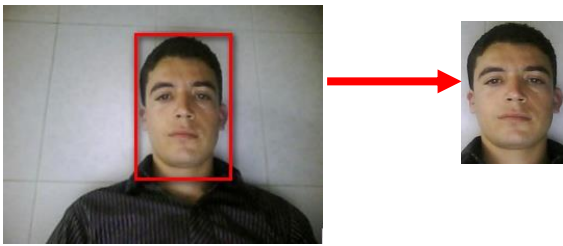


Fig. 2: Geometric form of the image right and photometric normalization left

2. BINARIZATION

A method for binarization is required to transform the image in black and white in order to facilitate the next steps including the extraction of features (Fig. 3).

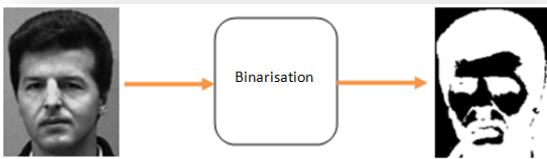


Fig. 3: Binarization of face image

II. FEATURE EXTRACTION

In pattern recognition and image processing, feature extraction is a particular form of reduced dimensionality. When the input of an algorithm is too important to be treated, the percentage of redundancy is very important causing a disturbance and images can't be treated well (lots of data but not much information), then the input data will be transformed into a reduced representation of all functions (also called feature vector). Transformation of input data is called feature extraction.

In this step we used the Legendre moments of our image to extract face features the most relevant.

The moments of Legendre, were first introduced by Teague [4]. They belong to the class of orthogonal moments and have been used in several applications of pattern recognition [17].

The two-dimensional Legendre moments of order (p + q), image intensity function f(x, y) are defined as [18]:

$$L_{pq} = \frac{(2p+1)(2q+1)}{4} \int_{-1}^1 \int_{-1}^1 P_p(x) P_q(y) f(x, y) dx dy$$

With p, q = 0, 1, 2, 3, 4 ...

Or $P_p(x)$ $P_q(y)$ are called Legendre polynomials defined as follows:

$$P(x) = \sum_{k=0}^p (-1)^{\frac{p-k}{2}} \frac{1}{2} \frac{(p+k)! x^k}{\left(\frac{p-k}{2}\right)! \left(\frac{p+k}{2}\right)! k!}$$

For p-k = pair, p, q = 0, 1, 2, 3 ... $x \in [-1, 1]$ degree $(P_p(x)) = p$.

Finally the moment of Legendre order (p + q) of a digital image f(x, y) is:

$$L_{pq} = \frac{(2p+1)(2q+1)}{4} \sum_x \sum_y P_p(x) P_q(y) f(x, y)$$

The recurrence relation of Legendre polynomials, $P_p(x)$ is given as follows [17]:

$$P(x) = \frac{(2p-1)xP_{-1}(x) - (p-1)P_{-2}(x)}{p}$$

Where $P_0(x) = 1$, $P_1(x) = x$ and $p > 1$. Since the region of definition of Legendre polynomials is within [-1, 1], a square image of N x N pixels with intensity function f(i, j), $0 \leq i, j \leq (N-1)$, across the region $-1 < x, y < 1$.

In the result, equation (1) can now be expressed as discrete as follows [17]:

$$L_{pq} = \lambda_{pq} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P_p(x_i) P_q(y_j) f(i, j)$$

Where the normalization constant is:

$$\lambda_{pq} = \frac{(2p+1)(2q+1)}{N^2}$$

x_i and y_j denote the normalized pixel coordinates in the interval [-1, 1], which are given by [17]:

$$x_i = \frac{2i}{N-1} - 1$$

And

$$y_j = \frac{2j}{N-1} - 1$$

The Legendre polynomials are characterized by the property of orthogonality, in fact:

$$\int_{-1}^1 P_p(x)P_q(x)dx = \frac{2}{(2p+1)}\delta_{pq}$$

δ_{pq} is the Kronecker symbol checking the following property:

$$\delta_{pq} = \begin{cases} 1 & \text{if } p = q \\ 0 & \text{ailleurs} \end{cases}$$

We are a square matrix whose elements represent the Legendre moments of our face image, we represent this matrix must enter an auto associative memory that must operate in the learning stage.

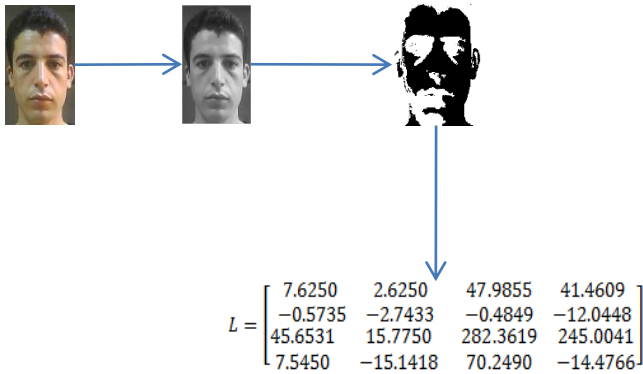


Fig. 4: Feature extraction by Legendre

III. Learning using Auto associative memories

Associative memory in the designation, the term memory refers to the storage function of these networks, and the term the associative addressing mode, since it must provide information to the network for one that is stored: c 'is a memory addressable by its contents.

With auto-associative memories, provide some of the information stored for the stored information (e.g. part of the face to get the entire face).

The applications of these memories are essentially the reconstruction of signals and their recognition.

Each neuron is connected to all others, and all entries. The transfer function is usually the identity. The evaluation can be done synchronously or asynchronously.

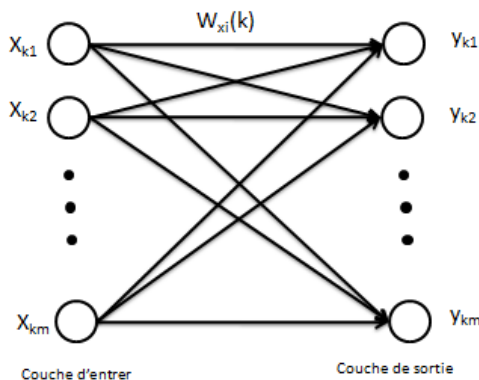


Fig. 5: Associative Memory

With

X_k : Input Vector.

$X_k = [x_{k1}, x_{k2}, \dots, x_{km}]$.

Y_k : output vector.

$Y_k = [y_{k1}, y_{k2}, \dots, y_{km}]$.

$W_{ki}(k)$: weight of memory.

And

k: Number of stimuli

m: number of neurons

For an auto associative memory input vector equal to the output $Y_k = X_k$.

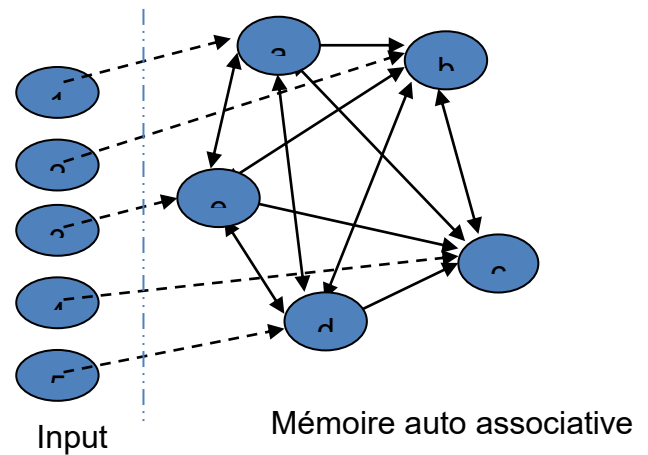


Fig. 6: Auto associative memory model

Learning by auto associative memory type is supervised. That is to say that the training set consists of pairs of vectors associated with input and output, and is based on the Hebb rule.

$$W = L * L^T$$

L: The Matrix features

On leaving the auto associative memory must be obtained another matrix O calculated as follows:

$$O = W * L$$

With an error:

$$e = \sum |t - o|$$

The error is calculated using iterations as shown in the figure below:

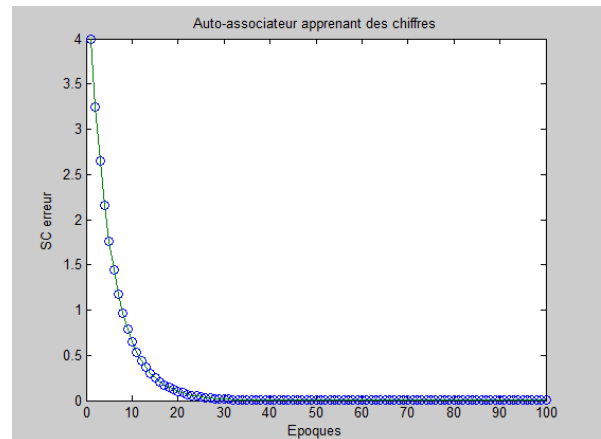


Fig. 8: The error based on iterations

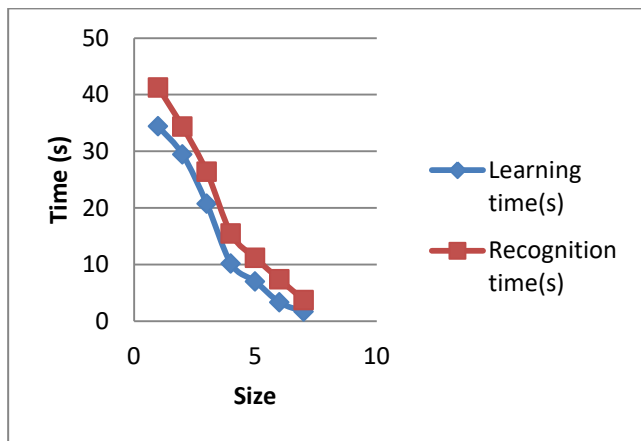
IV. EXPERIMENTAL RESULTS

In this part of the face recognition factor of time plays a very important role. Learning a large database requires significant time requires the attachment of the smallest size possible, taking into account the loss of information when the set size is small.

The table below shows the change in the learning time and recognition based on the size of the image for a face database of images.

Table 1: Time for learning and recognition based on the size

Size	Learning time(s)	Recognition time(s)
52 x 46	34,38	41,26
50 x 40	29,45	34,33
44 x 32	20,71	26,37
30 x 22	10,15	15,45
25 x 18	7,00	11,15
18 x 12	3,32	7,38
12 x 8	1,66	3,72



The decision is part of the system or on the edge of an individual belonging to all the faces or not, and if so what is its identity. So the decision is the culmination of the process. It can enhance the recognition rate (reliability) as determined by the rate of correctness of the decision.

In this step of the recognition, we repeat the same procedure as step of learning, wish means we must extract the matrix of characteristics for each test face image, and using the matrices stored auto associative memories by calculating the output matrix. Finally, we traverse the corresponding face image.

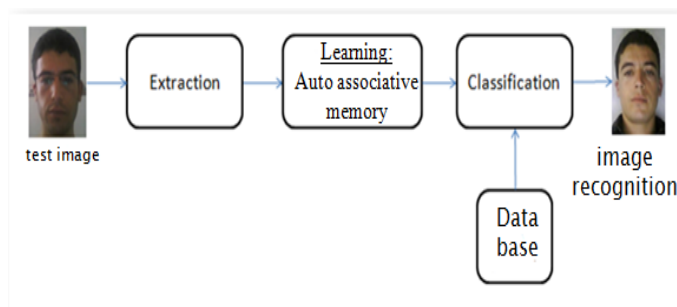


Fig. 8: The process of face recognition by associative memory self.

The table below shows the recognition rate for deferent size used:

Table 2: Recognition rate for each size

Size	Rate
52 x 46	100%
50 x 40	94%
44 x 32	90%
30 x 22	88%
25 x 18	90%
18 x 12	81.8%

The poor rate of recognition when the small size and depends on the lack of information extracted from a small face image.

Conclusion

In our present work it has led to encouraging result ant force which explains the auto associative memories in the field of face recognition. This work is based on the extraction of information from the Legendre moments of each image considered as a matrix of characteristics, which then turn on the auto associative memory models of learning and classification.

The results obtained for each image studied are quite encouraging and have shown that memories auto associative adapts well to changes in face images. However, its discriminatory ability is not very strong. In addition, the weaknesses of associative memories come from the weight calculation and estimation of activation function.

References

- [1]. E. Hjelm and B. K. Low. "Face detection: A survey", Computer Vision and Image Understanding, Vol. 83, n°3, pp. 236-274, 2001.
- [2]. F. Vermont: Localisation de visages, Lausanne Février 2005
- [3]. A LEMIEUX, système d'identification de personnes par vision numérique, université Laval, Québec décembre 2003.

- [4]. Hazim Mohamed Amir et Nabi Rachid, thème reconnaissance de visages, Universités d'Avignon et du pays du Vaucluse IUPGMT 2006 /2007.
- [5]. Chee-Way Chonga, P. Raveendranb and R. Mukundan, Translation and scale invariants of Legendre moments, Pattern Recognition 37 (2004) 119 – 129.
- [6]. BART KOSKO, Member, IEEE, Bidirectionnel Associative Memories, IEEE Transactions on systems, MAN, AND CYBERNETICS, vol. 18no.1, 1988.

How Scientific Research in Moroccan university can use Technology to ensure an efficient Governance

Younna Elhissi

Systems engineering laboratory
Sultan Moulay Slimane University, Morocco
y.elhissi@usms.ma

Abstract— In a context of E-Governance, modernization of higher education, scientific research at the Moroccan university is facing new challenges, such as :

- Orientation, enhancement and promotion of this component by responding to the needs of the socioeconomic community ;
- Development of the culture of communication and information by putting an end to practices that limit the dissemination of information ;
- Exploitation of new technologies in the management process of scientific research.

On the other hand, the technological revolution offers us several tools to support different fields and different activities, hence the establishment of an information system to ensure good governance of scientific research in the Moroccan university, is one of the good solutions that will allow the actors involved in the research activity to work independently, transparently and responsibly adopting a participative approach.

Based on moroccan experience, this work aims to show how an information system can be a good tool to insure the governance of Scientific Research in the Moroccan university.

Keywords: Moroccan University, Scientific Research, Information System, Management, Communication, Information Security.

I. INTRODUCTION

The technological revolution in the world encourages University to adopt new sharing and communication tools for meeting the needs of its socio-economic environment [1]. So it

Manuscript received October 9, 2001. (Write the date on which you submitted your paper for review.) This work was supported in part by the U.S. Department of Commerce under Grant BS123456 (sponsor and financial support acknowledgment goes here). Paper titles should be written in uppercase and lowercase letters, not all uppercase. Avoid writing long formulas with subscripts in the title; short formulas that identify the elements are fine (e.g., "Nd-Fe-B"). Do not write "(Invited)" in the title. Full names of authors are preferred in the author field, but are not required. Put a space between authors' initials.

F. A. Author is with the National Institute of Standards and Technology, Boulder, CO 80305 USA (corresponding author to provide phone: 303-555-5555; fax: 303-555-5555; e-mail: author@boulder.nist.gov).

S. B. Author, Jr., was with Rice University, Houston, TX 77005 USA. He is now with the Department of Physics, Colorado State University, Fort Collins, CO 80523 USA (e-mail: author@lamar.colostate.edu).

T. C. Author is with the Electrical Engineering Department, University of Colorado, Boulder, CO 80309 USA, on leave from the National Research Institute for Metals, Tsukuba, Japan (e-mail: author@nrim.go.jp).

permits the growth of the economy and the sustainable development of a country[2]. A good use of these tools allows universities to manage their research by the rationalization of human and material resources and the opening of a portal collaboration with the various stakeholders involved in scientific research. So the Knowledge, in real time, becomes a key factor in the success of the essential business strategy[3]. Currently the Information System "IS" is a necessary device that meets all these objectives. It is also a mean of connection to the internal and external world of the university, which contributes to the democratization of information. The logic of social development dictates the need to direct a managerial attention to knowledge management [4]. This "IS" highlights the important role of information in the organization's operating processes and systematically addresses major [5]. In this case, the value of good communication and effective exchange of information are important to participate in the management and decision making[6]. In Morocco, we opted for the establishment of an "IS" at universities to have an easy access to information related to scientific research, and this through the placement of functional blocks [7].

The establishment of an Information System as a management tool for the management and communication of research aims to:

- Allow all university stakeholders (teacher, researcher, students, administrator, and socio-economic partners) to use a computerized work environment;
- allow this category of users to access it simply and in a unified way to have, publish and share information;
- ensure interaction and rapprochement between the two academic and industrial poles
- involve all internal and external actors in the development of the region socially and economically;
- ensure a secure platform for all types of risk in order to protect and keep all information;
- involve the university's research structures in the Information System governance strategy (IT-Governance).

To do, a number of Fars topics of scientific research will be considered: in this work we present 3 of them : -management of the human resources section, -management of scientific research laboratory and management of deliverables.

II. METHODOLOGICAL FRAMEWORK

Within the framework of this work, we will focus on the MISSION project [8], a project coordinated by Hassan 1er University for a period of 3 years (2012-2015), it associates 22 university and institutional partners from Morocco and the EU. The project consists of improving the management of universities by setting up an Operational Information System Service "SSI".

This project funded by the European Commission through the Education, Audiovisual and Culture Executive Agency under the Tempus IV program. He is interested in the development of higher education in Morocco.

The MISSION project responds to Moroccan priorities in terms of governance reform, namely: Management, quality assurance and institutional and financial autonomy.

The main objective of the project is the establishment of an Operational Information System Service in Moroccan universities "MISSION". This solution is the integration of 5 items : Students affairs, Finance, Patrimony, Human Resources Management and Management of Scientific Research.



Figure 1 : Items treated by the MISSION project

Using an Integrated ERP (Enterprise Resource Planning) which is defined as being a configurable, modular and integrated IT application [9], universities can enter data and feed their departments with real-time information. This developed information system makes it possible to edit statistics, dashboards relating to key performance indicators specified in the specifications detailing the business process diagrams.

Given the importance of the scientific research « SR » module in the university and in the governance system of Moroccan universities, This work will look at this aspect of research in order to promote its activities and have data on its structures. To this end, we will focus on the platform set up by the MISSION project and which aims to manage research structures.

The aim is to show how well this platform meets the needs of the governance of scientific research at Hassan 1er University as a pilot university. We therefore present the SR-sections developed by the project, its results and its flaws relating to the research component.

III. HUMAN RESOURCE MANAGEMENT

Although this seems to be a simple rubric in a recording platform, it requires utmost importance, because it is the source of all information that will arise in the future. Good management of this database will allow us to manage the rest. This section must also be an internal resource sharing means to derive the maximum benefit and profitability. As already mentioned in the introduction IS should provide some comfort in the work and good career management. Before tackling this section we should first start by making an inventory of the existing, then raised the mechanisms of operation, so analyze before making a proposal for improvement.

This section will be devoted to group together all stakeholders in the scientific research, outline detailing the key information needed will be established. These stakeholders are called upon to identify and learn from the canvas to get a login and password, validated by the director of the research structure, these identifiers allows them to use the authorizations given by the different functions in the Information System. Once grouped, all this information will help provide real-time statistics on the entire body of research structures and stakeholders in terms of quality and quantity as well as information on the subject of scientific research topics at of the University. The figure bellow gives a global idea about the platform:

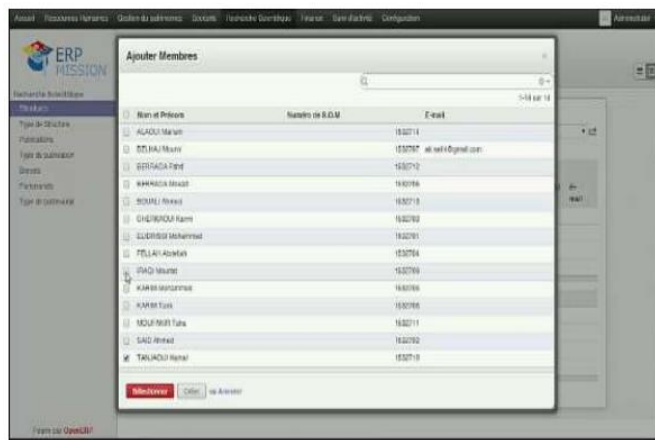


Figure 2 : Human Management Ressources platform

IV. MANAGEMENT OF SCIENTIFIC RESEARCH LABORATORY

The management of this section is a responsibility of the directors of research structures. We can find two types of information:

- Public information concerns the number of structures, their members, their research areas and even the equipment issued to each structure;

• Private information represents the status of the work of these structures. This information can only be accessed by those involved and becomes public after their completion.

Several actors are involved in scientific research either directly or indirectly we can identify the following actors:

- The Committee on Research, which brings together representatives of all stakeholders related to research
- Research units, each belonging to its corresponding section
- Research Teams

In this section the different research structures will be listed in detail: the number of stakeholders in each structure, the topics, the progress of the work of each axis or research subject and even material and equipment granted to each structure. This will allow us to have a global idea about the human and material resources in the university allocate for scientific research so permit to identify all needs and deficiency in this area for improved visibility and good management. On the other hand show the different possible opportunity to pool human and material resources for a better profitability of research in the university. This figure approach us about what will be the rubric by an IS:

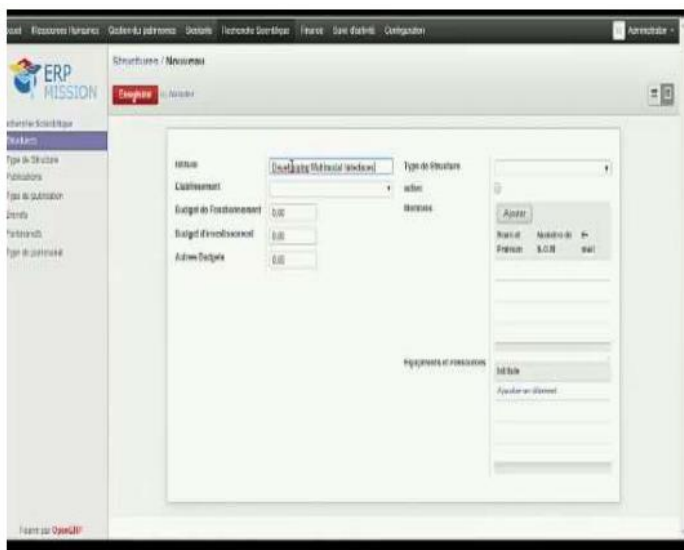


Figure 3 : Management of Research Laboratory platform

elsewhere and also communicate this type of information. The information system as a concept of management and communication can organize this part of research; it will allow each user entitled to access to provide a partial or a complete information about its research work either part of thesis, a summary of patent or a paper or also an article. This information should be published in the section so that the supervisor or other actor's structures are aware of the progress of work of its members. Until then, we were provided a portion of work, once it is finalized and validated by the responsible unit, we have the right to have and consult all of the information, it allows a publication of work, disclosure and it may cause interaction with other structures involved research areas in order to pool resources and better develop this research. The main objective is to have a general idea about the progress of his research work state, so it can generate working together with other research structures of members who are interested in the same issue or they work on the same axis searching but on other dimensions. In addition, this section will administer all activities in relations with doctoral students: the user (PhD) record any information regarding any training planned as part of his research he attended, according to an outline details this information is returned to supervisor belonging to one research structures for validation. Once it's done systematically it will give us an idea about the PhD student's activities during his years of research following the required charter.

As we can see in this figure:

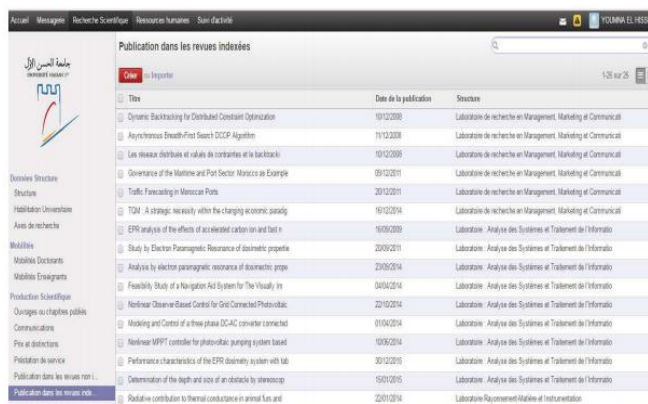


Figure 4 : Management of deliverables platform

V. DELIVERABLES MANAGEMENT: THESIS, RESEARCH PAPERS, PATENTS

Deliverables of university research structures is a criterion of scientific production, on which it is based to assess research in the university. These are the results of the work and activities of all the research structures within the university that can value them. Good management of this part evaluates the work of research to enhance the value within the university and

CONCLUSION

In Morocco, scientific research has long remained the weak link in the governance of universities, so several actions and projects have tried to resolve this problem in order to promote its activities, structure its units and communicate around its results.

These goals cannot be achieved without implementing information-based management. Available, reliable information shared and which will contribute to the

transmission of knowledge and knowledge which helps us to make the most effective decision.

Governance has emerged as an approach that can regularize research activity, optimize its resources and structure its units with the overall objective of continuously enhancing and evaluating scientific research in Moroccan universities.

The establishment of an Information System as a management tool for the management and communication of research aims to:

- Allow all university stakeholders (teacher, researcher, students, administrator, and socioeconomic partners) to use a computerized work environment;
- allow this category of users to access it easily and unified way to have, publish and share information;
- ensure interaction and rapprochement between the two academic and industrial poles
- involve all internal and external actors in the development of the region socially and economically;
- ensure a secure platform for all types of risk in order to protect and keep all information;
- involve the university's research structures in the Information System governance strategy (IT-Governance).

Hence the importance of measure the performance of this Information System to ensure an efficient governance of Scientific Research in moroccan university, next work will focus on about maturity security and governance of Information System.

I. BIBLIOGRAPHY

- [1] "Morocco's strategy for the development of Research 2025", Ministry of National Education, Higher, Education, Staff Training and Scientific Research, November 2009.
- [2] J. Urbanovic and L. Tauginiène, "Institutional Responsibility vs Individual Responsibility: Ethical Issues in the Management of Research Performance" Proceeding in the 1st World Congress of Administrative & Political Sciences (ADPOL-2012).
- [3] L. Caporiondo, "Mise en place d'un système d'information décisionnel dans une entreprise", http://eduscol.education.fr/ecogest/si/SID/@@document_wholed2
- [4] G.I. Petrova, V.M. Smokotin, A.A. Kornienko, I.A. Ershova and N.A. Kachalov, "Knowledge management as a strategy for the administration of education in the Research University", International Conference on Research Paradigms Transformation in Social Sciences, 2014.
- [5] H. Angot, « Système d'information de l'entreprise », de Boeck, Novembre 2006.
- [6] J. Taylor, "Gérer l'ingérable: La gestion de la recherche dans les universités à vocation de recherche", Politiques et gestion de l'enseignement supérieur, 2006
- [7] Project 4 Axis-II "Departmental action plan for 2013-2016", p. 36, Morocco, March 2012.

[8] E-Form Mission 2015, <https://Sites.google.com/A/Uhp.ac.ma/MISSION/Documentati on>

[9] P. Pérotin, "Les Progiciels De Gestion Intégrés, Instruments De L'intégration Organisationnelle ?" Thèse Université Montpellier II, 2004

An improved model for complex optimization problems based on the PSO algorithm

Maria ZEMZAMI
LGS Lab- Sultan Moulay Slimane
University. Beni Mellal- Morocco

maria.zemzami@gmail.com

Norelislam EL HAMI
LGS Lab- ENSA- Ibn Tofail
University. Kenitra - Morocco

norelislam@outlook.com

Nabil HMINA
LGS Lab- Sultan Moulay Slimane
University. Beni Mellal- Morocco

hmina5864@gmail.com

Abstract— recognized for many years as an effective method for solving difficult optimization problems, the Particle Swarm Optimization method however has drawbacks, the most studied are: the high computing time and the premature convergence. This paper highlights a new variant of the PSO method aimed at avoiding these two drawbacks of the method. This variant combines two approaches: the parallelization of the computational method and the organization of appropriate neighborhoods for the particles. The performance evaluation of the proposed model was made based on an experiment on a series of test functions. In the light of the analysis of the obtained results, we observe that the proposed model gives better results than those of the classic PSO in terms of quality of the solution and of the calculation time.

Keywords—*PSO, Optimization, Parallelization, neighborhood.*

I. INTRODUCTION

Optimization problems in the real world are usually complex. They not only contain the terms of constraints, of simple / multiple objectives, but their modeling is constantly evolving. Their resolution and iterative evaluation of objective functions require a long CPU time. The high computational cost for solving these problems has prompted the development of optimization algorithms with parallelization.

The Particle Swarm Optimization (PSO) algorithm is one of the most popular algorithms based on swarm intelligence, which is enriched with robustness, simplicity and global search capabilities. However, one of the main obstacles of the PSO method is its susceptibility to get trapped in local optima and; Like other evolutionary algorithms, the performance of PSO deteriorates as the size of the problem increases. Therefore, several efforts are made to improve its performance. Different scenarios are developed, either by adding new parameters to the basic algorithm [1], by hybridizing it with other meta-heuristics [2], or by proposing parallelization scenarios [3] - [6].

The basic architecture of PSO inherits a natural parallelism, and the responsiveness of fast processing machines has made this task very convenient. Therefore, parallel models based on the PSO meta-heuristic have

become very popular because parallelization offers an excellent path to improve system performance.

The present paper has five sections including the introduction. In the next section, a description the PSO algorithm is given. In section 3 our parallel approach of PSO is discussed. Section 4 describes experiment and results. The chapter finally concludes with Section 5.

II. REVIEW OF PARTICLE SWARM OPTIMIZATION

A. Concept

Particle swarm optimization (briefed as PSO) is a stochastic optimization meta-heuristic; based on population solutions, proposed in 1995 by James Kennedy (socio-psychologist) and Russel Eberhart (electrical engineer) for the resolution of the optimization problems, particularly continuous variable problem. [7].

PSO is based on the social behavior of individuals evolving in swarms, i.e., the "social interactions" between "agents" called "particles" representing a "swarm", in order to achieve a given objective in a common research space. Where each particle has a certain capacity for memorizing and processing information. In PSO, social behavior is modeled by a mathematical equation to guide particles during their process of movement [7].

The movement of a particle is influenced by three components: the inertia component, the cognitive component and the social component. Each of these components reflects part of the particle's movement equation.

1. The inertia component: the particle tends to follow its current direction of travel;
2. The cognitive component: the particle tends to move towards the best site through which it has already passed;
3. The social component: the particle tends to move towards the best site reached by its neighbors.

B. Algorithm

The basic algorithm of the PSO method proposed by [7] begins with a random initialization of the particles in their search space, by assigning them an initial position and speed. With each iteration of the algorithm, the particles move

according to the displacement equations, and the objective functions of the particles (fitness) are calculated so that the best position of all can be calculated.

The best position of each particle and the best position of the swarm are updated with each iteration. The process is repeated until the stop criterion is met (see Fig. 1).

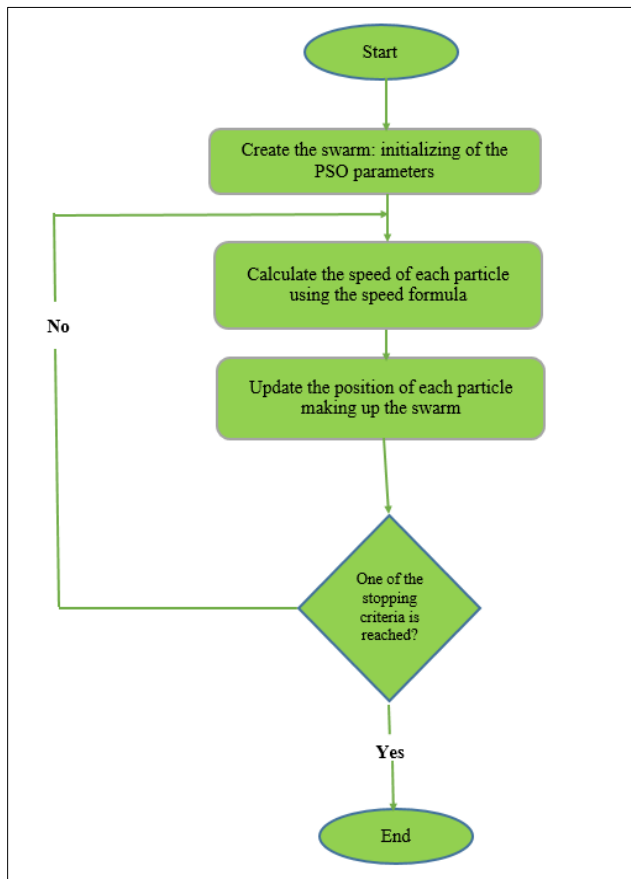


Fig. 1. Basic PSO Flowchart

C. PSO configuration

Like any meta-heuristic, PSO has a set of parameters that intervene and influence its performance. The choice of these parameters remains critical and generally depends on the problem posed [8] [9] but has a great influence on the convergence of the algorithm. Among these parameters, we can mention:

- **Number of particles:** One of the key parameters of the PSO method is the number of particles; it greatly influences the performance of the algorithm, especially in terms of calculation time, since the presence of each particle in the algorithm causes a calculation: evaluation of the position and displacement of the particle [10]. The quantity of particles allocated to the resolution of a problem depends on several parameters, namely: the dimension of the problem to be optimized (the size of the search space), the ratio between the computational capacities of the machine and the time maximum of research and especially the complexity of the optimization problem. The choice

of an adequate value for this parameter is not an easy task, since there is no rule to determine it, only a massive experiment by making many tests allows acquiring the necessary experience to the understanding of this parameter.

- **Dimension of the problem:** This parameter is a direct link with the problem to be optimized, it represents the space of research and evolution of the particles, and of course the dimension of the problem impacts the convergence of the algorithm, at the level of calculation time: dealing with a 2-dimensional problem does not take the same time as a 30-dimensional problem, and also at the level of the reliability of the results: the precision of the small-dimensional results is not the same as that linked to the problems with high dimension.
- **Particle arrangement:** Before starting the algorithm, the positions of the particles as well as their initial velocities must be initialized randomly according to a uniform law on [0..1]. This initial arrangement has an influence on the next displacement of each particle and therefore the convergence of the algorithm, especially in the case of the constitution of geographical neighborhoods. However, there is a set of automatic position generators, allowing different positions to be assigned to the whole swarm. Several studies exist in this direction [11]. The SOBOL sequence generator is one of the most efficient in the field, for a homogeneous arrangement of particles in a space of dimension n following a study made by [12] who used a certain sequence with small deviation to initialize the particles. The researchers used Halton, Sobol and Faure sequence generators to initialize the swarm. They then tested the offered variants using six standard test functions. They found that the performance of PSO with Sobol initialization is the best among all the techniques.
- **Stopping criteria:** The stopping criterion represents one of the key success of the PSO algorithm, choosing an optimal stopping criterion is not an easy task, and is not done randomly, but must be the result of a depth study of the problem and a massive experiment. This parameter differs according to the optimization problem and the constraints defined by the user, it is strongly advised to endow the algorithm with an exit gate since convergence towards the optimal solution is not guaranteed in all even if the experiments indicate the great performance of the method. As a result, several studies have been carried out in this direction [13], different propositions have taken place: the algorithm must then be executed as long as one of the convergence criteria has not been reached, this can be: the number maximum iterations; the global optimum is known a priori, one can define an "acceptable precision", the variation of the speed is close to 0. Other stopping criteria can be used according to the problem of

optimization posed and of the constraints defined by the user.

D. Neighborhood in PSO

The neighborhood constitutes the structure of the social network. The neighborhood of a particle represents with whom each of the particles will be able to communicate. There are two main types of neighborhoods:

Geographic neighborhood: this type of neighborhood represents geographic proximity, it is the most natural notion of neighborhood for particle swarms, neighbors are considered the closest particles. However, with each iteration, the new neighbors must be recalculated from a predefined distance in the search space. It is therefore a dynamic neighborhood that must be defined and updated at each iteration.

Social neighborhood: this type of neighborhood represents social proximity, neighborhoods are no longer the expression of distance but the expression of the exchange of information, they are defined at initialization and are not modified by the following. Once the network of social connections is established, there is no need to update it. Therefore, it is a static neighborhood.

- Neighborhood topology

The network of relationships between all the particles is known as the "swarm topology". The choice of a neighborhood topology is very important, several topology studies have been carried out in this regard [14], different combinations have been proposed, the most used of which are mentioned above [2]:

Ring topology (Fig. 2 (a)): each particle is connected to n particles, (usually $n = 3$), each particle tends to point towards the best in its local neighborhood.

Radius topology (Fig. 2 (b)): communication between particles is made via a central particle, only the latter adjusts its position towards the best, if there is improvement in its position, the information is then spread to its congeners.

Star topology (Fig. 2 (c)): each particle is connected to all the others, the social network is complete, i.e. The neighborhood optimum is the global optimum.

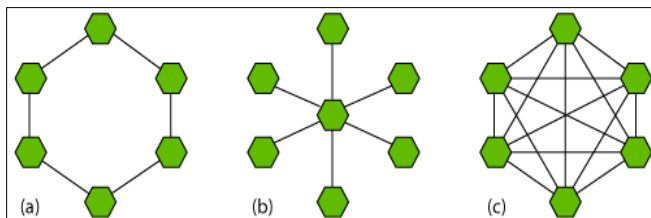


Fig. 2. Neighborhood topologies

III. PARALLEL MODELS BASED ON THE PSO ALGORITHM

The PSO algorithm quickly took its place among the most powerful methods in the family of meta-heuristics dedicated to solve complex optimization problems. Although this meta-heuristic is very popular due to its robustness, it nevertheless presents several drawbacks, the

most studied of which are the high computational time and premature convergence.

Parallelism is one of the concepts, which interested researchers in the field to remedy these drawbacks of the PSO method.

A. Related work

As mentioned previously, in the implementation of the classical algorithm of the PSO method, all the calculations are done in a sequential manner; this is where the idea of parallelization arises in order to improve the performance of the algorithm. Several scenarios are proposed, we distinguish:

In [15], the authors tested in 2006 in both synchronous and asynchronous PSO parallel algorithms for optimization of typical parameters of the wings of a transport plane. The result indicates that the asynchronous PSO algorithm performs better than the synchronous PSO in terms of parallel efficiency. [16] Implemented in 2006 an asynchronous parallel PSO algorithm for analytical and biomechanical testing problems. The experimental results obtained show that the asynchronous PSO is 3.5 times faster than the synchronous PSO algorithm.

In [17] in 2007, a parallel PSO approach named (MRPSO) based on the MapReduce parallel programming model, with the aim of addressing complex optimization problems. MRPSO has been applied to a set of test functions well known in the optimization field for their difficulty. Based on the results obtained, MRPSO can handle up to 256 processors for moderately difficult optimization problems and tolerates node failure.

In this study [18] carried out in 2010, two algorithms are developed for determining the pricing of options using particle swarm optimization. The first algorithm we developed is the Synchronous Option Valuation Algorithm Using PSO (SPSO), and the second is the Parallel Synchronous Valuation Algorithm. The pricing results obtained from these two algorithms are close compared to the classic Black-Scholes-Merton model for simple European options. A test of the synchronous parallel PSO algorithm in three architectures was performed on a shared memory machine using OpenMP, a distributed memory machine using MPI, and a homogeneous multicore architecture running MPI and OpenMP (hybrid model). The results show that the hybrid model handles the charge well when there is an increase in the number of particles in simulation while maintaining equivalent precision.

A parallel particle swarm optimization algorithm is described in [4], proposed in 2012 to solve the problem of coverage of pursuit-evasion games, where several pursuers must cooperate to cover the potential flight zone of an agile fraudster within a reasonable time. The area to be covered is complex and therefore difficult to calculate analytically. With the use of the proposed parallel PSO algorithm, maximum coverage is achieved in less time, considering the minimum number of pursuers. Calculation time can be further reduced by optimizing the fitness function according to the locality of the data. In addition, the use of a variable length of the communication data frame makes it possible to

reduce the communication time between processes when the number of processors increases (more than four in the test example). The simulation results show a comparison between the acceleration, the computation time before and after the optimization of the fitness function and the communication time between fixed and variable data frames. The positions and orientations of the prosecutors are also presented to show the effectiveness of the proposed parallel algorithm.

In [5], the authors in 2014 introduce several parallel functional skeletons, which, in a sequential implementation of the PSO method, automatically provide the corresponding parallel implementations. They use these skeletons and report some experimental results. They find that, despite the low effort required by programmers to use these skeletons, their empirical results show that the proposed skeletons achieve reasonable acceleration speeds.

In 2015, the authors of [19], developed the problem of Constraint Satisfaction Problems CSP which occur in different domains. Several methods are used to solve them. In particular, the PSO meta-heuristic, which effectively solves CSPs by drastically reducing the computational time needed to explore the solution search space. However, PSO is excessively expensive in the face of large instances. For this work, a particular interest was brought to the problems of satisfaction of maximum constraint (Max-CSP) by proposing a new approach of resolution, which allows solving efficiently Max-CSP, even with large instances. The goal was to implement a PSO-based method using the Graphics Processing Unit (GPU) architecture as a parallel computing framework. Two parallel models are offered; the first is a GPU parallel PSO for Max-CSP (GPU-PSO) and the second is a GPU distributed PSO for Max-CSP (GPU-DPSO). The experimental results show the efficiency of the two proposed approaches and their ability to exploit the GPU architecture.

In [6] two parallel strategies are proposed in 2019, based on several swarms to solve multi-objective optimization problems. The multiple swarms co-evolving in parallel and interacting through migration. Different policies for triggering migration are proposed and evaluated. An in-depth experimental evaluation of the algorithms is presented, as well as a study of the impact of the proposed methods on the convergence and diversity of research in many scenarios of multi-objective optimization. The first strategy is based on Pareto domination and the other on decomposition. Several swarms run on independent processors and communicate in broadcast over a fully connected network. A study of the impact of using synchronous and asynchronous communication strategies for the decomposition-based approach. Experimental results have been obtained for several reference problems. The conclusion was that parallelization has a positive effect on the convergence and diversity of the optimization process for multi-objective problems. However, there is no single strategy that works best for all categories of problems. In terms of scalability, for higher numbers of goals, parallel algorithms based on decomposition always show the best results.

B. Proposed parallel model

The scenario that we have adopted in this proposed model called PN-PSO (Parallel Neighborhood PSO) allows parallelizing the calculations by launching a set of threads on batches of particles positioned in different neighborhoods. Threads, (a sort of Java processes in our experiment), run in parallel with each iteration of the algorithm.

Each thread performs the processing of an iteration of its batch of particles, and waits for the other threads to complete their processing to update the neighborhoods and start a new iteration. This scenario is repeated until a satisfactory solution is obtained "reaching the stopping criterion".

The particularity of this model consists in taking advantage of the robustness of the PSO algorithm in the choice of the right parameterization in order to create diversity in the search (in our case: the distribution of particles in the search space and our notion of neighborhood) and in information sharing to facilitate convergence. Parallel computing speeds up calculations in order to have an "optimal" solution in an optimized computing time. Fig. 3 is a representation of the proposed approach [20].

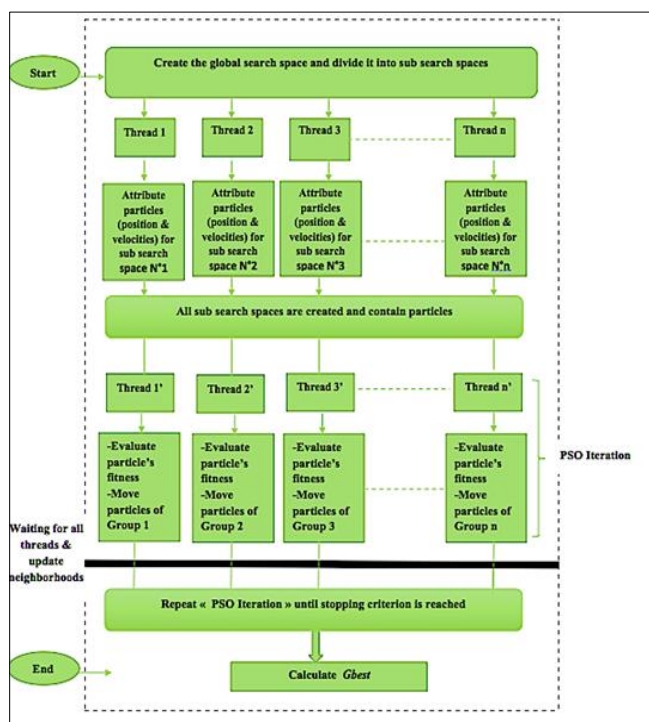


Fig. 3. The proposed Parallel PSO model

- Algorithm Framework

The main steps of the PN-PSO algorithm are as follows:

- Step 1: Create the global search space and divide it into sub spaces according to a value of step.
- Step 2: Generate randomly a set of particles by attributing their positions and speeds.
- Step 3: Each sub set of particles is attributed to one of the created thread.
- Step 4: Each thread evaluates the velocity and the position of all its own particles.
- Step 5: Wait for all threads and update neighborhoods.
- Step 6: If the stopping criterion is satisfied, stop, otherwise go to step 4.

IV. EXPERIMENT AND RESULTATS

This section deals with a description of the experiments and an analysis of the obtained results of the proposed model based on the PSO algorithm.

A. Description of experiments

The modification of the basic algorithm of PSO method for our two models involves three essential points: the notion of neighborhood, parameter adaptation, and parallel computing. These changes to the algorithm improve its performance.

Our algorithm was programmed in JAVA 1.8, and the experiments were done on a MacBook Pro OS X 10.13.15, Core i7, 16 GB machine. Threads are the technology used in Java to multitask applications. They share the same memory, as well as resources (in memory), for this, the threads risk competing and corrupting the system.

This is where concurrent programming comes in, bringing together a set of features and techniques to enable synchronization of tasks running in parallel. Java manages processes better than threads, but threads are used a lot more because they are better integrated into the Java language and less memory intensive. While Java is a robust language with many advantages (portability, inheritance, etc.), but especially to the concept of concurrent programming (parallelism and synchronization) we opted for this language in order to carry out our experiments and take advantage of the advantages of parallelism in terms of reducing the computation time, and also to make the most of the hardware resources of the machine.

To test our proposed parallel model, a set of ten test functions has been selected (see Table. 1) to evaluate the performance of the proposed model.

TABLE I. DESCRIPTION OF THE USED FUNCTIONS

Function	Range	f_{\min}
f_1 Rosenbrock	± 30	0
f_2 Himmelblau	± 30	-3.78396
f_3 Beale's	± 4.5	0
f_4 Easom	± 100	-1
f_5 McCormick	± 4.0	-1.9133
f_6 Three-hump camel	± 5.0	0
f_7 Hölder table	± 10	-19.2085
f_8 Matyas	± 10	0
f_9 Booth's	± 10	0
f_{10} Goldsteinprice	± 2.0	3

In the PSO algorithm each parameter has an important influence on the behavior of the particles and therefore on the convergence of the algorithm; and even if the PSO method presents satisfactory results, the choice of the right parameterization of the method remains a critical point as well as one of the keys to success for any PSO algorithm. In the descriptive section of the PSO method, we have presented some parameters that influence the behavior of particles in their movements in search of the optimum.

The parameters that we developed in our model is the use of several variable parameters that can be modified from the user interface dedicated for it. It all depends on the requirements of the optimization problem. Massive experimentation was carried out to find the appropriate set of parameters; it has given results, which we consider satisfactory.

It is important to note that a simple change in the value of a parameter can greatly change the result, and can even lead to premature convergence. For the present study, which deals with the moderate size, two-dimensional problems, the list of parameters, which gave sufficient good results are mentioned in the table below (see Table. 2).

TABLE II. DESCRIPTION OF THE USED PSO PARAMETERS

PSO parameters	Parallel model
Swarm size	30
Number of iterations	50-80
Acceleration coefficients	C1 = 1.25, C2=2.25, C3=2.25
Inertia factor	(0.4 – 0.2)
Communication topology	Ring
Number of threads	Depends on objective function
Stopping criteria	- The maximum number of iterations. - The maximum number of iterations without improvement of Gbest.

B. Results

The table below shows the results details of the average of 1000 executions: the values of execution time in milliseconds, the SR: the success rate which is the percentage of convergence of the function towards the right solution, the EvalIF which represents the number of evaluation of the objective function, and this for the basic PSO and the proposed parallel model on a set of ten functions.

TABLE III. EXPERIMENTAL RESULTS

From the results obtained, we have observed the performance of the PN-PSO algorithm in terms of the solution quality; it avoids the convergence of particles in local optima. The computation time in the proposed PN-PSO model is also lower than the sequential model.

V. CONCLUSIONS

This paper presents a parallel model based on the Particle Swarm Optimization meta-heuristic. The objective was to propose solutions to the two drawbacks of the method: premature convergence and the high computing time. In the literature, several improving versions of the PSO method are proposed either by adding new parameters, by parallelizing it or by hybridizing it with other meta-heuristics.

The proposed model that we have presented in this paper is based on two concepts: parallelization and neighborhood. The combination of these two notions improved the performance of the method in terms of quality of the solution and of computation time. Our PN-PSO model uses the notion of dynamic neighborhood, which allows to create diversity in the search as well as a better exploration of the search space in order to improve the quality of the solution and avoid the stagnation of the algorithm in a local optimum; parallel computing is used to speed up calculations in order to have an "optimal" solution in a reduced computing time.

REFERENCES

[1] Y. Cooren, Perfectionnement d'un algorithme adaptatif d'Optimisation par Essaim Particulaire. Applications en génie médical et en

électronique, Doctoral thesis, Paris 12 Val de Marne University, France, 2008.

[2] N. Elhami, Contribution aux méthodes hybrides d'optimisation heuristiques: Distribution et application à l'interopérabilité des systèmes d'information", Doctoral thesis, Mohammed V University Rabat, Morocco & Normandy University, France, 2013.

[3] K. Byung-I et G. Alan, Parallel asynchronous particle swarm optimization, in: International Journal For Numerical Methods In Engineering, vol. 67, pp. 578-595, 2006.

[4] Shiyuan Jin, Damian Dechev, Zhihua Qu, Parallel Particle Swarm Optimization (PPSO) on the Coverage Problem in Pursuit-Evasion Games, in : Conference: Proposed for presentation at the 20th High Performance Computing Symposium (HPC 2012), Orlando, FL, 2012.

[5] P. Rabanal, I. Rodríguez et F. Rubio, Parallelizing Particle Swarm Optimization in a Functional Programming Environment, Journal of Algorithms2014 : vol. 7, pp. 554–581, 2014.

[6] Arionde Campo Aurora T.R.Pozo Elias P.Duarte, Parallel multi-swarm PSO strategies for solving many objective optimization problems. Journal of Parallel and Distributed Computing. Vol 126, pp.13-33, 2019.

[7] J. Kennedy et R. Eberhart, Particle Swarm Optimization, in: Proceedings of the IEEE International Joint Conference on Neural Networks, IEEE Press, vol. 8, no. 3, pp. 1943–1948, 1995.

[8] K. E. Parsopoulos et M. N. Vrahatis, Recent approaches to global optimization problems through particle swarm optimization, Natural Computing: an international journal. 1(2-3), pp. 235-306, 2002.

[9] M. E. Hyass et P. Hyass, Good Parameters for Particle Swarm Optimization, Laboratories Technical Report no. HL1001. 2010.

[10] M. Zemzami, VARIATIONS SUR PSO : APPROCHES PARALLELES, JEUX DE VOISINAGES ET APPLICATIONS., Doctoral thesis, Ibn Tofail University Kenitra, Morocco & Normandy

Funtion	PN-PSO model			Basic PSO		
	SR %	Time (ms)	EvalF	SR%	Time (ms)	EvalF
F1	100	891.2	300123	99	967.1	301213.1
F2	100	15.4	9702.5	95	23.7	11089.3
F3	100	511.3	21760.2	100	702.4	23099.4
F4	100	10.3	6222.2	100	12.1	7680.1
F5	100	702.9	26201.3	100	911.4	28901.3
F6	100	831.6	279008	100	997.9	384098
F7	100	12.8	7907.1	100	14.3	8709.8
F8	100	345.5	120312	100	456.7	133098
F9	100	9.8	5780.3	100	11.6	7002
F10	100	13.4	8872	100	15.6	9995.4

University, France, 2019.

[11] Pant, M., Thangaraj, R., & Abraham, A, Particle swarm optimization using adaptive mutation, in: 19th International Conference on Database and Expert Systems , Washington, DC, USA, pp. 519-523, 2008.

[12] Nguyen, U. Q., Hoai, N. X., McKay, R., & Tuan, P. M, Initializing PSO with randomized low-discrepancy sequences: the comparative results, in: Proceedings of the IEEE Congress on Evolutionary Computation pp. 1985-1992, 2007.

[13] K. Zielinski et R. Laur, Stopping Criteria for Differential Evolution in Constrained Single-Objective Optimizationm Advanced in Differential Evolution, the series Studies in Computational Intelligence Vol. 143, pp. 111-138 Springer, Berlin Heidelberg. 2008.

[14] R. Mendes, Population Topologies and Their Influence in Particle Swarm Performance., Doctoral thesis, Minho Uviversity, Portugal, 2004.

- [15] Gerhard Venter et Jaroslaw Sobieszcanski, Parallel particle swarm optimization algorithm accelerated by asynchronous evaluations, *J. Aerosp. Comput. Inf. Commun.* 3(3), pp. 123–137, 2006.
- [16] Byung-II Koh, Alan D. George, Raphael Haftka et Benjamin J Fregly, Parallel asynchronous particle swarm optimization, *Communications in Numerical Methods in Engineering* 67(4) pp. 578-595, 2006.
- [17] Andrew W. McNabb ; Christopher K. Monson ; Kevin D. Seppi, Parallel PSO using MapReduce in : *IEEE Congress on Evolutionary Computation*. 2007.
- [18] Hari Prasain ; Girish Kumar Jha ; Parimala Thulasiraman ; Rupa Thulasiram, A parallel Particle swarm optimization algorithm for option pricing, in *IEEE: International Symposium on Parallel & Distributed Processing, Workshops and Phd Forum (IPDPSW)*. 2010.
- [19] Narjiss Dali et Sadok Bouamama, GPU-PSO : Parallel Particle Swarm Optimization approaches on Graphical Processing Unit for Constraint Reasoning: Case of Max-CSPs, in: *19th International Conference on Knowledge Based and Intelligent Information and Engineering Systems*. *Procedia Computer Science* 60 (2015) pp.1070 – 1080, 2015.
- [20] M.Zemzami, N.Elhami, M.Itmi et N.Hmina, A New Parallel Approach For The Exploitation Of The Search Space Based On PSO Algorithm, in: *4th International Colloquium in Information Science and Technology (CIST'16)*. Tangier. Morocco. 2016.

Proposed Solutions for Security of Smart Traffic Lights using IoT and Machine Learning

L'GHDAICHE Sara

*Network Telecoms and Electrical Engineering
Department*

*University Ibn Tofail, National School of Applied
Sciences Kenitra, Morocco*

lghdaiche.sara@gmail.com

ADDAIM Adnane

*Network Telecoms and Electrical Engineering
Department*

*University Ibn Tofail, National School of Applied
Sciences Kenitra, Morocco*

addaim@gmail.com

EL HASSAK Imad

*Network Telecoms and Electrical Engineering
Department*

*University Ibn Tofail, National School of Applied
Sciences Kenitra, Morocco*

Imad.elhassak@gmail.com

Abstract - The traffic management systems standards are being replaced by traffic lights modes, which automatically adjust control parameters and revise signal plans, and are now an integral part of modern road transport infrastructure. These advances of the Internet of Things (IoT), which allows communication and interaction with various devices in adaptive traffic signals, have been proven to be vulnerable to security breaches and could be easily exploited to allow an attacker to directly modify traffic signal indications. The vulnerabilities could allow full control of traffic control devices and could cause traffic disruption. In this article, a new deep learning model is proposed to analyze data of IoT smart cities. We offer a new model based on Long Short Term memory networks (LSTMs) for predicting anomalies in an intelligent traffic light. The proposed model shows promise and shows that the model can be used in other smart city prediction problems as well.

Keywords - Traffic lights, Security, IoT, Long short term memory (LSTM), Smart city, Predicting anomalies.

devices is changing rapidly as it crosses the total world population so the data generated by IoT devices is going to be huge. IoT is one of the most emerging technologies, but security and confidentiality are still considered as challenges in many fields of application because security problems in IoT networks are much more important with the increasing number of attacks.

Moreover, smart traffic devices face the following challenges:

- Work in real time.
- Dealing with a lot of vehicles using different communication models.
- Facing all failures and any attacks.
- Dealing with the transition time between the parties.

Nowadays some researchers are dealing with security issues in IoT, but as new technologies arrive, people are orienting more researches towards applications based on machine learning alongside that can solve the security problem in IoT, because before the calculation is done of information generated by an IoT application, it is necessary to pass through the verification process to avoid any malicious data or redundancy.

I. INTRODUCTION

Due to the increase in the number of vehicles in cities, the traffic lights modes are replaced by the traffic management systems standards, which automatically adjust control parameters and revise signaling plans [1], are now part of modern infrastructure. These advances in the Internet of Things (IoT) which enables communication and interaction with various devices in the area of adaptive traffic lights has been proven that the IoT is vulnerable to security breaches and which could be easily exploited and which would allow an attacker to directly modify the indications of traffic lights. The vulnerabilities could allow complete control of the traffic control devices and could cause traffic disorder.

The Internet of Things (IoT) is a smart network that shares information on the Internet. IoT allows vehicles to collect information from the road unit and obtain information on the route, time and traffic details [2]. The growing number of IoT

II. RELATED WORK

In this section, we look at some works that use new algorithms based on machine learning to manage security issues in IoT environments.

Currently, Deep learning is recognized as a relevant approach to intrusion detection in networks. The success of deep learning (DL) in various areas of big data has sparked several interests in the areas of cyber security [3].

The study presented in [4] found that NN was used to detect DoS attacks in IoT networks based on a multi-layer perception based control system.

Diro and al. [5] approached deep learning as a new intrusion detection technique due to its ability to extract high-level functionality for IoT. The authors proposed a deep learning-based IoT / Fog network attack detection system. Experience has shown that Distributed Attack Detection can better detect cyber-attacks than centralized algorithms due to sharing parameters and

also demonstrated that the deep model has gone beyond traditional machine learning systems.

Nguyen and al. [6] Introduces an advanced detection mechanism based on a deep learning approach in the cloud environment and they are shown that their proposed learning model can achieve high precision in the detection and isolation of cyber-attacks and surpass other existing machine learning methods.

DL models are regarded as powerful models for showing excellent performance on difficult learning tasks [7]. There are many studies proposing solutions for the prediction problem of DL in the literature. LSTM which is a special type of DL is the state of the art recurrent neural network (RNN) for supervised temporal sequence learning [8]. LSTM has a structure of loops that memorize previous events to make better use of its input [9].

Deep learning for malicious traffic detection has gained several notable achievements with various network models. For example, the authors in [10] proposed a novel network-based anomaly detection method which extracts behavior snapshots of the network and uses deep autoencoders to detect anomalous network traffic emanating from compromised IoT devices.

However, the performance of the work primarily relies on several self-generated synthetic data sets, which may lack the diversity of data exchange. In another research work [11], the authors proposed a malware traffic classification method using CNN by considering traffic data as images. The work is one of the first attempts to apply a representation learning approach for malware traffic classification from raw traffic.

III. DIFFERENT TRAFFIC MANAGEMENT

Before using the attack detection model, the data processing in the system is divided into two types:

Local processing (Figure 3) is carried out by an agent installed at each intersection. It receives data from two sensors installed in the lanes of an intersection (figure 2) and from the central server. After data collection, the processing agent selects a solution from the optimization methods, and then the agent sends the new configurations to the traffic lights. Once the new configurations are defined, the collection and processing process will continue until a traffic problem occurs. In this case, a camera will take images and process them using Machine Learning algorithms (recognition and convolution) in order to detect the presence or absence of an accident at this intersection [12].

Central processing (Figure 3) does not require a lot of resources in terms of throughput. The central server receives a lot of data and then we use the concept of big data to process it,

organize it and classify it according to the needs of the system. On the other hand, we use machine learning algorithms to process and analyse the images collected by the cameras in order to make decisions on the traffic situation in case of traffic problems (detect if there is an accident at an intersection or not) [12].

IV. DIFFERENT ATTACKS IN EACH LAYER OF IOT

Every layer of IoT is vulnerable to security threats and attacks. At every layer, IoT devices and services are vulnerable to Denial of Service (DoS) attacks, which render the device, resource, or network unavailable to authorized users.

This section provides a detailed analysis of the attack issues for each layer.

1. Perception Layer

There are three security concerns in the IoT perception layer [13]: first, the strength of the wireless signals, the signals of which are transmitted between the IoT sensor nodes, the effectiveness of which can be compromised by disturbing waves. Second, the sensor node in IoT devices can be intercepted by the owner and also by attackers, since IoT nodes generally operate in external environments, which involves physical attacks on sensors and IoT devices in which an attacker may damage the device. Third, the inherent nature of the network topology is dynamic because IoT nodes are often moved to different places.

The IoT perception layer mainly consists of sensors and RFID, due to which their storage capacity, their energy consumption and their computation capacity are very limited, which makes them sensitive to many types of threats and attacks. The malicious node injection attack was the most dangerous attack because it stops services and modifies data.

2. Network Layer

The network layer of the IoT is vulnerable to privacy attacks through traffic analysis, eavesdropping and passive surveillance. These attacks have a high probability of occurrence due to the mechanisms of remote access and data exchange of the devices. The network layer is very sensitive to human attack in the middle, which can be followed by eavesdropping. The key exchange mechanism in the IoT must be secure to prevent any intruder from listening and then committing identity theft. Attackers can also take advantage of the fact that everything is connected in order to obtain more information about users and to use this information for future criminal activities. The deepest attack is the most unsafe in the network layer. Because the attacker attracts all traffic to the

base station and can launch other threats such as selective forwarding, packet modification, or drop packages [13].

3. Application Layer

There are many issues related to application security. Large amounts of connected devices that share data will cause significant overhead on the applications that analyze the data, which may have a significant impact on the availability of services. Among the software attacks, the worm attack is the most dangerous [13]. It is a self-replicating program that harms the computer by using security holes in software and network hardware. It can steal information or delete files from the system; it can also change passwords without notifying it and cause the computer to lock, etc. These attacks cause significant damage because they modify data, drop packets, steal private information and the encryption key.

V. SOLUTION APPROACH

A. Long Short Term Memory (LSTM)

LSTM is a type of Recurrent Neural Network (RNN) that is capable of learning long-term additions. It was introduced in order to overcome the leakage gradient problem. In this neural network model, a memory block takes the place of each ordinary neuron in the hidden layer of the standard recurrent neural network [14].

The LSTM block shown in the following figure has an entry gate, a forgetting gate, and an exit gate which regulate the flow of information entering and leaving the cell. These doors block entry and exit as follows:

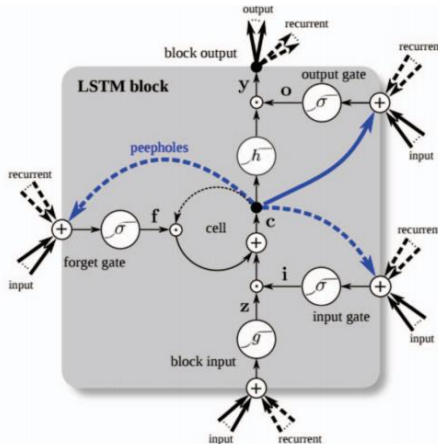


Figure 1. A Long Short-Term Memory Block [16]

LSTM has been applied to a variety of real world problems such as machine translation [17], speech recognition [18], and

image recognition [19]. In this study, a new anomaly prediction model based on LSTM is proposed on IoT data.

B. Proposed anomaly prediction model

In this section, we will present our approach (see figure 3) to have a good prediction of traffic anomalies to secure traffic lights in smart cities. We will explain all the elements that must be brought together to optimize the security of traffic management systems with new techniques.

In this proposed anomaly prediction model, we propose a deep learning model composed of LSTM neural networks. According to this model, the network adopts the architecture which gives the best result.

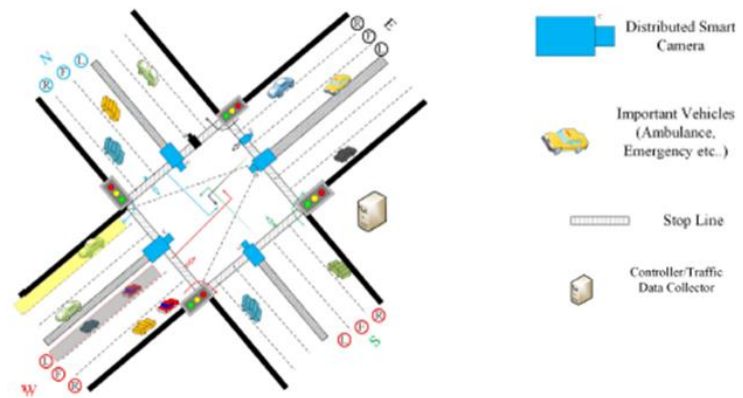


Figure 2. Distributed Smart-camera of IoT [1]

The first step of the processing system (see Fig. 3) is the collection of data from the IoT (Smart Camera) sensors installed in the streets (see Fig. 1) to determine the number of vehicles entering and exiting at an intersection. In addition to these sensors, there are others installed on public vehicles (Important Vehicles), these data collected from the traffic lights are among the most important parameters of this system.

After data collection, the next step is the processing of this data. In our system, we distinguish three stages of processing: the application of the Deep Learning Model, Labeling of this data and then taking the Decision. Basically, the model data is split into two parts: Training set and Test set.

Deep Learning Model (Figure.3), to create a new harmful traffic classification framework, LSTM is able to achieve this thanks to its closed cell. This closed cell has two states, open and closed, which make it act like a computer memory, as it makes decisions about what data it is allowed to write, read, and store in. This feature allows it to keep attack details in the training process and make detection decisions based on this information stored in the closed cell. Moreover, this algorithm detects DDoS attacks with high detection accuracy and low false alarm rates.

Labeling Unit (Figure 3) has been integrated into our proposed attack detection process. It is connected to the deep learning model and automatically receives the predicted data for evaluation. It's divide the data into two units, for example gives to normal information unit's (cars, public transport, ambulance,...) the number 0 and gives to the doubtful information unit's (anomalies or attacks) the number 1, in order to send it to the decision unit which makes the decision to avoid units that are equal to 1.

In the proposed model, a certain number of sequential values are taken into account and predictions are made for next step, denoted Eq. (1).

$$a(t-n), \dots, a(t-2), a(t-1), a(t) = a(t+1) \quad (1)$$

Where $a(t-n), \dots, a(t-2), a(t-1), a(t)$ are the observed values (input sequence) and $a(t+1)$ is the predicted value.

The LSTM-based anomaly prediction model takes advantage of its advantages of storing long historical data and achieving higher prediction accuracy even though it has a simple network structure. Consequently, the employment of the LSTM based prediction model to the IoT data is effective and promising.

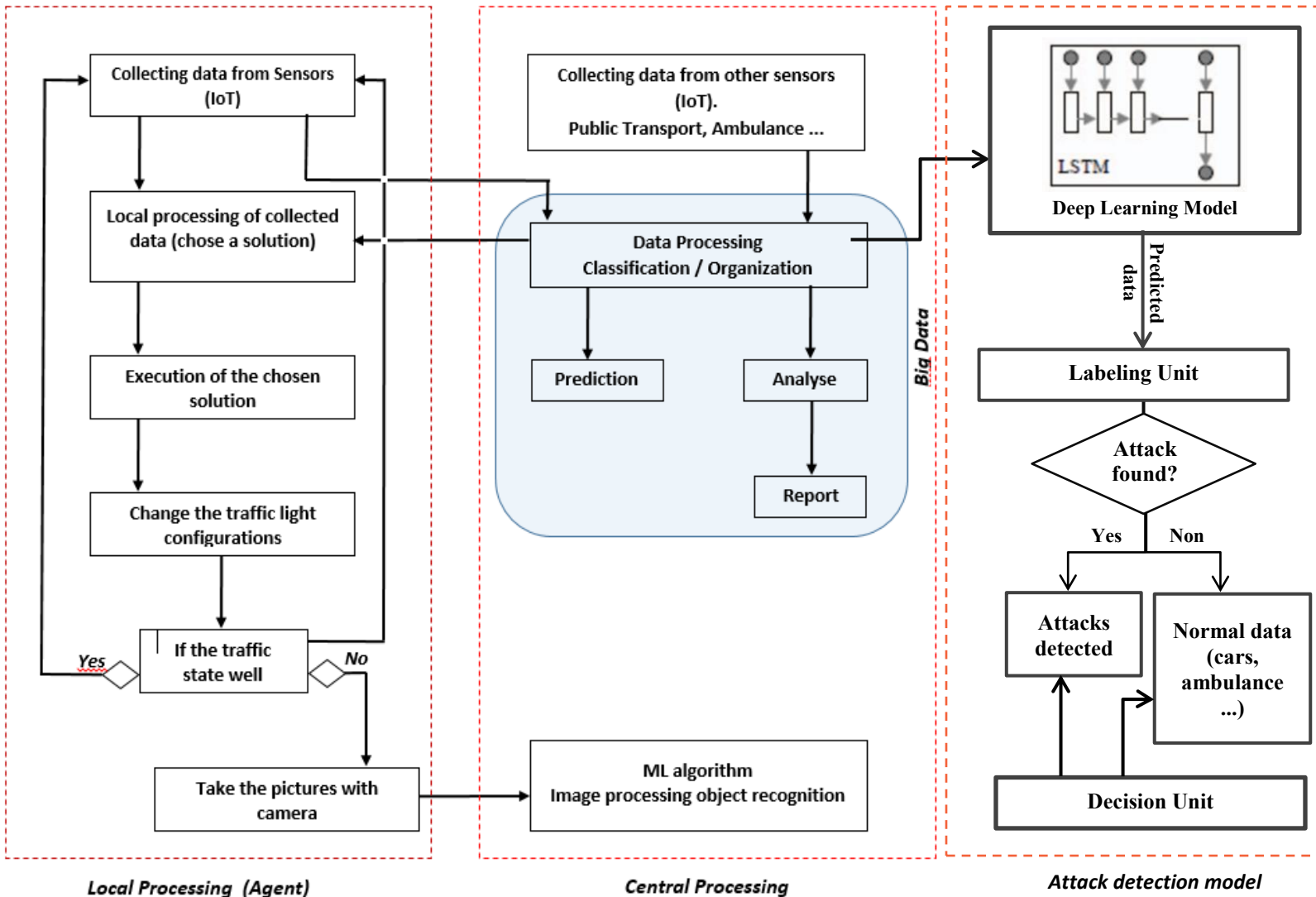


Figure 3. Proposed Anomaly Prediction Model

VI. CONCLUSION

In this article, we have proposed an anomaly prediction model for traffic classification using the Long-Term Memory (LSTM) and Embedding Word Network model. The main advantage of the proposed framework is that it does not require pre-processing packets, thus improving the acceleration of detection. This solution may be changed in the future as the popularity of deep learning attracts more attackers to exploit its vulnerabilities for hacking. Preventing adversary models against deep learning is thus one of the most promising security research topics.

REFERENCES

1. Willy Carlos Tchuitcheu, Christophe Bobda, Md Jubaer Hossain Pantho "Internet of smart-cameras for traffic lights optimization in smart cities".
2. H. Wang, J. Gu, S. Wang, An effective intrusion detection framework based on svm with feature augmentation, *Knowl. Based Syst.* 136 (2017) 130–139.
3. Kulkarni, R.V., Venayagamoorthy, G.K., 2009. Neural network based secure media access control protocol for wireless sensor networks. In: *Proc. Int. Joint Conf. Neural Networks*. June, Atlanta, GA, pp. 3437–3444.
4. H. H. Nguyen, N. Harbi, and J. Darmont, "An efficient local region and clustering-based ensemble system for intrusion detection," in *Proc.*
5. A .A. Diro, N. Chilamkurti, Distributed attack detection scheme using deep learning approach for internet of things, *Future Gener. Comput. Syst.* 82 (2018) 761–768.
6. Nguyen, K. K., Hoang, D. T., Niyato, D., Wang, P., & Dutkiewicz, E. (2018). Cyberattack detection in mobile cloud computing: A deep learning approach. *IEEE*.
7. I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to Sequence Learning with Neural Networks," *Computation and Language*, 2014.
8. F. J. Gomez and J. Schmidhuber, "Co-Evolving Recurrent Neurons Learn Deep Memory POMDPs," in 7th annual conference on Genetic and evolutionary computation, Washington, USA, 2005.
9. Y. Tang, J. Xu, K. Matsumoto, and C. Ono, "Sequence to-Sequence Model with Attention for Time Series Classification," presented at the IEEE 16th International Conference on Data Mining, 2016.
10. Meidan, Y.; Bohadana, M.; Mathov, Y.; Mirsky, Y.; Shabtai, A.; Breitenbacher, D. N-BaIoT—Network-Based Detection of IoT Botnet Attacks Using Deep Autoencoders. *IEEE Pervasive Comput.* 2018, 17, 11–22.
11. Wang,W.; Zhu, M.; Zeng, X.; Ye, X.; Sheng, Y. Malware Traffic Classification Using Convolutional Neural Networks for Representation Learning. In *Proceedings of the 31st International Conference on Information Networking*, Da Nang, Vietnam, 11–13 January 2017; pp. 712–717.
12. Elhassak, Addaim, " Proposed Solutions for Smart Traffic Lights using Machine Learning and Internet of Thing".
13. Rwan Mahmoud, Tasneem Yousuf, Fadi Aloul, Imran Zualkernan " Internet of Things (IoT) Security: Current Status, Challenges and Prospective Measures".
14. Z. C. Lipton, J. Berkowitz, and C. Elkan, "A Critical Review of Recurrent Neural Networks for Sequence Learning," 2015, 2015.
15. K. Greff, R. K. Srivastava, J. Koutn'ík, B. R. Steunebrink, and J. Schmidhuber, "Lstm:A Search Space Odyssey," presented at the *IEEE Transactions on Neural Networks and Learning Systems*, 2016.
16. D. Bahdanau, K. Cho, and Y. Bengio, "Neural Machine Translation by Jointly Learning to Align and Translate," in *International Conference on Learning Representations*, 2014.
17. A. Graves, A.-R. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," presented at the 2013 *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2013.
18. J. Donahue, L. A. Hendricks, S. Guadarrama, M. Rohrbach, S. Venugopalan, K. Saenko, et al., "Long Term Recurrent Convolutional Networks for Visual Recognition and Description," *Conference on Computer Vision and Pattern Recognition(CVPR)*, 2015.
19. A. Shahid, B. Khalid, S. Shaukat, H. Ali, M. Y. Qadri, *Internet of Things Shaping Smart Cities : A Survey*,

- Springer International Publishing, Cham, 2018, pp. 335{358.
20. H. Sak, A. Senior, and F. Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling," 2014.

Effect of CEV on option prices under jump-diffusion dynamics with stochastic volatility: A finite element method approach

1st Abdelilah Jraifi
MISCOM, ENSA Safi,
Cadi Ayyad University, Morocco
jraifi.abdelilah@gmail.com

2nd Aziz Darouichi
L2IS, FST Marrakech,
Cadi Ayyad University, Morocco
az.darouichi@uca.ma

3rd Ilias Elmouki
LERMA, EMI Rabat,
Mohammed V University, Morocco
i.elmouki@gmail.com

Abstract—In this paper, we propose a stochastic volatility jump-diffusion model with a constant elasticity of variance β in the variance equation. This model describes dynamics of the asset price S_t and its variance V_t , based on two stochastic differential equations (SDEs) with Poisson jumps. The resolution of these two SDEs, is essential to find the sought European option price which depends on both main two variables S_t and V_t through a partial integro-differential equation (PIDE). The existence and uniqueness of solution of this PIDE in a weighted Sobolev space, are established based on a variational formulation of the considered problem which we solve using the finite element method (FEM). Spatial differential operators are discretized using $P1$ elements, while the time stepping is performed using an explicit Euler scheme. Finally, we provide some numerical results based on the FEM to show the effect of different values of β on the option prices.

Index Terms—Jump-diffusion model, Option pricing, Partial integro-differential equation, Finite elements method

I. INTRODUCTION

In finance literature, it is common to represent the uncertainty of the economy by a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathcal{P})$ where \mathcal{F}_t is the filtration of available information at time t , and \mathcal{P} is the real probability measure. All processes that we shall consider in this section will be defined in this space. An asset price model with stochastic volatility, in which stochastic volatility follows the CEV model has been defined in [3] as follows

$$\begin{cases} dS_t = \mu S_t dt + S_t V_t dW_t \\ dV_t = p dt + \eta V_t^\beta dW_t' \end{cases} \quad (1)$$

where μS_t is the drift term, V_t is the volatility, ηV_t^β is the volatility term, while p , η and β are non-negative constants. W_t and W_t' are defined as two Brownian motions with $\langle W_t, W_t' \rangle = \rho$ (i.e. ρ is the correlation factor).

The numerical resolution of such problems, has been provided in [1] using the finite element method when there is no jump and also, where the authors have considered the jump-diffusion model (Bates's model [4]) with a stochastic volatility which follows the Heston's model [19] for options pricing. On the other hand, Eraker and Polson in [13],

extended Bates's work in 2003, by incorporating jumps in the volatility equation also. Their model is given by

$$\begin{cases} dS_t = \mu S_t dt + S_t \sqrt{V_t} dW_t + S_t - Y_t dN_t^s \\ dV_t = p dt + \eta \sqrt{V_t} dW_t' + Z_t dN_t^V \end{cases} \quad (2)$$

where N_t^s and N_t^V are independent Poisson processes.

Broadie and Kaya [6] have performed exact method for Eraker's et al model for the evaluation of an European call under the stock index $S\&P500$ by the Monte-Carlo method. In this paper, we would like to generalize the volatility term of the volatility in (2). In fact, we consider V_t^β instead of $\sqrt{V_t}$ in the volatility equation, and we propose the finite element method for comparing the obtained results in [6] and improving the convergence of the RMS error in dimension 2.

The rest of this paper is organized as follows: In section 2., we present the constant elasticity of variance with jumps (CEVJ) model for option pricing problem. In section 3., we introduce a weighted Sobolev space and the variational formulation for the considered problem. In section 4., we provide and compare the obtained numerical results by the two methods mentioned before, namely, the execution time and the RMS error for different values of β . Finally, we conclude our work in section 5.

II. MODEL DESCRIPTIONS

First, we assume there is a risk-neutral probability measure \mathbb{Q} , for more details, refer to [26] and references therein. We consider an European derivative on S_t , denoted by $w(t, S_t, V_t)$ with expiration date T and payoff function h , and by r ; the interest rate. Its price at a time t , will depend on time t , on the price of the underlying asset S_t , and on the volatility V_t . It is given by the risk-neutral expected discounted payoff

$$w(t, s, v) = \mathbb{E}^{\mathbb{Q}} \left[e^{-r(T-t)} h(S_T) / S_t = s, V_t = v \right], \quad (3)$$

In the risk-neutral world, the considered model are governed by the following dynamics:

$$\begin{cases} dS_t = (r - \lambda^s m) S_t dt + S_t \sqrt{V_t} dW_t + S_t - Y_t dN_t^s \\ dV_t = (p - R(t, s, v) \eta V_t^\beta) dt + \eta V_t^\beta dW_t' + Z_t dN_t^V \end{cases} \quad (4)$$

where S_t, V_t, p, η and β are defined above in (2). Moreover, r is the risk-free interest rate, N_t^s and N_t^v are independent Poisson processes with constant intensities λ^s and λ^v respectively. Y_t is the jump size of the asset price return with density $\phi_Y(y)$ and $E(Y_t) := m < \infty$, while Z_t is the jump size of the volatility with density $\phi_Z(z)$.

In addition, we have

$$R(t, s, v) = \rho \frac{\mu - r}{v_t} + \sqrt{1 - \rho^2} \theta^*(t, s, v),$$

where ρ is the correlation factor and $\theta^*(t, s, v)$ is an arbitrary function, for more details, see page 45 in [14]. We note that N_t^s and N_t^v are independent of standard Brownian motions W_t and W'_t .

Since every compound Poisson process can be represented as an integral form of a Poisson random measure (see page 82 "section 3.2" in [7], and equation (3.8) "section 3.3" in [7]), we have

$$\int_0^t Z_s^* dN_s = M_t = \int_0^t \int_{\mathbb{R}} z N(ds, dz)$$

where $N(dt, dz)$ is a Poisson random measure of the process

$$M_t = \sum_{i=1}^{N_t} Z_i^* \text{ with intensity measure } \lambda \phi(dz) dt.$$

Then, the dynamics on the right hand side of (4) can be rewritten as follows

$$\begin{cases} dS_t &= (r - \lambda^s m) S_t dt + S_t \sqrt{V_t} dW_t + \int_{\mathbb{R}} S_t - y N^s(dt, dy) \\ dV_t &= (p - R(t, s, v) \eta V_t^\beta) dt + \eta V_t^\beta dW'_t + \int_{\mathbb{R}} z N^v(dt, dz) \end{cases} \quad (5)$$

N^s is a Poisson random measure of the process $\sum_{i=1}^{N_t} Y_i$ with intensity measure $\nu^s(dy) dt = \lambda^s \phi_Y(y) dt$, and N^v is a Poisson random measure of the process $\sum_{j=1}^{N_t} Z_j$ with intensity

measure $\nu^v(dz) dt = \lambda^v \phi_Z(z) dt$.

In the following, we present a theoretical formulation of the considered problem (5), and we provide a proof of the existence and uniqueness of solution in this general case.

A. General formulation

Let us consider a financial asset whose price $\{S_t, t \geq 0\}$ is given by the following jump-diffusion stochastic differential equation (see page 10 in [8]):

$$dS_t = F(t, S_t, V_t) dt + G(t, S_t, V_t) dW_t + \int_{\mathbb{R}^n} \gamma(t, S(t^-), z) \tilde{N}^S(dt, dz) \quad (6)$$

where $S_0 = s_0 \in \mathbb{R}^n$, $F : [0, T] \times \mathbb{R}^n \times \mathbb{R}^d \rightarrow \mathbb{R}^n$, is a drift term, W_t is a m -dimensional Brownian motion, and $G : [0, T] \times \mathbb{R}^n \times \mathbb{R}^d \rightarrow M_{n \times m}(\mathbb{R})$, is the stochastic volatility. V_t is given by the following stochastic differential equation:

$$dV_t = a(t, S_t, V_t) dt + b(t, S_t, V_t) dW'_t + \int_{\mathbb{R}^d} \chi(t, V(t^-), \varsigma) \tilde{N}^V(dt, d\varsigma) \quad (7)$$

with $V_0 = v_0 \in \mathbb{R}^d$. The simplest models (see [9]) take a constant volatility, but these models are generally not smooth

enough to match real price. The operator $a : [0, T] \times \mathbb{R}^n \times \mathbb{R}^d \rightarrow \mathbb{R}^d$, is a drift term of volatility, $b : [0, T] \times \mathbb{R}^n \times \mathbb{R}^d \rightarrow M_{d \times m}(\mathbb{R})$, is the volatility of volatility, and W'_t is a linear combination of W_t and an independent Brownian motion B_t defined by : $W_t = \rho W'_t + \sqrt{1 - \rho^2} B_t$, where the correlation ρ is some constant in $[-1, 1]$.

$$\begin{aligned} \tilde{N}^S(dt, dz) &= (\tilde{N}_1^S(dt, dz), \dots, \tilde{N}_{l_S}^S(dt, dz)) \\ &= (N_1^S(dt, dz_1) - \nu_1^S(dz_1) dt, \dots, N_{l_S}^S(dt, dz_{l_S}) - \nu_{l_S}^S(dz_{l_S}) dt) \\ \tilde{N}^V(dt, d\varsigma) &= (\tilde{N}_1^V(dt, d\varsigma), \dots, \tilde{N}_{l_V}^V(dt, d\varsigma)) \\ &= (N_1^V(dt, d\varsigma_1) - \nu_1^V(d\varsigma_1) dt, \dots, N_{l_V}^V(dt, d\varsigma_{l_V}) - \nu_{l_V}^V(d\varsigma_{l_V}) dt), \end{aligned}$$

are respectively l_S, l_V independent compensated Poisson random measures, independents of $W()$. For each $k \in \{1, 2, \dots, l_S\}$ and $k' \in \{1, 2, \dots, l_V\}$ we have: $\tilde{N}_k^S(dt, dz) = N_k^S(dt, dz) - \nu_k^S(dz) dt$ and $\tilde{N}_{k'}^V(dt, d\varsigma) = N_{k'}^V(dt, d\varsigma) - \nu_{k'}^V(d\varsigma) dt$, where $\nu_k^S, \nu_{k'}^V$ are the Lévy measures (intensity measures) of the Poisson random measures $N_k^S(\cdot, \cdot)$ and $N_{k'}^V(\cdot, \cdot)$ respectively, see (Theorem 1.7, page 3 in [8], Appendix). We mention that $\{N_k^S\}$ and $\{N_{k'}^V\}$ are independent Poisson random measures with Lévy measures ν^S, ν^V respectively, for all $k \in \{1, 2, \dots, l_S\}$ and $k' \in \{1, 2, \dots, l_V\}$.

Moreover, we assume that the jump processes \tilde{N}^S and \tilde{N}^V are independent of standard Brownian motions W_t and W'_t .

$\gamma : [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow M_{n \times l_S}(\mathbb{R})$, $(t, s, z) \mapsto \gamma(t, s, z)$, and $\chi : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \rightarrow M_{d \times l_V}(\mathbb{R})$, $(t, v, \varsigma) \mapsto \chi(t, v, \varsigma)$, are respectively $n \times l_S$ and $d \times l_V$ matrix of measurable real valued functions which are adapted processes such that the integrals exist. For a detailed presentation of jump-diffusion model, we refer to [8].

$$\gamma = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \dots & \gamma_{1l_S} \\ \gamma_{21} & \gamma_{22} & \dots & \gamma_{2l_S} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \dots & \gamma_{nl_S} \end{pmatrix},$$

$$\chi = \begin{pmatrix} \chi_{11} & \chi_{12} & \dots & \chi_{1l_V} \\ \chi_{21} & \chi_{22} & \dots & \chi_{2l_V} \\ \vdots & \vdots & \ddots & \vdots \\ \chi_{d1} & \chi_{d2} & \dots & \chi_{dl_V} \end{pmatrix}.$$

$\gamma^{(k)} := (\gamma_{1k}, \gamma_{2k}, \dots, \gamma_{nk})$, $\chi^{(k')} := (\chi_{1k'}, \chi_{2k'}, \dots, \chi_{dk'})$ for $k \in \{1, 2, \dots, l_S\}$ and $k' \in \{1, 2, \dots, l_V\}$.

Note that each column $\gamma^{(k)}, \chi^{(k')}$ respectively of the $n \times l_S$ and $d \times l_V$ matrix $\gamma = [\gamma_{ij}]$, depends on z and ς only through the k^{th}, k'^{th} coordinate $z_k, \varsigma_{k'}$, i.e $\gamma^{(k)}(t, s, z) = \gamma^{(k)}(t, s, z_k)$. $\chi^{(k')}(t, v, \varsigma) = \chi^{(k')}(t, v, \varsigma_{k'})$, $z \in \mathbb{R}^{l_S}, \varsigma \in \mathbb{R}^{l_V}$.

If the following assumptions are satisfied:

- The functions F, G, a, b, γ, χ , are measurable.

For every $t \in [0, T]$, $s, s' \in \mathbb{R}^n$, and $v, v' \in \mathbb{R}^d$, and there exists two constants K_1, K_2 such that,

$$\begin{aligned} &\|F(t, s, v) - F(t, s', v')\| + \|G(t, s, v) - G(t, s', v')\| + \\ &\int_{\mathbb{R}^n} \|\gamma(t, s, z) - \gamma(t, s', z)\| \tilde{N}^S(dt, dz) \leq K_1 (\|s - s'\| + \|v - v'\|) \end{aligned}$$

- $\|F(t, s, v)\|^2 + \|G(t, s, v)\|^2 + \frac{\partial u}{\partial x_i} \in L^2_\alpha(U)$, equipped with the norm
- $\int_{\mathbb{R}^n} \|\gamma(t, s, z)\|^2 \tilde{N}^S(dt, dz) \leq K_1^2(1 + \|s\|^2)$
- $\|a(t, s, v) - a(t, s', v')\| + \|b(t, s, v) - b(t, s', v')\| + \int_{\mathbb{R}^n} \|\chi(t, v, \varsigma) - \chi(t, v', \varsigma)\|^2 \tilde{N}^V(dt, d\varsigma) \leq K_2(\|s - s'\| + \|v - v'\|)$
- $\|a(t, s, v)\|^2 + \|b(t, s, v)\|^2 + \int_{\mathbb{R}^n} \|\chi(t, v, \varsigma)\|^2 \tilde{N}^V(dt, d\varsigma) \leq K_2^2(1 + \|v\|^2)$
- S_0, V_0 are square-integrable.

Then, the solution of the system (6)-(7) is unique, for more details see ([10], [11], [12]).

We note that in our special case, we have here $F(t, S_t, V_t) = r - \lambda^s m$, $G(t, S_t, V_t) = S_t \sqrt{V_t}$, $a(t, S_t, V_t) = p - R(t, s, v) \eta V_t^\beta$ and $b(t, S_t, V_t) = \eta V_t^\beta$.

In the following, we formulate the variational problem associated to (5).

III. VARIATIONAL PROBLEM

In weighted Sobolev spaces, Bensoussan and Lions [5], considered that the value of the function w defined in (3), can be characterized as the solution of the following PIDE.

$$\begin{cases} \frac{\partial w}{\partial t}(t, s, v) - \mathcal{L}w(t, s, v) = 0 & \forall t \in [0, T], s \in \mathbb{R}, v \in \mathbb{R}^+ \\ w(T, s, v) = h(s) & s \in \mathbb{R}, v \in \mathbb{R}^+ \end{cases} \quad (8)$$

with

$$\begin{aligned} \mathcal{L}w &= \frac{1}{2} v^2 s^2 \frac{\partial^2 w}{\partial s^2} + \rho \eta v^{\beta+1} s \frac{\partial^2 w}{\partial s \partial v} + \frac{1}{2} \eta^2 v^{2\beta} \frac{\partial^2 w}{\partial v^2} \\ &+ (r - \lambda^s m) s \frac{\partial w}{\partial s} + (p - R(t, s, v) \eta v^\beta) \frac{\partial w}{\partial v} - \\ &(r - \lambda^s m) w + \int_{\mathbb{R}} [u(s + sy, v) - u(s) - sy u'_s(s, v)] \nu^s(dz) \\ &+ \int_{\mathbb{R}} [u(s, v + z) - u(s) - z u'_v(s, v)] \nu^v(dz) \end{aligned} \quad (9)$$

Then, we use the variational formulation in order to solve the partial integro-differential equation (PIDE) by the finite element method (FEM).

For this, let be $U = [0, +\infty[\times]a, b[$ with $0 < a < b < +\infty$ is a domain in \mathbb{R}^2 . Let $\delta = (s, v)$ be a vector in U , with the Euclidean norm defined by the formula $|\delta| = \sqrt{s^2 + v^2}$.

We introduce some weighted Sobolev spaces, $L^2_\alpha(U)$ is a space of measurable functions u and 2^{th} integrable for the measure $e^{-\alpha|\delta|} d\delta$ on U where $\alpha > 0$ and $d\delta = ds dv$. The variational formulation of (8) consists of finding a continuous function u defined on the time interval $[0, T]$ with value in the following weighted Sobolev space (see [1])

$W^{1,2}_{\alpha,1}([0, T] \times U) \equiv W^{1,2}_{\alpha,1}$ the space of functions u in $L^2(0, T; W^{1,2}_\alpha(U))$ such that $\frac{\partial u}{\partial t} \in L^2(0, T; L^2_\alpha(U))$. $W^{1,2}_\alpha(U)$ is the space of functions u in $L^2_\alpha(U)$ such that

$$\|u\|_{W^{1,2}_\alpha(U)} := \|u\|_\alpha = \left(\int_U |u|^2 e^{-\alpha|\delta|} d\delta + \int_U \sum_{i=1}^2 \left| \frac{\partial u}{\partial x_i} \right|^2 e^{-\alpha|\delta|} d\delta \right)^{1/2} \quad (10)$$

Multiplying the PIDE in (8) with a test function $u' \in \mathcal{D}(U)$, we obtain the associated variational formulation, with $u(t, s, v) = w(T - t, s, v)$:

$$\begin{cases} \left(\frac{\partial u}{\partial t}, u' \right) + (\mathcal{L}u, u') = 0 & \forall t \in [0, T] \text{ and } (s, v) \in U \\ (u(0, \cdot, \cdot), u') = (h, u') & \forall (s, v) \in U \end{cases} \quad (11)$$

using the Green's formula and the Dirichlet boundary conditions, we get:

$$\begin{cases} \left(\frac{\partial u}{\partial t}, u' \right) - a(u, u') = 0 & \forall t \in [0, T] \text{ and } (s, v) \in U \\ (u(0, \cdot, \cdot), u') = (h, u') & \forall (s, v) \in U \end{cases} \quad (12)$$

with

$$\begin{aligned} a(u, u') &= \frac{1}{2} \int_U v^2 s^2 \frac{\partial u}{\partial s} \frac{\partial u'}{\partial s} e^{-\alpha|\delta|} d\delta + \int_U v^2 s \frac{\partial u}{\partial s} u' e^{-\alpha|\delta|} d\delta \\ &- \int_U (r - \lambda^s m) s \frac{\partial u}{\partial s} u' e^{-\alpha|\delta|} d\delta + \int_U \frac{\eta^2}{2} v^{2\beta} \frac{\partial u}{\partial v} \frac{\partial u'}{\partial v} e^{-\alpha|\delta|} d\delta \\ &+ \int_U \beta \eta^2 v^{2\beta-1} \frac{\partial u}{\partial v} u' e^{-\alpha|\delta|} d\delta - \int_U (p - \eta v^\beta R) \frac{\partial u}{\partial v} u' e^{-\alpha|\delta|} d\delta \\ &+ \int_U \rho \eta v^{\beta+1} s \frac{\partial u}{\partial v} \frac{\partial u'}{\partial s} e^{-\alpha|\delta|} d\delta + \int_U \rho \eta v^{\beta+1} \frac{\partial u}{\partial v} u' e^{-\alpha|\delta|} d\delta \\ &+ \int_U (r - \lambda^s m) u u' e^{-\alpha|\delta|} d\delta - \int_U \frac{\alpha s^3 v^2}{2|\delta|} \frac{\partial u}{\partial s} u' e^{-\alpha|\delta|} d\delta \\ &- \int_U \rho \alpha \eta \frac{sv^{\beta+1}}{|\delta|} \frac{\partial u}{\partial s} u' e^{-\alpha|\delta|} d\delta - \int_U \frac{\alpha \eta^2 v^{2\beta+1}}{2|\delta|} \frac{\partial u}{\partial v} u' e^{-\alpha|\delta|} d\delta \\ &- \int_U \left(\int_{\mathbb{R}} [u(s + sy, v) - u(s) - sy \frac{u}{\partial s}] \nu^s(dy) \right) u' e^{-\alpha|\delta|} d\delta \\ &- \int_U \left(\int_{\mathbb{R}} [u(s, v + z) - u(v) - z \frac{u}{\partial v}] \nu^v(dz) \right) u' e^{-\alpha|\delta|} d\delta \end{aligned} \quad (13)$$

where

$$(\mathcal{L}u, u') = -a(u, u')$$

For $h \in W^{1,2}_\alpha \cap L^\infty$, the variational problem (12) admits a unique solution in $W^{1,2}_{\alpha,1} \cap L^\infty$. This solution has the probabilistic representation (3), for more details see ([1], [2]). In the next section, we present and discuss some numerical results.

IV. NUMERICAL RESULTS

With the growing complexity of models and derivatives, the numerical methods associated with assessing financial options became an important field of research over the last decade. In the present section, we will implement a European call using the associated PIDE (12) to the CEVJ model (5) for option pricing. We will present the results of the simulations given by finite elements method using FreeFem++ software. The numerical experiments were performed on a Sony Vaio Laptop with an Intel® Pentium® CPU P6100@ 2.00 GHZ processor and 4 Go of RAM, running Windows 7 (64 bits).

We consider a European call for the *S&P500* stock index whose true value is equal to 6.8619 for this purpose. The parameters used in our numerical experiments are $K = 100$, $p = 0.00346$, $\rho = -0.82$, $R_{moy} = 3.14$, $\eta = 0.05$, $r = 3.19\%$, $T = 1 \text{ year}$.

Then, we will solve and compare the resolution approach of variational problem for the pricing of the considered European option by using the finite elements approximation in space, and an explicit Euler discretization in time.

For the numerical simulation, we consider the problem (12) on a bounded domain $\Omega = (S_{min}, S_{max}) \times (v_{min}, v_{max})$, where (v_{min}, v_{max}) does not contain zero. The corresponding variational problem is then given by

$$\begin{cases} \left(\frac{\partial u}{\partial t}, u' \right) - a(u, u') = 0, \quad \forall t \in [0, T], \quad (s, v) \in \Omega \\ (u(0, \cdot, \cdot), u') = (h, u') \quad \forall (s, v) \in \Omega \end{cases} \quad (14)$$

where $a(u, u')$ is given by (13). The resolution of this problem using FreeFem++ with *P1* finite elements, provides the following numerical results as illustrated in Table I and Table II.

TABLE I
COMPUTING TIMES OF FEM IN SECOND UNIT, ASSOCIATED TO CEV VALUES UTILIZED FOR EACH NUMBER OF TIME STEPS

No of time steps	Comput times (sec)	Weight of space α	Power of volatility β
10	2.0256	0.30217	2/5
20	12.7682	0.82002	1/3
40	18.8287	2.1362469	2/7
80	37.3328	2.2957453	1/4
160	58.1834	2.3424291	2/9

In Table I, we present different values of CEV, namely β , associated to each number of time steps. This number of iterations has been changed in a way to show the time when it has been observed there is a reduction or minimization of the difference between the values of estimated and true option prices based on different statistical factors as it will be discussed hereafter.

TABLE II
BIAS, MEAN SQUARED ERROR (MSE) AND THE VALUES OF THE OPTION PRICE ASSOCIATED TO EACH CEV VALUE

Power of volatility β	Bias	MSE	Estimated price of option	True price of option
2/5	0.5023	0.327447	7.3084	6.8061
1/3	0.4451	0.001161	6.9512	
2/7	0.2167	0.000196	7.0228	
1/4	0.1246	0.000004	6.9307	
2/9	0.0835	0.0000002	6.8896	

In Table II, we show the numerical values of the estimated option price for different values of CEV. We can deduce from this table that as more the value of CEV is small, as more the value of the bias and the mean squared error (MSE) become small, while the estimated option price becomes closer to the true option price when β is chosen small.

V. CONCLUSION

The aim of this paper, is to show the effect of CEV on the pricing of the European option under jump-diffusion model with stochastic volatility. The study has been based on the formulation of a variational problem resolved using the finite elements method (FEM) for some values of β or more precisely when it is strictly smaller than 1/2. We concluded that the obtained values of the option price are closer to the true market values of the European option exercised under the stock index *S&P500* on March 2, 2014.

REFERENCES

- [1] R. Aboulaich, F. Bagheri, A. Jraifi, Option pricing for a stochastic volatility jump-diffusion model, *International Journal of Mathematics and Statistics* 13, p. 1-19, 2013.
- [2] R. Aboulaich, F. Bagheri, A. Jraifi, Numerical approximation for options pricing of a stochastic volatility jump-diffusion model, *Int. J. Applied. Math. Stat* 50, p. 69-82, 2013.
- [3] R. Aboulaich, L. Hadji, A. Jraifi, Option pricing with constant elasticity of variance (CEV) model, *Applied Mathematical Sciences* 7, p. 5443-5456, 2013.
- [4] D.S. Bates, Jumps and stochastic volatility : exchange rate processes implicit in deutsche mark option, *The Review of financial Studies* 9, p. 69-107, 1996.
- [5] A. Bensoussan, J.L. Lions, *Contrôle Impulsionnel et Inéquations Quasi-Variationnelles*, Dunod, Paris, 1980.
- [6] M. Broadie, O. Kaya, Exact simulation of stochastic volatility and other affine jump diffusion processes, *Operations Research* 54, p. 217-231, 2006.
- [7] R. Cont, P. Tankov, *Financial modeling with jump processes*, edition of Chapman & Hall/CRC Financial Mathematics, Boca Raton, FL, Volume 1, 2004.
- [8] B. Øksendal, A. Sulem, *Applied stochastic control of jump diffusions (2nd Edition)*, Universitext, Springer-Verlag, New York, 2007.
- [9] X. Zhang, Numerical analysis of American option pricing in a jump-diffusion model, *Mathematics of operation research* 22, p. 668-690, 1997.
- [10] I.I.Gikhman, A.V. Skorokhod, *Stochastic differential equations*, Springer Verlag, New York, 1972.
- [11] J. Jacod, A.N. Shiryaev, *Limit theorems for stochastic processes*, *Grundlehren der Mathematische Wissenschaften*, Springer Verlag, New York, 1987.
- [12] M. Johannes, N. Polson, J. Stroud, Filtering of stochastic differential equations with jumps, Working paper, Columbia University 11, p. 1-44, 2002.
- [13] B. Eraker, M. Johannes, N. Polson, The impact of jumps in equity index volatility and returns, *Journal of Finance* 58, p. 1269-1300, 2008.
- [14] J.P. Fouque, G. Papanicolaou, K.R. Sircar, *Derivatives in Financial Markets with stochastic Volatility*, Cambridge University Press, 2000.
- [15] K.O. Friedrichs, The identity of weak and strong extensions of differential operators, *Rans, Amer, Math, Soc* 55, p. 132-151, 1987.
- [16] J.W. Gao, Optimal Portfolios for DC Pension Plans under a CEV Model, *Insurance : Mathematics and Economics* 44, p. 479-490, 2009a.
- [17] J.W. Gao, Optimal Investment Strategy for Annuity Contracts under the Constant Elasticity of Variance (CEV) Model, *Insurance : Mathematics and Economics* 45, p. 9-18, 2009b.
- [18] J.W. Gao, An extended CEV model and the Legendre transform-dual-asymptotic solutions for annuity contracts, *Insurance : Mathematics and Economics* 46, p. 511-530, 2010.
- [19] S. Heston, A Closed-Form solution for Options with stochastic volatility with applications to bond and currency options, *Review of Financial Studies* 6, p. 327-343, 1993.
- [20] Y.L. Hsu, T.I. Lin, C.F. Lee, Constant Elasticity of Variance (CEV) Option Pricing Model, Integration and Detailed Derivation, *Mathematics and Computer in Simulation* 79, p. 60-71, 2008.
- [21] J. Hull, A. White, The pricing of options on assets with stochastic volatilities, *Journal of Finance* 42, p. 281-300, 1989.
- [22] J.L. Lions, E. Magenes, *Méthode Numérique pour les Equations aux Dérivées Partielles en Finance*, Addison-Wesley Reading, Massachusetts, 1994.

- [23] A.E. Lindsay, D.R. Brecher, Simulation of the CEV process and the local martingale property, *Mathematics and Computers in Simulation* 82, p. 868-878, 2012.
- [24] J.D. Macbeth, L.J. Merville, Tests of the Black-Scholes and Cox Call Option Valuation Models, *Journal of Finance* 35, p. 285-300, 1980.
- [25] P. Protter, *Stochastic Integration and Differential Equations*, Second Edition, Springer-Verlag, 2003.
- [26] H. Geman, N. El Karoui, J.C. Rochet, Changes of Numeraire, Changes of Probability Measure and Option Pricing, *Journal of Applied Probability* 32, p. 443-458, 1995.
- [27] J. Xiao, H. Zhai, C. Qin, H. Zhai, C. Qin, The Constant Elasticity of Variance (CEV) Model and the Legendre Transform-Dual Solution for Annuity Contracts, *Insurance : Mathematics and Economics* 40, p. 302-310, 2007.

Weak solution to a class of quasilinear elliptic System in Orlicz-Sobolev Spaces

1st Hamza EL-HOUARI
Sultan Moulay Slimane University
Faculty of Science and Technology.
Beni Mellal, Morocco.
h.elhouari94@gmail.com

2nd L. Saâdia CHADLI
Sultan Moulay Slimane University
Faculty of Science and Technology.
Beni Mellal, Morocco.
sa.chadli@yahoo.fr

3rd MOUSSA HICHAM
Departement of Mathematics LMACS
Sultan Moulay Slimane University
Beni Mellal, Morocco.
hichammoussa23@gmail.com

Abstract—The purpose of this paper is to investigate on the existence of a weak solution to the following quasilinear system driven by the M -Laplacian

$$\begin{cases} (-\Delta_{m_1})u = F_u(x, u, v) & \text{in } \Omega, \\ (-\Delta_{m_2})v = F_v(x, u, v) & \text{in } \Omega, \\ u = v = 0 & \text{on } \partial\Omega, \end{cases} \quad (1)$$

where Ω is a bounded open subset in \mathbb{R}^N and $(-\Delta_m)$ is the M -Laplacian operator.

Index Terms—Orlicz-Sobolev spaces, M -Laplacian Operator, Variational problem, Elliptic system.

I. INTRODUCTION

A natural question is to see what results can be recovered when the standard Laplace operator is replaced by the fractional m -Laplacian. In the recent years many others has been an increasing interest in studying non-local problems with p -structure due to its accurate description of models involving anomalous diffusion.

This type of operators arises in many different applications, such as, continuum mechanics, phase transition phenomena, population dynamics, minimal surfaces and game theory, as they are the typical outcome of stochastically stabilization of Levy processes, see for example [7], [12].

In this paper we deal with the existence of a solution to the following quasilinear elliptic system problem

$$\begin{cases} (-\Delta_{m_1})u = F_u(x, u, v) & \text{in } \Omega, \\ (-\Delta_{m_2})v = F_v(x, u, v) & \text{in } \Omega, \\ u = v = 0 & \text{on } \partial\Omega, \end{cases} \quad (1.1)$$

where Ω is a bounded open subset in \mathbb{R}^N and $(-\Delta_{m_i})$ is the M -Laplacian operator defined by

$$(-\Delta_{m_i})u := -\operatorname{div}(m_i(|\nabla u| \nabla u)), \quad i = 1, 2. \quad (1.2)$$

When we take $m_1(t) = |t|^{p-2}$, $m_2(t) = |t|^{q-2}$ ($p, q > 1$). Then the system (1.1) reduces to the following (p, q) -Laplacian system :

$$\begin{cases} (-\Delta)_p u = F_u(x, u, v) & \text{in } \Omega, \\ (-\Delta)_q v = F_v(x, u, v) & \text{in } \Omega, \\ u = v = 0 & \text{on } \partial\Omega. \end{cases} \quad (1.3)$$

The existence of solutions for systems like (1.3) have also received a wide range of interests. For this we find in the literature have many researchers have studied this type of systems using some important methods, such as variational method, Nehari manifold and fibering method, three critical points theorem (see for instance [2]–[4]).

In [13], Huentutripay-Manásevich studied an eigenvalue problem to the following system:

$$\begin{cases} -\operatorname{div}(m_1(|\nabla u| \nabla u)) = \lambda F_u(x, u, v) & \text{in } \Omega, \\ -\operatorname{div}(m_2(|\nabla v| \nabla v)) = \lambda F_v(x, u, v) & \text{in } \Omega, \\ u = v = 0 & \text{on } \partial\Omega. \end{cases}$$

For a certain λ , the authors translated the existence of solution into a suitable minimizing problem and proved the existence of solution under some reasonable restriction.

Liben, Zhang and Fang in [15] studied the problem (1.1) by using the Mountain Pass Theorem and they obtained the following result:

Theorem 1.1: [Theorem 3.1 [15]] Assumes that the following conditions hold:

$(\phi_1)'$: $m_i \in C(0, +\infty)$; $tm_i(t) \rightarrow 0$ as $t \rightarrow 0$; $tm_i(t) \rightarrow +\infty$ as $t \rightarrow +\infty$.

$(\phi_2)'$: $tm_i(t)$ are strictly increasing.

$(\phi_3)'$:

$$1 < l_i := \inf_{t>0} \frac{t^2 m_i(t)}{M(t)} \leq \sup_{t>0} \frac{t^2 m_i(t)}{M(t)} := n_i < N,$$

where

$$M_i(t) = \int_0^{|t|} sm_i(s) ds, \quad \text{for all } t \in \mathbb{R},$$

and F satisfies:

$(F_0)'$: $F : \Omega \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ is a C^1 function such that $F(x, 0, 0) = 0$, for all $x \in \Omega$.

$(F_1)'$: There exist two continuous functions $\Psi_{i=1,2} : [0, +\infty) \rightarrow \mathbb{R}$, which satisfy that

$$\Psi_i(t) := \int_0^{|t|} \psi_i(s) ds, \quad \text{for all } t \in \mathbb{R},$$

are two N -functions increasing essentially more slowly than $M_{i=1,2}^*$ near infinity, respectively, where M_i^* is the Sobolev

conjugate function of M_i , which will be specified later. Moreover,

$$n_i < l_{\Psi_i} := \inf_{t>0} \frac{t\psi_i(t)}{\Psi_i(t)} \leq \sup_{t>0} \frac{t\psi_i(t)}{\Psi_i(t)} := n_{\Psi_i} < \infty,$$

such that

$$\begin{cases} |F_u((x, u, v))| \leq c_1(1 + \psi_1(|u|) + \overline{\Psi}_1^{-1}(\Psi_2(v))), \\ |F_v((x, u, v))| \leq c_1(1 + \psi_2(|v|) + \overline{\Psi}_2^{-1}(\Psi_1(u))), \end{cases}$$

for all $(x, u, v) \in \Omega \times \mathbb{R} \times \mathbb{R}$, where c_1 is a positive constant, $\overline{\Psi}$ denote the complements of $\Psi_{i=1,2}$, respectively.

(F₂)'

$$\lim_{|(u,v)| \rightarrow +\infty} \frac{F(x, u, v)}{M_1(u) + M_2(v)} = \infty, \quad \text{uniformly for all } x \text{ in } \Omega.$$

(F₃)':

$$\lim_{|(u,v)| \rightarrow 0} \sup \frac{|F(x, u, v)|}{\lambda_1 m_1(u) + \lambda_2 m_2(v)} = c_0$$

(F₄)': There exists a continuous function $\gamma : [0, \infty) \rightarrow \mathbb{R}$ such that

$$\Gamma(t) := \int_0^{|t|} \gamma(s) ds, \quad \text{for all } t \in \mathbb{R},$$

is an N-function with

$$1 < l_\Gamma := \inf_{t>0} \frac{t\gamma(t)}{\Gamma(t)} \leq \sup_{t>0} \frac{t\gamma(t)}{\Gamma(t)} := n_\Gamma < +\infty,$$

and functions $H_i(t) := |t| \frac{l_i l_\Gamma}{l_i l_\Gamma - 1}$, $t \in \mathbb{R}$ increase essentially more slowly than M_i^* near infinity, respectively, such that

$$\Gamma \left(\frac{F(x, u, v)}{|u|^{l_1} + |v|^{l_2}} \right) \leq c_2 \overline{F}(x, u, v), \quad x \in \mathbb{R}, \quad |(u, v)| \geq r,$$

where c_2, r are two strictly positive constants and

$$\overline{F}(x, u, v) := \frac{1}{n_1} F_u(x, u, v)u + \frac{1}{n_2} F_v(x, u, v)v - F(x, u, v),$$

$\forall (x, u, v) \in \Omega \times \mathbb{R} \times \mathbb{R}$. Then the flowing system has a nontrivial solution

$$\begin{cases} -\text{div}(m_1(|\nabla u|)\nabla u) = F_u(x, u, v) & \text{in } \Omega, \\ -\text{div}(m_2(|\nabla v|)\nabla v) = F_v(x, u, v) & \text{in } \Omega, \\ u = v = 0 & \text{on } \partial\Omega. \end{cases}$$

We get motivated by Theorem 1.1 above and by relaxing hypotheses $(\phi_3)'$, $(F_1)'$ and $(F_2)'$ we shall prove the existence of solution to our problem (I.1).

Let $M : \mathbb{R} \rightarrow \mathbb{R}^+$ be an N-function, i. e., M is an even and convex function such that

$$M(t) > 0 \quad \text{for } t \neq 0, \quad \lim_{t \rightarrow 0} \frac{M(t)}{t} = 0 \quad \text{and} \quad \lim_{t \rightarrow \infty} \frac{M(t)}{t} = \infty,$$

Equivalently, M admits the representation:

$$M(t) = \int_0^{|t|} m(s) ds,$$

where $m : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is a non-decreasing and right continuous function, with

$$m(0) = 0, \quad m(t) > 0 \quad \text{for } t > 0 \quad \text{and} \quad \lim_{t \rightarrow \infty} m(t) = \infty. \quad (I.4)$$

The N-function \overline{M} complementary to M is defined by $\overline{M}(t) = \int_0^{|t|} \overline{m}(s) ds$, where $\overline{m} : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ satisfies (I.4).

The relationship that relates M and \overline{M} is shown in

$$\overline{M}(t) := \sup_{t \geq 0} \{ts - M(t)\}. \quad (I.5)$$

We shall show that the representation giving by

$$M_{i=1,2}(t) := \int_0^{|t|} r m_i(r) dr \quad \text{for all } t \in \mathbb{R}, \quad (I.6)$$

where $m_{i=1,2}$ verified (I.4) exists and it's an N-functions.

Proof I.2: By theorem 1.1 in [11]. Every convex function H which satisfies the condition $H(a) = 0$ can be represented

in the form $H(t) = \int_a^{|t|} h(r) dr$ for all $t \in \mathbb{R}$, where

$h(t)$ is a non-decreasing right-continuous function. Note that $h(r) = r m(r)$ then we have by definition of m in (I.4) that h is a non-decreasing and right continuous function for all $t \geq 0$ and we have that

$$h(0) = 0, \quad h(t) = t m(t) > 0 \quad \text{for } t > 0 \quad \text{and} \quad \lim_{t \rightarrow \infty} h(t) = \lim_{t \rightarrow \infty} t m(t) = \infty. \quad (I.7)$$

Then M_i defined in (I.6) is an N-function.

We suppose through our paper that M_i above are satisfying Δ_2 -condition globally. Then by lemma 2.5 we have for all $t > 0$ that

$$1 < l_i := \inf_{t>0} \frac{t^2 m_i(t)}{M_i(t)} \leq \sup_{t>0} \frac{t^2 m_i(t)}{M_i(t)} := n_i < N \quad (I.8)$$

Related to function F our hypotheses are the following:

F satisfies:

(F₁): $F : \Omega \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ is a C^1 function such that $F(x, 0, 0) = 0$ for all $x \in \Omega$ and there exist two N-functions increasing essentially more slowly than $M_{i=1,2}$ near infinity $\Psi_{i=1,2} : \mathbb{R} \rightarrow \mathbb{R}^+$, which satisfy that

$$n_i < l_{\Psi_i} := \inf_{t>0} \frac{t\psi_i(t)}{\Psi_i(t)} \leq \sup_{t>0} \frac{t\psi_i(t)}{\Psi_i(t)} := n_{\Psi_i} < \infty, \quad (I.9)$$

where

$$\Psi_i(t) := \int_0^{|t|} \psi_i(r) dr, \quad \text{for all } t \in \mathbb{R}.$$

Moreover,

$$\begin{cases} |F_u((x, u, v))| \leq c_1(1 + \psi_1(|u|) + \overline{\Psi}_1^{-1}(\Psi_2(v))), \\ |F_v((x, u, v))| \leq c_1(1 + \psi_2(|v|) + \overline{\Psi}_2^{-1}(\Psi_1(u))), \end{cases} \quad (I.10)$$

for all $(x, u, v) \in \Omega \times \mathbb{R} \times \mathbb{R}$, where c_1 is a positive constant, $\overline{\Psi}$ denote the complements of $\Psi_{i=1,2}$, respectively.

(F₂)

$$\lim_{|(u,v)| \rightarrow 0} \frac{|F(x, u, v)|}{M_1(u) + M_2(v)} = 0, \quad \text{uniformly for all } x \text{ in } \Omega. \quad (I.11)$$

And

$$\lim_{|(u,v)| \rightarrow +\infty} \frac{F(x, u, v)}{M_1(u) + M_2(v)} = \infty, \quad \text{uniformly for all } x \text{ in } \Omega. \quad (I.12)$$

(F₃): There exists a continuous function $\gamma : [0, \infty) \rightarrow \mathbb{R}$ such that

$$1 < l_\Gamma := \inf_{t>0} \frac{t\gamma(t)}{\Gamma(t)} \leq \sup_{t>0} \frac{t\gamma(t)}{\Gamma(t)} := n_\Gamma < +\infty, \quad (I.13)$$

where

$$\Gamma(t) := \int_0^{|t|} \gamma(r)dr, \quad \text{for all } t \in \mathbb{R},$$

is an N-function and functions $H_i(t) := |t|^{\frac{l_i l_\Gamma}{l_\Gamma - 1}}, t \in \mathbb{R}$ increase essentially more slowly than M_i near infinity, respectively, such that

$$\Gamma\left(\frac{F(x, u, v)}{|u|^{l_1} + |v|^{l_2}}\right) \leq c_2 \bar{F}(x, u, v), \quad x \in \mathbb{R}, \quad |(u, v)| \geq r, \quad (I.14)$$

where c_2 and r are tow strictly positive constants and

$$\bar{F}(x, u, v) := \frac{1}{n_1} F_u(x, u, v)u + \frac{1}{n_2} F_v(x, u, v)v - F(x, u, v),$$

$$\forall (x, u, v) \in \Omega \times \mathbb{R} \times \mathbb{R}.$$

A. Examples

We set some examples which are in at least one M and \bar{M} can be not reflexive:

- 1) $m(t) = |t|^{p-1}$ where $1 < p < \infty$. This is a case of polynomial growth, M and \bar{M} satisfy the Δ_2 -condition. We are in a reflexive situation, the classical theory of monotone operators in reflexive Banach can be applied.
- 2) $m(t) = \text{sgn } t \log(1 + |t|)$ This is a case of slow growth, M satisfies the Δ_2 -condition but \bar{M} does not.
- 3) $m(t) = \text{sgn } t \cdot (e^{|t|} - 1)$. This is a case of rapid growth, M does not satisfy the Δ_2 -condition but \bar{M} does. For further examples we refer to ([11] p 28).

This paper is organized as follows: In the second Section, we recall some well-known properties and results on Orlicz and Orlicz Sobolev spaces. Third Section we present the existence of a solution to the problem (I.1) and its proof which relies on the Mountain Pass Theorem.

II. SOME PRELIMINARY RESULTS AND HYPOTHESES

In this section, we list some basic properties of the Orlicz-Sobolev Space. We refer the reader to [8], [10], [11] for further references and for some of the proofs of the results in this section.

Let Ω be an open subset of $\mathbb{R}^N, N \in \mathbb{N}$.

Definition 2.1: The N -function M satisfies a Δ_2 -condition globally, if for some constant $k > 2$,

$$M(2t) \leq k M(t), \quad \text{for every } t > 0. \quad (II.1)$$

The Orlicz space $L_M(\Omega)$ is defined as the set of equivalence classes of real-valued measurable functions u on Ω such that:

$$\int_\Omega M\left(\frac{u(x)}{\lambda}\right) dx < +\infty \text{ for some } \lambda > 0. \quad (II.2)$$

Notice that $L_M(\Omega)$ is a Banach space under the so-called Luxemburg norm, namely

$$\|u\|_M = \inf \left\{ \lambda > 0 / \int_\Omega M\left(\frac{u(x)}{\lambda}\right) dx \leq 1 \right\}. \quad (II.3)$$

In $L_M(\Omega)$ we define the Orlicz norm $\|u\|_{(M)}$ by

$$\|u\|_{(M)} = \sup \int_\Omega u(x)v(x) dx, \quad (II.4)$$

where the supremum is taken over all $v \in E_{\bar{M}(\Omega)}$ such that $\|v\|_{\bar{M}, \Omega} \leq 1$. An important inequality in $L_M(\Omega)$ is the following:

$$\int_\Omega M(u(x)) dx \leq \|u\|_{(M)} \text{ for all } u \in L_M(\Omega) \text{ such that } \|u\|_{(M)} \leq 1, \quad (II.5)$$

wherefrom we readily deduce

$$\int_\Omega M\left(\frac{u(x)}{\|u\|_{(M)}}\right) dx \leq 1 \text{ for all } u \in L_M(\Omega) \setminus \{0\}. \quad (II.6)$$

It can be shown that the norm $\|\cdot\|_{(M)}$ is equivalent to the Luxemburg norm $\|\cdot\|_{M, \Omega}$. Indeed,

$$\|u\|_{M, \Omega} \leq \|u\|_{(M)} \leq 2\|u\|_{M, \Omega} \text{ for all } u \in L_M(\Omega). \quad (II.7)$$

We have the following inequality

$$\|u\|_{(M), \Omega} \leq \int_\Omega M(u(x))dx + 1 \text{ for all } u \in L_M(\Omega). \quad (II.8)$$

Also, the Hölder inequality holds

$$\int_\Omega |u(x)v(x)| dx \leq \|u\|_{M, \Omega} \|v\|_{\bar{M}} \text{ for all } u \in L_M(\Omega) \text{ and } v \in L_{\bar{M}}(\Omega),$$

in particular, if Ω has finite measure, Hölder's inequality yields the continuous inclusion $L_M(\Omega) \subset L^1(\Omega)$.

For Orlicz spaces Young inequality reads as follows:

$$st \leq M(s) + \bar{M}(t) \text{ for all } t, s \geq 0 \text{ and } x \in \Omega. \quad (II.9)$$

Lemma 2.2: ([10]) If $M_2 \ll M_1$ then

$$L_{M_1}(\Omega) \hookrightarrow L_{M_2}(\Omega)$$

We now turn to the Orlicz-Sobolev Space defined by

$$W^1 L_M(\Omega) := \left\{ u \in L_M(\Omega) : \frac{\partial u}{\partial x_i} \in L_M(\Omega), i = 1, \dots, N \right\}$$

equipped with the norm $\|u\|_{1, M} = \|u\|_M + \|\nabla u\|_M$ and

$$W^1 E_M(\Omega) := \left\{ u \in E_M(\Omega) : \frac{\partial u}{\partial x_i} \in E_M(\Omega), i = 1, \dots, N \right\}$$

Thus $W^1 L_M(\Omega)$ and $W^1 E_M(\Omega)$ are tow Banach spaces under the Luxemburg norm. Denote

$$W_0^1 L_M(\Omega) = \overline{C^\infty(\Omega)}^{\|\cdot\|_{1, M}}.$$

Definition 2.3: Let $(u_k) \in L_M(\Omega)$ and $u \in L_M(\Omega)$. We say that u_k converges to u for the modular convergence in $L_M(\Omega)$ if for some $\lambda > 0$, $\int_{\Omega} M\left(\frac{u_k - u}{\lambda}\right) dx \rightarrow 0$. The fact that M satisfies a Δ_2 -condition globally implies that

$$u_k \rightarrow u \text{ in } L_M(\Omega) \Leftrightarrow \int_{\Omega} M((u_k - u)) dx \rightarrow 0. \quad (\text{II.10})$$

Theorem 2.4 (Generalized Poincaré Inequality): Let Ω be a bounded open subset of \mathbb{R}^N and let M be an N-function. Then there exists a positive constant μ such that,

$$\|u\|_M \leq \mu \|u\|_{0,M}, \quad \forall u \in W_0^1 L_M(\Omega). \quad (\text{II.11})$$

Notation : In this work we note $W_0^1 L_M(\Omega)$ by $W_0^{1,M}(\Omega)$ and $W^1 L_M(\Omega)$ by $W^{1,M}(\Omega)$

Next, we give some inequalities which will be used in our proofs. For the proofs, we refer the reader to the papers [1], [9].

Lemma 2.5: Let $\xi_0(t) = \min\{t^l, t^n\}$, $\xi_1(t) = \max\{t^l, t^n\}$, $t \geq 0$, M is an N-function, then the following conditions are equivalent :

1)
$$1 < l := \inf_{t>0} \frac{tM(t)}{M(t)} \leq \sup_{t>0} \frac{tM(t)}{M(t)} := n < N. \quad (\text{II.12})$$

2)
$$\xi_0(t)M(\rho) \leq M(\rho t) \leq \xi_1(t)M(\rho), \quad \forall t, \rho \geq 0.$$

3) M satisfies a Δ_2 -condition globally.

Lemma 2.6: If (II.12) hold then

$$\xi_0(\|u\|_{M,\Omega}) \leq \int_{\Omega} M(u) dx \leq \xi_1(\|u\|_{M,\Omega}), \quad \forall u \in L_M(\Omega).$$

Lemma 2.7: Let \bar{M} be the complement of M and $\xi_2(t) = \min\{t^{\bar{l}}, t^{\bar{n}}\}$, $\xi_3(t) = \max\{t^{\bar{l}}, t^{\bar{n}}\}$, $t \geq 0$ where $\bar{l} = \frac{l}{l-1}$ and $\bar{n} = \frac{n}{n-1}$. If M is an N-function and (II.12) holds with $l > 1$, then \bar{M} satisfies:

1)
$$\bar{n} = \inf_{t>0} \frac{t\bar{M}'(t)}{\bar{M}(t)} \leq \sup_{t>0} \frac{t\bar{M}'(t)}{\bar{M}(t)} = \bar{l}.$$

2)
$$\xi_2(t)\bar{M}(\rho) \leq \bar{M}(\rho t) \leq \xi_3(t)\bar{M}(\rho), \quad \forall t, \rho \geq 0.$$

3)
$$\xi_2(\|u\|_{\bar{M}}) \leq \int_{\Omega} \bar{M}(u) dx \leq \xi_3(\|u\|_{\bar{M}}), \quad \forall u \in L_{\bar{M}}(\Omega).$$

Lemma 2.8: Let $\xi_4(t) = \min\{t^{l^*}, t^{n^*}\}$, $\xi_5(t) = \max\{t^{l^*}, t^{n^*}\}$, $t \geq 0$ where $l^* = \frac{lN}{N-l}$ and $n^* = \frac{nN}{N-n}$. If M is an N-function and (II.12) in Lemma 2.7 hold with $l, n \in (1, N)$, then M^* satisfies:

1)
$$l^* = \inf_{t>0} \frac{t(M^*)'(t)}{M^*(t)} \leq \sup_{t>0} \frac{t(M^*)'(t)}{M^*(t)} = n^*.$$

2)
$$\xi_4(t)M^*(\rho) \leq M^*(\rho t) \leq \xi_5(t)M^*(\rho), \quad \forall t, \rho \geq 0.$$

3)
$$\xi_4(\|u\|_{M^*}) \leq \int_{\Omega} M^*(u) dx \leq \xi_5(\|u\|_{M^*}), \quad \forall u \in L_{M^*}(\Omega).$$

Lemma 2.9: Under the assumption of Lemma 2.8, the embedding from $W_0^{1,M}(\Omega)$ into $L_{M^*}(\Omega)$ is continuous and into $L_{\Phi}(\Omega)$ is compact for any N-function Φ increasing essentially more slowly than M^* near infinity.

Lemma 2.10: M increases essentially more slowly than M^* near infinity, i.e,

$$\lim_{t \rightarrow \infty} \frac{M(kt)}{M^*(t)} = 0 \quad \text{for every constant } k > 0.$$

Proof 2.11: by 2) Lemma 2.5 and 2) Lemma 2.8,

$$0 \leq \frac{M(kt)}{M^*(t)} \leq \frac{M(k)\xi_1(t)}{M^*(1)\xi_2(t)} = \frac{M(k)t^n}{M^*(1)t^{l^*}}$$

for $1 \leq t$. Since $n < l^*$, we have the result.

Due to the nature of M -Laplacian operator defined in I.2 we need to consider the Orlicz-Sobolev framework and we will examine some specific techniques to Orlicz and the Orlicz-Sobolev spaces. For that we define

$$W := W_0^{1,M_1}(\Omega) \times W_0^{1,M_2}(\Omega)$$

equipped with the following norm $\|u, v\| = \|\nabla u\|_{M_1} + \|\nabla v\|_{M_2}$.

We can see that W is a separable and reflexive Banach space.

Definition 2.12: We define a weak solution (u, v) in W to the problem (I.1) by

$$\langle -\Delta_{m_1} u, \bar{u} \rangle + \langle -\Delta_{m_2} v, \bar{v} \rangle = \int_{\Omega} F_u(x, u, v) \bar{u} dx + \int_{\Omega} F_v(x, u, v) \bar{v} dx \quad (\text{II.13})$$

for all $(\bar{u}, \bar{v}) \in W$, where

$$\langle -\Delta_{m_1} u, \bar{u} \rangle = \langle \mathcal{H}_1(u), \bar{u} \rangle := \int_{\Omega} m_1(|\nabla u|) \nabla u \nabla \bar{u} dx,$$

and

$$\langle -\Delta_{m_2} v, \bar{v} \rangle = \langle \mathcal{H}_2(v), \bar{v} \rangle := \int_{\Omega} m_2(|\nabla v|) \nabla v \nabla \bar{v} dx.$$

Now define the operators $\mathcal{H}_i : W_0^{1,M_i}(\Omega) \rightarrow (W_0^{1,M_i}(\Omega))^*$ by

$$\langle \mathcal{H}_i(u), \bar{u} \rangle := \int_{\Omega} m_i(|\nabla u|) \nabla u \nabla \bar{u} dx.$$

Lemma 2.13: [5] The function \mathcal{H}_i is of type (S_+) . i.e. given a sequence (u_k) converges weakly to u in $W_0^{1,M_i}(\Omega)$ and

$$\limsup_{n \rightarrow \infty} \langle \mathcal{H}_i(u_k), u_k - u \rangle \leq 0. \quad (\text{II.14})$$

Then (u_k) converge strongly to $u \in W_0^{1,M_i}(\Omega)$.

We observe that the energy functional I on W corresponding to system (I.1) is

$$I(u, v) := \int_{\Omega} M_1(|\nabla u|)dx + \int_{\Omega} M_2(|\nabla v|) - \int_{\Omega} F(x, u, v)dx,$$

for all $(u, v) \in W$. Denote by $I_i (i = 1, 2) : W \rightarrow \mathbb{R}$ the functionals

$$I_1(u, v) := \int_{\Omega} M_1(|\nabla u|)dx + \int_{\Omega} M_2(|\nabla v|)dx$$

and

$$I_2(u, v) = \int_{\Omega} F(x, u, v)dx.$$

Then $I(u, v) = I_1(u, v) - I_2(u, v)$.

The function I_1 is well-defined and of class $C^1(W, \mathbb{R})$ and we have the following representation

$$\langle I'(u, v), (\bar{u}, \bar{v}) \rangle = \langle \mathcal{H}_1(u), \bar{u} \rangle + \langle \mathcal{H}_2(v), \bar{v} \rangle - \int_{\Omega} F_u(x, u, v)\bar{u}dx - \int_{\Omega} F_v(x, u, v)\bar{v}dx,$$

for all $(\bar{u}, \bar{v}) \in W$. Then, the critical points of I on W are weak solutions of system I.1.

III. MAIN RESULTS

In this section, we present the following existence result by using mountain pass theorem, see [14].

Theorem 3.1: Assume that $(F_0) - (F_2)$ and (F_3) hold. Then system (I.1) possesses a nontrivial weak solution.

Remark 3.2: Under assumptions (F_1) and (F_3) , by Lemma 2.10, the following embeddings $W_0^{s, m_i}(\Omega) \rightarrow L^{\Psi_i}(\Omega)$, $W_0^{s, m_i}(\Omega) \rightarrow L^{l_i \tilde{l}_\Gamma}(\Omega)$ and $W_0^{s, m_i}(\Omega) \rightarrow L^{l_i \tilde{m}_\Gamma}(\Omega)$ are compact where $\tilde{l} = \frac{l_\Gamma}{l_\Gamma - 1}$ and $\tilde{m} = \frac{m_\Gamma}{m_\Gamma - 1}$.

Remark 3.3: By 2) in Lemma 2.5, assumptions (F_2) and (F_3) show

$$\lim_{|(u, v)| \rightarrow +\infty} \tilde{F}(x, u, v) \rightarrow +\infty, \quad \text{uniformly for all } x \in \Omega.$$

Remark 3.4: Based on the Youngs inequality (II.9), $F(x, 0, 0) = 0$ and the fact

$$F(x, u, v) = \int_0^u F_s(x, s, 0)ds + \int_0^v F_t(x, 0, t)dt + F(x, 0, 0),$$

$$\forall (x, u, v) \in \Omega \times \mathbb{R} \times \mathbb{R}.$$

By (I.10) and 2) in Lemma 2.5, show that there exists a constant $c_4 > 0$ such that

$$|F(x, u, v)| \leq c_4(\Psi_1(u) + \Psi_2(v)), \quad \forall (x, u, v) \in \Omega \times \mathbb{R} \times \mathbb{R}. \tag{III.1}$$

Theorem 3.5: Let E be a real Banach space with its dual space E^* , and suppose that $J \in C^1(E, \mathbb{R})$ satisfies

$$\max\{J(0), J(e)\} \leq \alpha < \beta \leq \inf_{\|u\|=\rho} J(u),$$

for some $\alpha > \beta$, $\rho > 0$ and $e \in E$ with $\|e\| > \rho$. Let $c \geq \beta$ be characterized by

$$c = \inf_{\gamma \in \Gamma} \max_{t \in [0, 1]} J(\gamma(t))$$

where $\Gamma = \{\gamma \in C([0, 1], E) : \gamma(0) = 0, \gamma(1) = e\}$ is the set of continuous paths joining 0 and e , then there exists a sequence $\{u_k\} \subset E$ such that

$$J(u_k) \rightarrow c \geq \beta \text{ and } \|J'(u_k)\|_{E^*}(1 + \|u_k\|) \rightarrow 0 \text{ as } n \rightarrow \infty. \tag{III.2}$$

This kind of sequence is usually called a Cerami sequence.

Definition 3.6: Let $J : W_0^{1, M}(\Omega) \rightarrow \mathbb{R}$ is a class C^1 . We say that a sequence u_k in a Banach Space $W_0^{1, M}(\Omega)$ is a Cerami sequence (in short $(C)_c$) at the level $c \in \mathbb{R}$ for the functional J when

$$J(u_k) \rightarrow c \text{ and } (1 + \|u_k\|)\|J'(u_k)\| \rightarrow 0.$$

Lemma 3.7: Let E be a real Banach Space and $I \in C^1(E, \mathbb{R})$ satisfying (PS)-condition. Suppose $I(0) = 0$ and (I_1) there are constants $\rho, \alpha > 0$ such that $I|_{\partial B_\rho} \geq \alpha$. (I_2) there is an $e \in E \setminus B_\rho$ such that $I(e) \leq 0$.

Then I possesses a critical value $c \geq \alpha$.

Lemma 3.8: Suppose that (F_1) hold. Then there are constants $\rho, \alpha > 0$ such that $I|_{\partial B_\rho} \geq \alpha$.

Proof 3.9: By equation (I.10) there exists $c_4 > 0$ such that

$$|F(x, u, v)| \leq (1-\epsilon)(\lambda_1 M_1(u) + \lambda_2 M_2(v)) + c_4(1 + \Psi_1(u) + \Psi_2(v)),$$

where $\|u, v\| \leq 1$, by Poincaré's Inequality and Lemma 2.5 we obtain

$$\begin{aligned} I(u, v) &= \int_{\Omega} M_1(|\nabla u|)dx + \int_{\Omega} M_2(|\nabla v|)dx - \int_{\Omega} F(x, u, v)dx \\ &\geq \epsilon \min\{\|\nabla u\|_{M_1}^{l_1}, \|\nabla u\|_{M_1}^{n_1}\} + \epsilon \min\{\|\nabla v\|_{M_2}^{l_2}, \|\nabla v\|_{M_2}^{n_2}\} \\ &\quad - c_4 \int_{\Omega} \Psi_1(u)dx - c_4 \int_{\Omega} \Psi_2(v)dx \\ &\geq \|\nabla u\|_{M_1}^{n_1} (\epsilon - c_4 \|\nabla u\|_{M_1}^{l_{\Psi_1} - n_1}) \\ &\quad + \|\nabla v\|_{M_2}^{n_2} (\epsilon - c_4 \|\nabla v\|_{M_2}^{l_{\Psi_2} - n_2}), \end{aligned}$$

since $1 < n_i < l_{\Psi_i}$ we can choose positive constants ρ and α small enough such that $I(u, v) > \alpha$ for all $(u, v) \in W$ with $\|(u, v)\| = \rho$.

Lemma 3.10: Suppose that (F_3) hold. Then there is a point $(u, v) \in W \setminus B_\rho$ such that $I(u, v) \leq 0$.

Proof 3.11: By (F_3) and the fact that F is continuous, then for any given constant $G > 0$, there exists a constant $C_G > 0$ such that

$$F(x, u, v) \geq G(M_1(u) + M_2(v)) - C_G \quad \forall (x, u, v) \in \Omega \times \mathbb{R} \times \mathbb{R}. \tag{III.3}$$

Now, choose $u_0 \in C_c^1(\Omega) \setminus \{0\}$ with $0 \leq u_0(x) \leq 1$. Then $(u_0, 0) \in W$, and by (III.3) and 2) in Lemma 2.5, when $t > 0$ we have

$$\begin{aligned} I(tu_0, 0) &= \int_{\Omega} M_1(t|\nabla u_0|)d\mu - \int_{\Omega} F(x, tu_0, 0)dx \\ &\leq M_1(t)(\|\nabla u_0\|_{l_1}^{l_1} + \|\nabla u_0\|_{n_1}^{n_1} - M\|u_0\|_{n_1}^{n_1}) + C_G|\Omega|, \end{aligned}$$

Since $G > 0$ is arbitrary and $\lim_{t \rightarrow \infty} M_1(t) = +\infty$, we can

choose $G > \frac{\|\nabla u_0\|_{l_1}^{l_1} + \|\nabla u_0\|_{n_1}^{n_1}}{\|u_0\|_{n_1}^{n_1}}$ and large t such that $I(tu_0, 0) \leq 0$ and $\|(tu_0, 0)\| > \rho$.

Lemma 3.12: Suppose that $(F_0) - (F_2)$ and (F_3) hold. Then $(C)_c$ -sequence in W is bounded.

Proof 3.13: Let $\{u_k, v_k\} \in W$ be a $(C)_c$ -sequence of I in W , then for n large enough by (III.2), we obtain

$$\begin{aligned} c + 1 &\geq I(u_k, v_k) - \langle I'(u_k, v_k), \left(\frac{1}{n_1}u_k, \frac{1}{n_2}v_k\right) \rangle \\ &= \int_{\Omega} M_1(|\nabla u_k|)dx + \int_{\Omega} M_2(|\nabla v_k|)dx - \int_{\Omega} F(x, u_k, v_k)dx \\ &\quad - \frac{1}{n_1} \int_{\Omega} m_1(|\nabla u_k|)|\nabla u_k|^2 dx - \frac{1}{n_2} \int_{\Omega} m_2(|\nabla v_k|)|\nabla v_k|^2 dx \\ &\geq \int_{\Omega} \bar{F}(x, u_k, v_k)dx, \end{aligned}$$

by contradiction, we prove the boundedness of sequence $\{(u_k, v_k)\}$. Suppose that there exists a sub-sequence of $\{(u_k, v_k)\}$, still denoted by $\{(u_k, v_k)\}$, such that

$$\|(u_k, v_k)\| = \|\nabla u_k\|_{M_1} + \|\nabla v_k\|_{M_2} \rightarrow +\infty.$$

Next, we discuss the problem in two cases.

Case I: Suppose that $\|\nabla u_k\|_{M_1} \rightarrow +\infty$ and $\|\nabla v_k\|_{M_2} \rightarrow +\infty$. Let $\bar{u}_k = \frac{u_k}{\|\nabla u_k\|_{M_1}}$ and $\bar{v}_k = \frac{v_k}{\|\nabla v_k\|_{M_2}}$. Then $\{(\bar{u}_k, \bar{v}_k)\}$ is bounded in separable, reflexive Banach space W . Passing to a subsequence less denoted by $\{(\bar{u}_k, \bar{v}_k)\}$ by Remark 3.2, there exists a point $(\bar{u}, \bar{v}) \in W$ such that:

- (a) $\bar{u}_k \rightharpoonup \bar{u}$ in $W_0^{1, M_1}(\Omega)$; $\bar{u}_k \rightarrow \bar{u}$ in $L^{l_1 \bar{\Gamma}}(\Omega)$ and $L^{l_2 \bar{\Gamma}}(\Omega)$; $\bar{u}_k \rightarrow \bar{u}$ in a.e in Ω .
- (b) $\bar{v}_k \rightharpoonup \bar{v}$ in $W_0^{1, M_2}(\Omega)$; $\bar{v}_k \rightarrow \bar{v}$ in $L^{l_1 \bar{\Gamma}}(\Omega)$ and $L^{l_2 \bar{\Gamma}}(\Omega)$; $\bar{v}_k \rightarrow \bar{v}$ in a.e in Ω .

Firstly, we assume that $[\bar{u} \neq 0] := [x \in \Omega : \bar{u}(x) \neq 0]$ or $[\bar{v} \neq 0] := [x \in \Omega : \bar{v}(x) \neq 0]$ has nonzero Lebesgue measure. It is clear that

$$|u_k| = |\bar{u}_k| \|\nabla u_k\|_{M_1} \rightarrow +\infty \quad \text{in } [\bar{u} \neq 0],$$

and

$$|v_k| = |\bar{v}_k| \|\nabla v_k\|_{M_2} \rightarrow +\infty \quad \text{in } [\bar{v} \neq 0].$$

Then, by (III.4) and Fatou's Lemma, we have

$$c + 1 \geq \int_{\Omega} \bar{F}(x, u_k, v_k)dx \rightarrow +\infty,$$

which is a contradiction. Next, we assume that both $[\bar{u} \neq 0]$ and $[\bar{v} \neq 0]$ have zero Lebesgue measure, that is $\bar{u} = 0$ in $W_0^{1, M_1}(\Omega)$ and $\bar{v} = 0$ in $W_0^{1, M_2}(\Omega)$. By Lemma 2.6, we have

$$\begin{aligned} &\min\{\|\nabla u_k\|_{M_1}^{l_1}, \|\nabla u_k\|_{M_1}^{n_1}\} \\ &+ \min\{\|\nabla u_k\|_{M_2}^{l_2}, \|\nabla v_k\|_{M_2}^{n_2}\} \\ &\leq I(u_k, v_k) + \int_{\Omega} F(x, u_k, v_k)dx. \end{aligned} \tag{III.5}$$

When k is large enough, that is

$$\|\nabla u_k\|_{M_1}^{l_1} + \|\nabla u_k\|_{M_2}^{l_2} \leq I(u_k, v_k) + \int_{\Omega} F(x, u_k, v_k)dx,$$

which is equivalent to

$$\begin{aligned} 1 &\leq \frac{I(u_k, v_k)}{\|\nabla u_k\|_{M_1}^{l_1} + \|\nabla u_k\|_{M_2}^{l_2}} + \int_{|u_k, v_k| \leq R} \frac{F(x, u_k, v_k)}{\|\nabla u_k\|_{M_1}^{l_1} + \|\nabla v_k\|_{M_2}^{l_2}} dx \\ &\quad + \int_{|u_k, v_k| > R} \frac{F(x, u_k, v_k)}{\|\nabla u_k\|_{M_1}^{l_1} + \|\nabla v_k\|_{M_2}^{l_2}} dx = o_k(1) \\ &\quad + \int_{|u_k, v_k| \leq R} \frac{F(x, u_k, v_k)}{\|\nabla u_k\|_{M_1}^{l_1} + \|\nabla u_k\|_{M_2}^{l_2}} dx \\ &\quad + \int_{|u_k, v_k| > R} \frac{F(x, u_k, v_k)}{\|\nabla u_k\|_{M_1}^{l_1} + \|\nabla v_k\|_{M_2}^{l_2}} dx, \end{aligned} \tag{III.4}$$

where R is a positive constant such that $R > r$ (see (F_4)), bearing in mind that $|(u, v)| > R$ and by (F_2) we have

$$F(x, u, v) \geq 0, \quad x \in \Omega.$$

For $|(u, v)| \leq R$ and the fact that F is continuous, there exists a constant $C_R > 0$ such that

$$|F(x, u, v)| < C_R, \quad \forall x \in \Omega, \tag{III.7}$$

then

$$\int_{|u_k, v_k| \leq R} \frac{F(x, u_k, v_k)}{\|\nabla u_k\|_{M_1}^{l_1} + \|\nabla v_k\|_{M_2}^{l_2}} dx \leq o_k(1). \tag{III.8}$$

Besides, it follows from Höder's inequality that

$$\begin{aligned} &\int_{|u_k, v_k| > R} \frac{F(x, u_k, v_k)}{\|\nabla u_k\|_{M_1}^{l_1} + \|\nabla v_k\|_{M_2}^{l_2}} dx \\ &\leq 2 \left\| \frac{F(x, u_k, v_k)}{|u_k|^{l_1} + |v_k|^{l_2}} \chi_{\{|(u_k, v_k)| > R\}} \right\|_{\Gamma} \\ &\quad \times \left(\|\bar{u}_k\|_{\Gamma}^{l_1} + \|\bar{v}_k\|_{\Gamma}^{l_2} \right) \chi_{\{|(u_k, v_k)| > R\}} \Big|_{\bar{\Gamma}}, \end{aligned} \tag{III.9}$$

where χ denotes the characteristic function which satisfies

$$\chi_{\{|(u_k(x), v_k(x))| > R\}} = \begin{cases} 1 & \text{for } x \in \{x \in \Omega : |(u_k(x), v_k(x))| > R\} \\ 0 & \text{for } x \in \{x \in \Omega : |(u_k(x), v_k(x))| \leq R\} \end{cases}$$

For k large enough, by (I.14), (III.4) and the fact that \bar{F} is continuous, we obtain

$$\begin{aligned} &\int_{\Omega} \Gamma \left(\frac{F(x, u_k, v_k)}{|u_k|^{l_1} + |v_k|^{l_2}} \chi_{\{|(u_k, v_k)| > R\}} \right) dx \\ &\leq c_2 \int_{\Omega} \bar{F}(x, u_k, v_k)dx + C \leq c_2(c + 1) + C \end{aligned}$$

Then, for k large enough, by Lemma 2.6, there exists a constant $c_6 > 0$ such that

$$\left\| \frac{F(x, u_k, v_k)}{|u_k|^{l_1} + |v_k|^{l_2}} \chi_{\{|(u_k, v_k)| > R\}} \right\|_{\Gamma} \leq c_6. \tag{III.10}$$

Moreover, it is easy to see that

$$\begin{aligned} &\|(|\bar{u}_k|^{l_1} + |\bar{v}_k|^{l_2}) \chi_{\{|(u_k, v_k)| > R\}}\|_{\bar{\Gamma}} \leq \|\bar{u}_k\|_{\bar{\Gamma}}^{l_1} \\ &\quad + \|\bar{v}_k\|_{\bar{\Gamma}}^{l_2} \leq \|\bar{u}_k\|_{\bar{\Gamma}}^{l_1} + \|\bar{v}_k\|_{\bar{\Gamma}}^{l_2}. \end{aligned}$$

By Lemma 2.5, Lemma 2.7 and (F_3) implies that N-function $\bar{\Gamma}$ satisfies a Δ_2 -condition globally. Then by (II.10), $\|u\|_{\bar{\Gamma}} \rightarrow 0$ as $\int_{\Omega} \bar{\Gamma}(|u|)dx \rightarrow 0$. It follows from Lemma 2.7, (a) and (b) in case 1 that

$$\int_{\Omega} \bar{\Gamma}(|\bar{u}_k|^{l_1})dx + \int_{\Omega} \bar{\Gamma}(|\bar{v}_k|^{l_2})dx \leq o_k(1),$$

which implies

$$\|(|\bar{u}_k|^{l_1} + |\bar{v}_k|^{l_2})\chi_{\{|(u_k, v_k)| > R\}}\|_{\bar{\Gamma}} \leq \|\bar{u}_k\|_{\bar{\Gamma}}^{l_1} + \|\bar{v}_k\|_{\bar{\Gamma}}^{l_2} \tag{III.11}$$

$$= o_k(1).$$

By combining (III.8), (III.9), (III.10), (III.11) with (III.6) we get a contradiction.

Case2. Suppose that $\|\nabla u_k\|_{M_1} \leq C$ or $\|\nabla v_k\|_{M_2} \leq C$, for some $C > 0$ and all $k \in \mathbb{N}$. Without loss of generality, we assume that $\|\nabla u_k\|_{M_1} \rightarrow +\infty$ and $\|\nabla v_k\|_{M_2} \leq C$, for some $C > 0$ and for all $k \in \mathbb{N}$. Let $\bar{u}_k = \frac{u_k}{\|\nabla u_k\|_{M_1}}$ and $\bar{v}_k = \frac{v_k}{\|\nabla u_k\|_{M_1}}$ then $\|\bar{v}_k\|_{0, M_2} \rightarrow 0$ and $\|\bar{u}_k\|_{0, M_1} \rightarrow 1$. By Remark 3.2, there exists a point $(\bar{u}, \bar{v}) \in W$ such that:

(c) $\bar{u}_k \rightarrow \bar{u}$ in $W_0^{1, M_1}(\Omega)$, $\bar{u}_k \rightarrow \bar{u}$ in $L^{l_1 \bar{\Gamma}}(\Omega)$ and $L^{l_2 \bar{\Gamma}}(\Omega)$ $\bar{u}_k \rightarrow \bar{u}$ in a.e in Ω ,

(d) $\bar{v}_k \rightarrow \bar{v}$ in $W_0^{1, M_2}(\Omega)$; $\bar{v}_k \rightarrow \bar{v}$ in $L^{l_2 \bar{\Gamma}}(\Omega)$ and $L^{l_2 \bar{\Gamma}}(\Omega)$; $\bar{v}_k \rightarrow \bar{v}$ in a.e in Ω . Similarly, we firstly assume that $[\bar{u} \neq 0]$ has nonzero Lebesgue measure. We can see that

$$|u_k| = |\bar{u}_k| \|u_k\|_{s, M_1} \rightarrow +\infty, \quad \text{in } [\bar{u} \neq 0].$$

Then, by (III.4) and Fatou's Lemma, we get a contradiction by

$$c + 1 \geq \int_{\Omega} \bar{F}(x, u_k, v_k)dx \rightarrow +\infty.$$

Next, we suppose that $[\bar{u} \neq 0]$ has zero Lebesgue measure, that is $\bar{u} = 0$ in $W_0^{1, M_1}(\Omega)$. By Lemma 2.7 and (c) and (d) in case 2 we have

$$\min \left\{ \| |v_k|^{l_2} \|_{\bar{\Gamma}}, \| |v_k|^{l_2} \|_{\bar{\Gamma}} \right\} \leq \bar{\Gamma}(1)$$

$$\times \left(\int_{\Omega} |v_k|^{l_2 \bar{\Gamma}} dx + \int_{\Omega} |v_k|^{l_2 \bar{\Gamma}} dx \right) dx \rightarrow C,$$

Then there exists a constant $L > 0$ such that

$$\| |v_k|^{l_2} \|_{\bar{\Gamma}} \leq L, \quad \forall k \in \mathbb{N}. \tag{III.12}$$

When k large enough, (III.5) changed into

$$\|\nabla u_k\|_{M_1}^{l_1} + K \leq I(u_k, v_k) + \int_{\Omega} F(x, u_k, v_k) + K,$$

where K is a positive constant with $K > 4Lc_6$ (see (III.10) and (III.12)). Then by (III.7), (III.10), (III.11), (III.12) and Höder's Inequality, above estimate means

$$\begin{aligned} 1 &\leq \frac{I(u_k, v_k) + K}{\|\nabla u_k\|_{M_1}^{l_1} + K} + \int_{\Omega} \frac{F(x, u_k, v_k)}{\|\nabla u_k\|_{M_1}^{l_1} + K} dx \\ &\leq o_k(1) + 2c_6(o_k(1) + \frac{L}{K}) < o_k(1) + \frac{1}{2}, \end{aligned}$$

which is a contradiction.

Lemma 3.14: Suppose that $(F_1) - (F_3)$ hold. Then I satisfies $(C)_c$ -condition.

Proof 3.15: Let $\{(u_k, v_k)\}$ be any $(C)_c$ -sequence of I in W . Lemma 2.13 shows $\{(u_k, v_k)\}$ is bounded. Passing to a subsequence denote by $\{(u_k, v_k)\}$, there exists a point $(u, v) \in W$ such that:

(e) $u_k \rightarrow u$ in $W_0^{1, M_1}(\Omega)$, $u_k \rightarrow u$ in $L^{\Psi_1}(\Omega)$,

$u_k \rightarrow u$ a.e Ω .

(f) $v_k \rightarrow v$ in $W_0^{1, M_2}(\Omega)$, $v_k \rightarrow v$ in $L^{\Psi_2}(\Omega)$, $v_k \rightarrow v$ a.e Ω . then we have

$$\begin{aligned} \langle \mathcal{H}_1(u_k), u_k - u \rangle &= \int_{\Omega} m_1(|\nabla u_k|) \nabla u_k \nabla (u_k - u) dx \\ &= \langle I'(u_k, v_k), (u_k - u, 0) \rangle + \int_{\Omega} F_u(x, u_k, v_k) dx. \end{aligned} \tag{III.13}$$

Equation (III.2) shows that

$$\left| \langle I'(u_k, v_k), (u_k - u, 0) \rangle \right| \leq \|I'(u_k, v_k)\|_{W_0^{-1, M_1}} \|u_k - u\|_{W_0^{1, M_1}} \rightarrow 0. \tag{III.14}$$

By (F_1) and Höder's inequality, we get

$$\begin{aligned} \left| \int_{\Omega} F_u(x, u_k, v_k)(u_k - u) dx \right| & \\ \leq 2c_1 \|1 + \psi_1(|u_k|) + \bar{\Psi}_1^{-1}(\Psi_2(|v_k|))\|_{\bar{\Psi}_1} \|u_k - u\|_{\Psi_1}. \end{aligned} \tag{III.15}$$

Condition (F_1) shows that functions Ψ_1 and $\bar{\Psi}_1$ are N-functions satisfying Δ_2 -condition globally, which together with the convexity of N-function, Lemma 2.6 and the boundedness of $\{(u_k, v_k)\}$, imply that

$$\begin{aligned} &\int_{\Omega} \bar{\Psi}_1(1 + \psi_1(|u_k|)) + \bar{\Psi}_1^{-1}(\Psi_2(|v_k|)) dx \\ &\leq C \int_{\Omega} (\Psi_1(u_k) + \Psi_2(v_k)) dx \leq C, \end{aligned}$$

which, together with Lemma 2.6 again, shows that

$$\|1 + \psi_1(|u_k|) + \bar{\Psi}_1^{-1}(\Psi_2(|v_k|))\|_{\bar{\Psi}_1} \leq C, \tag{III.16}$$

for some $C > 0$. Moreover, (e) and (f) shows that

$$\|u_k - u\|_{\Psi_1} \rightarrow 0. \tag{III.17}$$

then, combining (III.13), (III.14), (III.15), (III.16) and (III.17) we obtain

$$\langle \mathcal{H}_1, u_k - u \rangle \rightarrow 0 \text{ as } n \rightarrow \infty.$$

or \mathcal{H} is of the class (S+), then $u_k \rightarrow u$ in $W^{1, M_1}(\Omega)$ and $v_k \rightarrow v$ in $W^{1, M_2}(\Omega)$ Therefore $(u_k, v_k) \rightarrow (u, v)$ in W .

Proof 3.16 (Proof of theorem 3.1): By Lemmas 3.10, 2.13, 3.14 and the obvious fact $I(0) = 0$, all conditions of Lemma 3.7 hold. Then system (I.1) possesses a nontrivial weak solution which is a critical point of I .

REFERENCES

- [1] R. A. Adams, J. F. Fournier, Sobolev spaces, Second edition, Pure and Applied Mathematics (Amsterdam), Elsevier/Academic Press, Amsterdam, (2003).
- [2] K. Adriouch, A. El Hamidi, The Nehari manifold for systems of nonlinear elliptic equations, *Nonlinear Anal.*, 64 (2006), 2149-2167.
- [3] G. A. Afrouzi, S. Heidarkhani, Existence of three solutions for a class of Dirichlet quasilinear elliptic systems involving the (p_1, \dots, p_n) -Laplacian, *Nonlinear Anal.*, 70 (2009), 135-143.
- [4] L. Boccardo, D. Guedes de Figueiredo, Some remarks on a system of quasilinear elliptic equations, *NoDEA Nonlinear Differential Equations Appl.*, 9 (2002), 309-323.
- [5] L. Boccardo, F. Murat Almost everywhere convergence of the gradients of solutions to elliptic and parabolic equations, *Nonlinear Anal.* 19(6), (1992), 581–597.
- [6] O. Claudiandor, L. Alves and M. Marcos. A Lieb type result and application involving a class of non-reflexive Orlicz-Sobolev space.
- [7] L. Diening, Theoretical and numerical results for electrorheological fluids, Ph.D. thesis, University of Freiburg, Germany (2002)
- [8] T. K. Donaldson and N. S. Trudinger, Orlicz-Sobolev Spaces and imbedding theorems. *J. Functional Analysis*, 8 (1971), 52-75.
- [9] N. Fukagai, M. Ito, K. Narukawa, Positive solutions of quasilinear elliptic equations with critical Orlicz-Sobolev nonlinearity on \mathbb{R}^N , *Funkcial. Ekvac.*, 49 (2006), 235-267
- [10] A. Kufner, O. John and S. Fucik, *Function Spaces*, Noordhoff, Leyden, 2013.
- [11] M. A. Krasnosel'skii and Ya. B. Rutickii, *Convex functions and Orlicz Spaces*. Noordhoff. Groningen, 1969.
- [12] T. C. Halsey, *Electrorheological fluids*, *Science*, 258 (1992), 761-766.
- [13] J. Huentutripay, R. Manásevich, Nonlinear eigenvalues for a quasilinear elliptic system in Orlicz-Sobolev Spaces, *J. Dynam. Differential Equations*, 18 (2006), 901-929.
- [14] P. H. Rabinowitz, *Minimax methods in critical point theory with applications to differential equations*, CBMS Regional Conference Series in Mathematics, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, (1986).
- [15] L. Wang, X. Zhang, H. Fang, Existence and multiplicity of solutions for a class of quasilinear elliptic systems in Orlicz-Sobolev Spaces. *J. Nonlinear Sci. Appl.*, 10 (2017), 3792-3814.

Entropy solutions for a nonlinear parabolic problem with lower order terms in Musielak-Orlicz spaces

1st SABIKI HAJAR

*Systems Engineering Laboratory
Sultan Moulay Slimane University*

National School of Business and Management Beni Mellal, Morocco.
sabikihajar@gmail.com

Abstract—We establish an approximation and compactness results in inhomogeneous Musielak-Orlicz-Sobolev spaces, then we shall give the proof of existence results for the entropy solutions of the following nonlinear parabolic problem

$$\begin{cases} \frac{\partial u}{\partial t} - \operatorname{div}(a(x, t, u, \nabla u)) - \operatorname{div}(\Phi(x, t, u)) = f & \text{in } Q_T \\ u(x, 0) = u_0(x) & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \times (0, T). \end{cases}$$

Where $Q_T = \Omega \times (0, T)$ and the growth and the coercivity conditions on the monotone vector field a are prescribed by a generalized N -function M . We assume any restriction on M , therefore we work with Musielak-Orlicz spaces which are not necessarily reflexive. The lower order term $\Phi : \Omega \times (0, T) \times \mathbb{R} \rightarrow \mathbb{R}^N$ is a Carathéodory function, for a.e. $(x, t) \in Q_T$ and for all $s \in \mathbb{R}$, satisfying only a growth condition and the right hand side f belongs to $L^1(Q_T)$.

Index Terms—Non-linear Parabolic problems; Musielak-Orlicz spaces; Entropy Solutions; Non-coercive Problems; Lower order term.

I. INTRODUCTION

In the last decade, there has been an increasing interest in the study of various mathematical problems in modular spaces. These problems have many consideration in applications (see [14], [38], [41]) and have resulted in a renewal interest in Lebesgue and Sobolev spaces with variable exponent, Musielak, Orlicz space, the origins of which can be traced back to the work of Orlicz in the 1930s. In the 1950s, this study was carried on by Nakano [34] who made the first systematic study of spaces with variable exponent. Later, Polish and Czechoslovak mathematicians investigated the modular function spaces (see for example Musielak [32], Kovacik and Rakosnik [26]). The study of variational problems where the function $a(\cdot)$ satisfies the non-polynomial growth conditions instead of having the usual p -structure arouses much interest with the development of applications to electro-rheological fluids as an important class of non-Newtonian fluids (sometimes referred to as smart fluids). The electro-rheological fluids are characterized by their ability to drastically change the mechanical properties under the influence of an external electromagnetic field. A mathematical model of electro-rheological fluids was proposed by Rajagopal and Ruzicka (we refer to [37], [38] for more details). Another important application is related to image processing [39] where

this kind of the diffusion operator is used to underline the borders of the distorted image and to eliminate the noise.

In point of mathematical physics view, it is hard task to show the existence of classical solutions, i.e., solutions which are continuously differentiable as many times as the order of derivatives in equations under consideration. However, the concept of weak solutions is not enough to give a formulation to all problems and does not provide uniqueness and stability properties. Hence, as a certain more general idea, we can use the notion of entropy solution which we have to assume in addition to the weak formulation of the problem certain inequalities.

In this work, we deal with the existence result of the entropy solutions for the following nonlinear parabolic problem without assuming any restriction on the N -function M

$$\begin{cases} \frac{\partial u}{\partial t} - \operatorname{div}(a(x, t, u, \nabla u)) - \operatorname{div}(\Phi(x, t, u)) = f & \text{in } Q_T \\ u(x, 0) = u_0(x) & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \times (0, T), \end{cases} \quad (1)$$

where the data f belongs to $L^1(Q_T)$, $Au = -\operatorname{div}(a(x, t, u, \nabla u))$ is a Leray-Lions operator defined on $W_0^{1,x}L_M(Q_T)$. The lower order term $\Phi : \Omega \times (0, T) \times \mathbb{R} \rightarrow \mathbb{R}^N$ is a Carathéodory function, for a.e. $(x, t) \in Q_T$ and for all $s \in \mathbb{R}$, satisfying only a growth condition and not necessarily coercive.

The notion of renormalized solution has been introduced by Lions and Di Perna [15] for the study of Boltzmann equation (see also P.-L. Lions [29] for a few applications to fluid mechanics models). This notion was then adapted to elliptic version by Boccardo, J.-L. Diaz, D. Giachetti, F.Murat [13] and F. Murat [31]. At the same the equivalent notion of entropy solutions has been developed independently by Bénylan and al. [11] for the study of nonlinear elliptic problems.

The study of the parabolic equations in Orlicz spaces have been a topic for many years, starting from the work of Donaldson [16] and with later results of Benkirane, Elmahi and Meskine, (see [7], [17], [18]). All of them concern the case of classical spaces, namely Orlicz spaces with an N -function dependent only on x without the dependence on (t, x) . We prove our result without any restriction on the growth of an N -function, in particular the Δ_2 -condition for an N -function and its conjugate. This results in a need of

formulating the approximation theorem and extensively using the notion of modular convergence. The fundamental studies in this direction are due to Gossez for the case of elliptic equations [20], [21]. The appearance of (x, t) -dependence in an N -function requires the studies on the uniform boundedness of the convolution operator. Existence of entropy solution with L^1 -data has been proved by Leone and Porretta in [28] for the Dirichlet problem associated to the nonlinear elliptic equation $-\operatorname{div}(a(x, u, \nabla u)) = f$ in the classical Sobolev spaces $W^{1,p}(\Omega)$. In [36] the existence and uniqueness of entropy solutions of the problem (1) has been studied by Prignet where $\Phi = 0$ and $Au = -\operatorname{div}(a(x, t, u, \nabla u))$ is a Leray-Lions operator in divergence form acting on $W^{1,p}(\Omega)$. The existence of renormalized solutions of the problem (1) in Orlicz spaces has been proved in [24].

As far as we know, there's not much papers concerned with the nonlinear parabolic equations with obstacle in Musielak-Orlicz spaces with L^1 data, in the context of renormalized solution we refer to the work of Gwiazda, Wittbold and al. in [22] where the existence proof related to a nonlinear parabolic problem with L^1 -data in Musielak spaces requires a very technical construction of multistage approximation of the solution. In particular it is based on nonlinear semi-group theory of m -accretive operators, but the authors assume that \bar{M} satisfies the Δ_2 -condition and the proof was based on the modular Poincaré inequality, we refer also to [23] for the elliptic case without Δ_2 -condition on M .

Other difficulties associated with the existence of entropy solutions of the problem (1) lie in the fact that the term $\operatorname{div}(\Phi(x, t, u))$ can not be managed by the divergence theorem and the general Musielak function M does not have to satisfy the suitable condition Δ_2 which induces a loss of reflexivity of our framework setting.

Our main goal of this paper is to prove the existence of an entropy solution of the problem (1) in the sense of Definition 5.1. (see section IV) for a general N -function M .

II. PRELIMINARY

In this section we list briefly some definitions and facts about Musielak-Orlicz-Sobolev spaces. A standard reference is [32]. We also include the definition of inhomogeneous Musielak-Orlicz-Sobolev spaces and some preliminaries lemmas to be used later on this paper.

Musielak-Orlicz spaces: Let Ω be a domain in \mathbb{R}^d , $d \in \mathbb{N}$.

Definition 2.1: Let $M: \Omega \times \mathbb{R} \mapsto \mathbb{R}$ be a function such that:

- (i) For almost all (a. a.) $x \in \Omega$, $M(x, \cdot)$ is an N -function, that is, convex and even in \mathbb{R} , increasing in \mathbb{R}^+ , $M(x, 0) = 0$, $M(x, s) > 0$ for all $s > 0$,

$$\lim_{s \rightarrow 0} \frac{M(x, s)}{s} = 0, \quad \lim_{s \rightarrow \infty} \frac{M(x, s)}{s} = \infty.$$

- (ii) For all $s \in \mathbb{R}$, $M(\cdot, s)$ is a measurable function.

A function $M(x, s)$ which satisfies the conditions (i) and (ii) is called a **Musielak-Orlicz function**, a generalized N -function or a generalized modular function.

From now on, $M: \Omega \times \mathbb{R} \mapsto \mathbb{R}$ will stand for a general Musielak-Orlicz function. In some situations, the growth order with respect to t of two given Musielak-Orlicz functions M and P are comparable. This concept is detailed in the next definition.

Definition 2.2: Let $P: \Omega \times \mathbb{R} \mapsto \mathbb{R}$ be another Musielak-Orlicz function.

- Assume that there exist two constants $\epsilon > 0$ and $s_0 \geq 0$ such that for a. a. $x \in \Omega$ one has $P(x, s) \leq M(x, \epsilon s)$ for all $s \geq s_0$. Then we write $P \prec M$ and we say that M dominates P globally if $s_0 = 0$ and near infinity if $s_0 > 0$.
- We say that P grows essentially less rapidly than M at $s = 0$ (respectively, near infinity), and we write $P \ll M$, if for every positive constant k_0 we have

$$\limsup_{t \rightarrow 0} \sup_{x \in \Omega} \frac{P(x, k_0 s)}{M(x, s)} = 0 \text{ (respectively, } \limsup_{t \rightarrow \infty} \sup_{x \in \Omega} \frac{P(x, k_0 s)}{M(x, s)} = 0).$$

We will also use the following notation: $M_x(s) = M(x, s)$, for a. a. $x \in \Omega$ and all $s \in \mathbb{R}$, and we associate its inverse function with respect to $s \geq 0$, denoted by M_x^{-1} , that is,

$$M_x^{-1}(M(x, s)) = M(x, M_x^{-1}(s)) = s, \text{ for all } s \geq 0.$$

Remark 2.3: It is easy to check that $P \ll M$ near infinity if and only if

$$\lim_{s \rightarrow \infty} \frac{M_x^{-1}(k_0 s)}{P_x^{-1}(s)} = 0 \text{ uniformly for } x \in \Omega \setminus \Omega_0$$

for some null subset $\Omega_0 \subset \Omega$. \square

We introduce the functional $\varrho_{M, \Omega}$ given by

$$\varrho_{M, \Omega}(u) = \int_{\Omega} M(x, u(x)) \, dx,$$

for any Lebesgue measurable function $u: \Omega \mapsto \mathbb{R}$. The set

$$\mathcal{L}_M(\Omega) = \{u: \Omega \mapsto \mathbb{R} \text{ measurable such that } \varrho_{M, \Omega}(u) < \infty\}$$

is called the **Musielak-Orlicz class** related to M in Ω or simply the Musielak-Orlicz class.

The **Musielak-Orlicz space** $L_M(\Omega)$ is the vector space generated by $\mathcal{L}_M(\Omega)$, that is, $L_M(\Omega)$ is the smallest linear space containing the set $\mathcal{L}_M(\Omega)$. Equivalently,

$$L_M(\Omega) = \{u: \Omega \mapsto \mathbb{R} \text{ measurable such that } \varrho_{M, \Omega}(u/\alpha) < \infty, \text{ for some } \alpha > 0\}.$$

For a Musielak-Orlicz function M , we introduce its **complementary function**, denoted by \bar{M} , as

$$\bar{M}(x, r) = \sup_{s \geq 0} \{rs - M(x, s)\},$$

that is $\bar{M}(x, r)$ is the complementary to $M(x, s)$ in the sense of Young with respect to the variable r . It turns out that \bar{M} is another Musielak-Orlicz function and the following Young-Fenchel inequality holds

$$|sr| \leq M(x, s) + \bar{M}(x, r) \text{ for all } s, r \in \mathbb{R} \text{ and a. a. } x \in \Omega. \tag{2}$$

In the space $L_M(\Omega)$ we define the following two norms:

$$\|u\|_{M,\Omega} = \inf \left\{ \lambda > 0 / \int_{\Omega} M(x, u(x)/\lambda) dx \leq 1 \right\},$$

which is called the **Luxemburg norm**, and the so-called **Orlicz norm**, namely

$$\|u\|_{(M),\Omega} = \sup_{\varrho_{M,\Omega}(v) \leq 1} \int_{\Omega} u(x)v(x) dx.$$

where the supremum is taken over all $v \in \mathcal{L}_{\bar{M}}(\Omega)$ such that $\varrho_{\bar{M},\Omega}(v) \leq 1$. An important inequality in $L_M(\Omega)$ is the following:

$$\int_{\Omega} M(x, u(x)) dx \leq \|u\|_{(M),\Omega} \tag{3}$$

for all $u \in L_M(\Omega)$ such that $\|u\|_{(M),\Omega} \leq 1$, where from we readily deduce

$$\int_{\Omega} M\left(x, \frac{u(x)}{\|u\|_{(M),\Omega}}\right) dx \leq 1 \text{ for all } u \in L_M(\Omega) \setminus \{0\}. \tag{4}$$

From the definition of the Orlicz norm and (2) it is easy to obtain the inequality

$$\|u\|_{(M),\Omega} \leq 1 + \int_{\Omega} M(x, u(x)) dx, \text{ for all } u \in L_M(\Omega). \tag{5}$$

It can be shown that the norm $\|\cdot\|_{(M),\Omega}$ is equivalent to the Luxemburg norm $\|\cdot\|_{M,\Omega}$. Indeed,

$$\|u\|_{M,\Omega} \leq \|u\|_{(M),\Omega} \leq 2\|u\|_{M,\Omega} \text{ for all } u \in L_M(\Omega). \tag{6}$$

Also, Hölder's inequality holds

$$\int_{\Omega} |u(x)v(x)| dx \leq \|u\|_{M,\Omega} \|v\|_{\bar{M},\Omega}$$

for all $u \in L_M(\Omega)$ and $v \in L_{\bar{M}}(\Omega)$. Most properties verified by the classical Orlicz spaces cannot be extended to the Musielak-Orlicz spaces unless we assume certain supplementary hypotheses on the generalized N -function M . To this end, we first introduce the two following assumptions.

$$\varrho_{M,\Omega}(\lambda\chi_K) < \infty \text{ for any } \lambda \geq 0 \text{ and any compact set } K \subset \bar{\Omega}. \tag{7}$$

$$\left\{ \begin{array}{l} \text{There exist two positive constants } \lambda_0 \text{ and } c_0 \text{ such that} \\ \text{ess inf}_{\Omega} M(x, \lambda_0) \geq c_0. \end{array} \right. \tag{8}$$

In (7), χ_A stands for the characteristic function of a measurable set A . The assumption (7) assures that any bounded measurable function with compact support in $\bar{\Omega}$ is in the class $\mathcal{L}_M(\Omega)$. In this situation, we may introduce the space $E_M(\Omega)$ as the closure in $L_M(\Omega)$ of the bounded measurable functions with compact support in $\bar{\Omega}$. The space $E_M(\Omega)$ is then the largest linear space such that $E_M(\Omega) \subset \mathcal{L}_M(\Omega)$, this inclusion being in general strict. Notice that if Ω is bounded then (7) implies the inclusion $L^\infty(\Omega) \subset \mathcal{L}_M(\Omega)$.

On the other hand, the assumption (8) implies that any function in $L_M(\Omega)$ is locally integrable in Ω . This is stated in the following result.

Lemma 2.4: Assume (8). Then the inclusion $L_M(\Omega) \subset L^1_{loc}(\Omega)$ holds true. Moreover, if $|\Omega| \stackrel{\text{def}}{=} \text{meas}(\Omega) < \infty$, then $L_M(\Omega) \subset L^1(\Omega)$ with continuous inclusion, that is

$$\|u\|_{L^1(\Omega)} \leq C_1 \|u\|_{(M),\Omega} \text{ for all } u \in L_M(\Omega),$$

where $C_1 = \lambda_0(|\Omega| + 1/c_0)$.

Proof 2.5: According to the convexity of $M(x, \cdot)$ we obtain

$$sM(x, \lambda_0) \leq \lambda_0 M(x, s) \text{ for all } s \geq \lambda_0 \text{ and a. a. } x \in \Omega.$$

Let $u \in L_M(\Omega)$ and $A \subset \Omega$ a measurable set with $|A| < \infty$. Take $\alpha > 0$ such that $\varrho_{M,\Omega}(u/\alpha) < \infty$. Then,

$$\begin{aligned} \int_A \left| \frac{u}{\alpha} \right| &= \int_{A \cap \{|u| < \alpha\lambda_0\}} \left| \frac{u}{\alpha} \right| + \int_{A \cap \{|u| \geq \alpha\lambda_0\}} \left| \frac{u}{\alpha} \right| \\ &\leq \lambda_0 |A| + \frac{1}{c_0} \int_{A \cap \{|u| \geq \alpha\lambda_0\}} \left| \frac{u}{\alpha} \right| M(x, \lambda_0) \\ &\leq \lambda_0 |A| + \frac{\lambda_0}{c_0} \int_{\Omega} M\left(x, \frac{u}{\alpha}\right) < \infty, \end{aligned}$$

and thus $u \in L^1(A)$. If $|\Omega| < \infty$, we may take $A = \Omega$ and $\alpha = \|u\|_{(M),\Omega}$ in the estimate above. Using (4) it yields

$$\int_{\Omega} |u| \leq \lambda_0 \left(|\Omega| + \frac{1}{c_0} \right) \|u\|_{(M),\Omega}.$$

From now on, we will assume both assumptions (7) and (8) in this paper.

Strong convergence in $L_M(\Omega)$ is rather strict. For most purposes, a mild concept of convergence will be enough, namely, that of **modular convergence**.

Definition 2.6: We say that a sequence $(u_n) \subset L_M(\Omega)$ is modular convergent to $u \in L_M(\Omega)$ if there exists a constant $\lambda > 0$ such that

$$\lim_{n \rightarrow \infty} \varrho_{M,\Omega}((u_n - u)/\lambda) = 0.$$

Musielak-Orlicz-Sobolev spaces: According to Lemma 2.4, any function in $L_M(\Omega)$ is locally integrable and, in particular, may be considered as a distribution. This allows us to introduce the so-called Musielak-Orlicz-Sobolev spaces. For any fixed nonnegative integer m we define

$$W^m L_M(\Omega) = \{u \in L_M(\Omega) / D^\alpha u \in L_M(\Omega) \text{ for all } \alpha, |\alpha| \leq m\}$$

where $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_d) \in \mathbb{Z}$, $\alpha_j \geq 0$, $j = 1, \dots, d$, with $|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_d$ and $D^\alpha u$ denote the distributional derivative of multiindex α . The space $W^m L_M(\Omega)$ is called the Musielak-Orlicz-Sobolev space (of order m).

Let $u \in W^m L_M(\Omega)$, we define $\varrho_{M,\Omega}^{(m)}(u) = \sum_{|\alpha| \leq m} \varrho_{M,\Omega}(D^\alpha u)$, and

$$\|u\|_{M,\Omega}^{(m)} = \inf \{ \lambda > 0 / \varrho_{M,\Omega}^{(m)}(u/\lambda) \leq 1 \},$$

$$\|u\|_{m,M,\Omega} = \sum_{|\alpha| \leq m} \|D^\alpha u\|_{M,\Omega}.$$

The functional $\varrho_{M,\Omega}^{(m)}$ is convex in $W^m L_M(\Omega)$, whereas the functionals $\|\cdot\|_{M,\Omega}^{(m)}$ and $\|\cdot\|_{m,M,\Omega}$ are equivalent norms on $W^m L_M(\Omega)$. The pair $(W^m L_M(\Omega), \|\cdot\|_{M,\Omega}^{(m)})$ is a Banach space under the assumption (8).

The space $W^m L_M(\Omega)$ is identified to a subspace of the product $\Pi_{|\alpha|\leq m} L_M(\Omega) = \Pi L_M$, this subspace is $\sigma(\Pi L_M, \Pi E_{\bar{M}})$ closed.

Let $W_0^m L_M(\Omega)$ be the $\sigma(\Pi L_M, \Pi E_{\bar{M}})$ closure of $\mathcal{D}(\Omega)$ in $W^m L_M(\Omega)$. Let $W^m E_M(\Omega)$ be the space of functions u such that u and its distribution derivatives up to order m lie in $E_M(\Omega)$, and $W_0^m E_M(\Omega)$ is the (norm) closure of $\mathcal{D}(\Omega)$ in $W^m L_M(\Omega)$.

Since we are going to work with two generalized N -functions, say P and M , such that $P \ll M$, we will consider the following assumptions for both complementary functions \bar{P} and \bar{M} :

$$\lim_{|\xi| \rightarrow \infty} \operatorname{ess\,inf}_{x \in \Omega} \frac{\bar{M}(x, \xi)}{|\xi|} = \infty, \tag{9}$$

and

$$\lim_{|\xi| \rightarrow \infty} \operatorname{ess\,inf}_{x \in \Omega} \frac{\bar{P}(x, \xi)}{|\xi|} = \infty. \tag{10}$$

Remark 2.7: From Remark 2.1 in [22] we have that the assumptions (9) and (10) provide the following:

$$\sup_{\xi \in B(0,R)} \operatorname{ess\,sup}_{x \in \Omega} M(x, \xi) < \infty \text{ for all } 0 < R < +\infty, \tag{11}$$

and

$$\sup_{\xi \in B(0,R)} \operatorname{ess\,sup}_{x \in \Omega} P(x, \xi) < \infty \text{ for all } 0 < R < +\infty. \tag{12}$$

Also notice that (11) implies (7).

Definition 2.8: We say that a sequence $(u_n) \subset W^1 L_M(\Omega)$ converges to $u \in W^1 L_M(\Omega)$ for the **modular convergence** in $W^1 L_M(\Omega)$ if, for some $h > 0$,

$$\lim_{n \rightarrow \infty} \varrho_{M,\Omega}^{(1)}((u_n - u)/h) = 0.$$

The following spaces of distributions will also be used:

$$W^{-1} L_{\bar{M}}(\Omega) = \left\{ f \in \mathcal{D}'(\Omega) / f = \sum_{|\alpha| \leq 1} (-1)^{|\alpha|} D^\alpha f_\alpha \right. \\ \left. \text{for some } f_\alpha \in L_{\bar{M}}(\Omega) \right\}$$

and

$$W^{-1} E_{\bar{M}}(\Omega) = \left\{ f \in \mathcal{D}'(\Omega) / f = \sum_{|\alpha| \leq 1} (-1)^{|\alpha|} D^\alpha f_\alpha \right. \\ \left. \text{for some } f_\alpha \in E_{\bar{M}}(\Omega) \right\}.$$

Lemma 2.9: If $P \ll M$ and $u_n \rightarrow u$ for the modular convergence in $L_M(\Omega)$, then $u_n \rightarrow u$ strongly in $E_P(\Omega)$. In particular, $L_M(\Omega) \subset E_P(\Omega)$ and $L_{\bar{P}}(\Omega) \subset E_{\bar{M}}(\Omega)$ with continuous injections.

Proof 2.10: Let $\epsilon > 0$ be given. Let $\lambda > 0$ be such that

$$\int_{\Omega} M\left(x, \frac{u_n - u}{\lambda}\right) \rightarrow 0, \text{ as } n \rightarrow \infty.$$

Therefore, there exists $h \in L^1(\Omega)$ such that

$$M\left(x, \frac{u_n - u}{\lambda}\right) \leq h \text{ and } u_n \rightarrow u \text{ a. e. in } \Omega$$

for a sub-sequence still denoted (u_n) . Since $P \ll M$, then for all $r > 0$ there exists $t_0 > 0$ such that

$$\frac{P(x, rt)}{M(x, t)} \leq 1, \text{ a. e. in } \Omega \text{ and for all } t \geq t_0.$$

For $r = \frac{\lambda}{\epsilon}$ and $t = \frac{t'}{\lambda}$, we get

$$\frac{P(x, \frac{t'}{\epsilon})}{M(x, \frac{t'}{\lambda})} \leq 1, \text{ when } t' \geq t_0 \lambda.$$

Then

$$P\left(x, \frac{u_n - u}{\epsilon}\right) \leq M\left(x, \frac{u_n - u}{\lambda}\right) + \sup_{t' \in B(0, t_0 \lambda)} \operatorname{ess\,sup}_{x \in \Omega} P(x, t'/\epsilon) \\ \leq h + \sup_{t' \in B(0, t_0)} \operatorname{ess\,sup}_{x \in \Omega} P(x, t'/\epsilon) \text{ for a. a. } x \in \Omega.$$

Since $h + \sup_{t' \in B(0, t_0 \lambda)} \operatorname{ess\,sup}_{x \in \Omega} P(x, \frac{t'}{\epsilon}) \in L^1(\Omega)$ (from Remark 2.7), it yields, by the Lebesgue dominated convergence theorem,

$$P\left(x, \frac{u_n - u}{\epsilon}\right) \rightarrow 0 \text{ in } L^1(\Omega),$$

hence, for n big enough, we have $\|u_n - u\|_{P,\Omega} \leq \epsilon$. That is, $u_n \rightarrow u$ in $L_P(\Omega)$.

The continuous injection $L_M(\Omega) \subset E_P(\Omega)$ is trivial since the convergence in $L_M(\Omega)$ implies the modular convergence in this space. On the other hand, since $P \ll M$ is equivalent to $\bar{M} \ll \bar{P}$, this yields the continuous injection $L_{\bar{P}}(\Omega) \subset E_{\bar{M}}(\Omega)$.

Lemma 2.11: (Lemma 2.2 in [30]) Let $(w_n) \subset L_M(\Omega)$, $w \in L_M(\Omega)$, $(v_n) \subset L_{\bar{M}}(\Omega)$ and $v \in L_{\bar{M}}(\Omega)$. If $w_n \rightarrow w$ in $L_M(\Omega)$ for the modular convergence and $v_n \rightarrow v$ in $L_{\bar{M}}(\Omega)$ for the modular convergence, then

$$\lim_{n \rightarrow \infty} \int_{\Omega} w_n v \, dx = \int_{\Omega} w v \, dx \text{ and } \lim_{n \rightarrow \infty} \int_{\Omega} w_n v_n \, dx = \int_{\Omega} w v \, dx.$$

Lemma 2.12: Let $(u_n) \subset E_M(\Omega)$ with $u_n \rightarrow u$ in $E_M(\Omega)$. Then there exist $h \in \mathcal{L}_M(\Omega)$ and a subsequence $(u_{n'})$ such that (a. e. stands for ‘almost everywhere’)

$$|u_{n'}(x)| \leq h(x) \text{ a. e. in } \Omega, \text{ and } u_{n'} \rightarrow u(x) \text{ a. e. in } \Omega.$$

Proof 2.13: If $u_{n'} = u$ for some subsequence $(u_{n'})$, then the result is trivial. Thus, we may assume that for $n \geq 1$ large enough (and some subsequence, if necessary, still denoted in the same way) it is $0 < 2\|u_n - u\|_{(M)} \leq 1$. Then

$$\|M_x(2(u_n - u))\|_{L^1(\Omega)} = \int_{\Omega} M_x\left(2\|u_n - u\|_{(M)} \frac{u_n - u}{\|u_n - u\|_{(M)}}\right) \\ \leq 2\|u_n - u\|_{(M)} \int_{\Omega} M_x\left(\frac{u_n - u}{\|u_n - u\|_{(M)}}\right) \\ \leq 2\|u_n - u\|_{(M)}.$$

Thus, $\|M_x(2(u_n - u))\|_{L^1(\Omega)} \rightarrow 0$, as $n \rightarrow \infty$. Therefore there exists $h_1 \in L^1(\Omega)$ and a subsequence $(u_{n'})$ such that

$u_{n'} \rightarrow u(x)$ a. e. in Ω and $M_x(2(u_{n'}(x) - u(x))) \leq h_1(x)$ or
a. e. in Ω , which implies that

$$|u_{n'}| \leq |u(x)| + \frac{1}{2} M_x^{-1}(h_1(x)) \stackrel{\text{def}}{=} h(x).$$

Since

$$\int_{\Omega} M_x \left(|u(x)| + \frac{1}{2} M_x^{-1}(h_1(x)) \right) \leq \frac{1}{2} \int_{\Omega} M_x(2u(x)) + \frac{1}{2} \int_{\Omega} h_1(x) < \infty,$$

we finally obtain $h \in \mathcal{L}_M(\Omega)$.

Lemma 2.14: (Cf. [4]) Let Ω be a bounded and Lipschitz-continuous domain in \mathbb{R}^d and let M and \bar{M} be two complementary Musielak-Orlicz functions in $\Omega \times \mathbb{R}$ which satisfy the following conditions:

(i) There exists a constant $A > 0$ such that for all $x, y \in \Omega$ with $0 < |x - y| \leq \frac{1}{2}$ one has

$$\frac{M(x, s)}{M(y, s)} \leq s^{-\frac{A}{\log|x-y|}} \text{ for all } s \geq 1. \quad (13)$$

(ii) There exists a constant $C > 0$ such that

$$\bar{M}(x, 1) \leq C \text{ a. e. in } \Omega. \quad (14)$$

Then the space $\mathcal{D}(\Omega)$ is dense in $L_M(\Omega)$ with respect to the modular convergence, $\mathcal{D}(\Omega)$ is dense in $W_0^1 L_M(\Omega)$ for the modular convergence and $\mathcal{D}(\Omega)$ is dense in $W^1 L_M(\Omega)$ for the modular convergence.

Remark 2.15: By taking $s = 1$ in (13) it yields that $M(x, 1) = \text{constant}$ for a. a. $x \in \Omega$. In particular, the condition (8) is obviously verified and also

$$\int_{\Omega} M(x, 1) dx < \infty.$$

Remark 2.16: (Cf. [9]) Let $p: \Omega \mapsto (1, \infty)$ be a measurable function such that there exists a constant $A > 0$ such that for all points $x, y \in \Omega$ with $|x - y| < 1/2$, one has the inequality

$$|p(x) - p(y)| \leq \frac{A}{\log|x-y|}.$$

Then the following Musielak-Orlicz functions satisfy the assumption (13):

- 1) $M(x, s) = s^{p(x)}$;
- 2) $M(x, s) = s^{p(x)} \log(1 + s)$;
- 3) $M(x, s) = s \log(1 + s) (\log(e - 1 + s))^{p(x)}$.

Poincaré's inequality does not hold in generalized Orlicz-Sobolev spaces unless the Musielak-Orlicz function $M(x, s)$ verifies some structural assumption. To this end, we introduce the following definition [3].

Definition 2.17: A generalized function $M(x, s)$ is said to satisfy the Y -condition on a non-empty bounded interval $(a, b) \subset \mathbb{R}$, if either

$$(Y_0) \begin{cases} \text{there exist } s_0 \geq 0 \text{ and } 1 \leq i \leq N \text{ such that the function} \\ x_i \in (a, b) \mapsto M(x, s) \text{ changes constantly its} \\ \text{monotony on both sides of } s_0 \text{ (that is, for } s \geq s_0 \\ \text{and } 0 \leq s < s_0), \end{cases}$$

$$(Y_{\infty}) \begin{cases} \text{there exists } 1 \leq i \leq N \text{ such that for all } s \geq 0, \\ \text{the function } x_i \in (a, b) \mapsto M(x, s) \\ \text{is monotone on } (a, b). \end{cases}$$

Here, x_i stands for the i -th component of $x \in \Omega$.

Lemma 2.18: (Poincaré's inequality [3]) Let Ω be a bounded and Lipschitz-continuous domain in \mathbb{R}^d and let M and \bar{M} be two complementary Musielak-Orlicz functions in $\Omega \times \mathbb{R}$. Assume that M verifies (13) and the Y -condition, and also that \bar{M} verifies (7) and (14). Then there exists a constant $C_0 = C_0(\Omega, M) > 0$ such that

$$\|u\|_{M, \Omega} \leq C_0 \|\nabla u\|_{M, \Omega}, \text{ for all } u \in W_0^1 L_M(\Omega). \quad (15)$$

From this point on we will always assume that the hypothesis of Lemma 2.18 hold true.

Remark 2.19: Let M be a Musielak-Orlicz function such that (15) is verified and let $u \in W_0^1 L_M(\Omega)$. Assume that, for some constant $C \geq 0$, one has $\int_{\Omega} M(x, \nabla u) dx \leq C$. Then, $\|u\|_{1, M, \Omega} \leq C'$ where $C' = (C_0 + 1) \max(C, 1)$. Indeed, since $\|u\|_{1, M, \Omega} = \|u\|_{M, \Omega} + \|\nabla u\|_{M, \Omega}$, by using (15), we get

$$\|u\|_{1, M, \Omega} \leq C_0 \|\nabla u\|_{M, \Omega} + \|\nabla u\|_{M, \Omega} \leq (C_0 + 1) \|\nabla u\|_{M, \Omega}.$$

Now, if $C \geq 1$, according to the convexity of $M(x, \cdot)$, it yields

$$\int_{\Omega} M\left(x, \frac{\nabla u}{C}\right) dx \leq \frac{1}{C} \int_{\Omega} M(x, \nabla u) dx \leq \frac{C}{C} = 1,$$

this means that $C \in \{\lambda > 0, \int_{\Omega} M(x, \nabla u/\lambda) dx \leq 1\}$, hence $\|\nabla u\|_{M, \Omega} \leq C$. On the other hand, if $C < 1$, then $\int_{\Omega} M(x, \nabla u) dx \leq C < 1$, which yields $\|\nabla u\|_{M, \Omega} \leq 1$.

Inhomogeneous Musielak-Orlicz-Sobolev spaces. When dealing with parabolic equations in the context of Musielak-Orlicz-Sobolev spaces we need to introduce some particular spaces which take into account the different orders of differentiation with respect to the spatial variables and the time variable.

Let Ω be a bounded and open subset of \mathbb{R}^d and let $Q_T = \Omega \times (0, T)$ for some $T > 0$. Let $M = M(x, s)$ be a Musielak-Orlicz function in $\Omega \times \mathbb{R}$ (here we do not consider a more general case where $M = M(x, t, s)$, $(x, t) \in Q_T$). For each $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{Z}^d$, $\alpha_j \geq 0$, $j = 1, \dots, d$, we denote by D_x^α the distributional derivative on Q_T of multiindex α with respect to the variable $x \in \mathbb{R}^d$. The inhomogeneous Musielak-Orlicz-Sobolev spaces of order one are defined as follows:

$$W^{1, x} L_M(Q_T) = \{u \in L_M(Q_T) / D_x^\alpha u \in L_M(Q_T) \text{ for all } \alpha, |\alpha| \leq 1\}$$

and

$$W^{1, x} E_M(Q_T) = \{u \in E_M(Q_T) / D_x^\alpha u \in E_M(Q_T) \text{ for all } \alpha, |\alpha| \leq 1\}$$

This last space is a subspace of the first one, and both are Banach spaces under the assumption (8) and with norm

$$\|u\| = \sum_{|\alpha| \leq 1} \|D_x^\alpha u\|_{M, Q_T}.$$

These spaces are considered as subspaces of the product space $\Pi L_M(Q_T)$ which has $(d + 1)$ copies.

We shall also consider the weak-* topologies $\sigma(\Pi L_M(Q_T), \Pi E_{\bar{M}}(Q_T))$ and $\sigma(\Pi L_M(Q_T), \Pi L_{\bar{M}}(Q_T))$. If $u \in W^{1,x}L_M(Q_T)$ then the function $t \rightarrow u(t)$ is defined on $(0, T)$ with values in $W^1L_M(\Omega)$. If, further, $u \in W^{1,x}E_M(Q_T)$ then this function is a $W^1E_M(\Omega)$ -valued and is strongly measurable. The space $W^{1,x}L_M(Q_T)$ is not in general separable. If $u \in W^{1,x}L_M(Q_T)$, we cannot conclude that the function $u(t)$ is measurable on $(0, T)$. However, the scalar function $t \rightarrow \|u(t)\|_{M,\Omega}$ is in $L^1(0, T)$. The space $W_0^{1,x}E_M(Q_T)$ is defined as the (norm) closure in $W^{1,x}E_M(Q_T)$ of $\mathcal{D}(Q)$. We can easily show as in [4] that when Ω is a Lipschitz-continuous domain then each element u of the closure of $\mathcal{D}(Q_T)$ with respect of the weak-* topology $\sigma(\Pi L_M, \Pi E_{\bar{M}})$ is limit, in $W^{1,x}L_M(Q_T)$, of some subsequence $(u_n) \subset \mathcal{D}(Q_T)$ for the modular convergence; i. e., there exists $\lambda > 0$ such that for all α with $|\alpha| \leq 1$

$$\int_{Q_T} M\left(x, \frac{D_x^\alpha u_n - D_x^\alpha u}{\lambda}\right) dx dt \rightarrow 0 \text{ as } n \rightarrow \infty,$$

and, in particular, this implies that (u_n) converges to u in $W^{1,x}L_M(Q_T)$ for the weak-* topology $\sigma(\Pi L_M, \Pi L_{\bar{M}})$. Consequently

$$\overline{\mathcal{D}(Q_T)}^{\sigma(\Pi L_M, \Pi L_{\bar{M}})} = \overline{\mathcal{D}(Q_T)}^{\sigma(\Pi L_M, \Pi E_{\bar{M}})}.$$

This space will be denoted by $W_0^{1,x}L_M(Q_T)$. Furthermore,

$$W_0^{1,x}E_M(Q_T) = W_0^{1,x}L_M(Q_T) \cap \Pi E_{\bar{M}}(Q_T).$$

Poincaré's inequality also holds in $W_0^{1,x}L_M(Q_T)$, i. e. there exists a constant $C > 0$ such that for all $u \in W_0^{1,x}L_M(Q_T)$ one has

$$\sum_{|\alpha| \leq 1} \|D_x^\alpha u\|_{M, Q_T} \leq C \sum_{|\alpha|=1} \|D_x^\alpha u\|_{M, Q_T}. \quad (16)$$

The dual space of $W_0^{1,x}E_M(Q_T)$ will be denoted by $W^{-1,x}L_{\bar{M}}(Q_T)$, and it can be shown that

$$W^{-1,x}L_{\bar{M}}(Q_T) = \left\{ f = \sum_{|\alpha| \leq 1} D_x^\alpha f_\alpha / f_\alpha \in L_{\bar{M}}(Q_T), \right\}.$$

for all α .

This space will be equipped with the usual quotient norm

$$\|f\| = \inf \sum_{|\alpha| \leq 1} \|D_x^\alpha f_\alpha\|_{\bar{M}, Q_T}$$

where the infimum is taken over all possible functions $f_\alpha \in L_{\bar{M}}(Q_T)$ from which the decomposition $f = \sum_{|\alpha| \leq 1} D_x^\alpha f_\alpha$ holds true.

We also denote by $W^{-1,x}E_{\bar{M}}(Q_T)$ the subspace of $W^{-1,x}L_{\bar{M}}(Q_T)$ consisting of those linear forms which are $\sigma(\Pi L_M, \Pi E_{\bar{M}})$ -continuous. It can be shown that

$$W^{-1,x}E_{\bar{M}}(Q_T) = \left\{ f = \sum_{|\alpha| \leq 1} D_x^\alpha f_\alpha / f_\alpha \in E_{\bar{M}}(Q_T) \right\}.$$

III. Compactness results

In the sequel, we will make use of the following results which concern mollification with respect to time and space variables and some trace results. For a function $u \in L^1(Q_T)$ we introduce the function $\tilde{u} \in L^1(\Omega \times \mathbb{R})$ as $\tilde{u}(x, s) = u(x, s)\chi_{(0, T)}$ and define, for all $\mu > 0$, $t \in [0, T]$ and a.e. $x \in \Omega$, the function u_μ given as follows

$$u_\mu(x, t) = \mu \int_{-\infty}^t \tilde{u}(x, s) \exp(\mu(s - t)) ds. \quad (17)$$

Lemma 3.1: ([2]).

- 1) Let $u \in L_M(Q_T)$. Then $u_\mu \in C([0, T]; L_M(\Omega))$ and $u_\mu \rightarrow u$ as $\mu \rightarrow +\infty$ in $L_M(Q_T)$ for the modular convergence.
- 2) Let $u \in W^{1,x}L_M(Q_T)$. Then $u_\mu \in C([0, T]; W^1L_M(\Omega))$ and $u_\mu \rightarrow u$ as $\mu \rightarrow +\infty$ in $W^{1,x}L_M(Q_T)$ for the modular convergence.
- 3) Let $u \in E_M(Q_T)$ (respectively, $u \in W^{1,x}E_M(Q_T)$). Then $u_\mu \rightarrow u$ as $\mu \rightarrow +\infty$ strongly in $E_M(Q_T)$ (respectively, strongly in $W^{1,x}E_M(Q_T)$).
- 4) Let $u \in W^{1,x}L_M(Q_T)$ then $\frac{\partial u_\mu}{\partial t} = \mu(u - u_\mu) \in W^{1,x}L_M(Q_T)$.
- 5) Let $(u_n) \subset W^{1,x}L_M(Q_T)$ and $u \in W^{1,x}L_M(Q_T)$ such that $u_n \rightarrow u$ strongly in $W^{1,x}L_M(Q_T)$ (respectively, for the modular convergence). Then, for all $\mu > 0$, $(u_n)_\mu \rightarrow u_\mu$ strongly in $W^{1,x}L_M(Q_T)$ (respectively, for the modular convergence).

Lemma 3.2: [2] Let M be a Musielak function. Let Y be a Banach space such that the following continuous imbedding holds $L^1(\Omega) \subset Y$. Then, for all $\epsilon > 0$ and all $\lambda > 0$ there is C_ϵ such that for all $u \in W^{1,x}L_M(Q_T)$ with $\frac{\nabla u}{\lambda} \in K_M(Q_T)$

$$\|u\|_{L^1(\Omega)} \leq \epsilon \lambda \left(\int_{Q_T} M(x, \frac{\nabla u}{\lambda}) dx dt + T \right) + C_\epsilon \|u\|_{L^1(0, T; Y)}. \quad (18)$$

Lemma 3.3: [2] Let Y be a Banach space such that $L^1(\Omega) \subset Y$ with continuous imbedding.

If F is bounded in $W_0^{1,x}L_M(Q_T)$ and is relatively compact in $L^1(0, T; Y)$ then F is relatively compact in $L^1(Q_T)$.

Lemma 3.4: (cf. [33]) Let $Q_T = \Omega \times (0, T)$, let M a Musielak-Orlicz function, $E_M(\Omega)$ the Musielak-Orlicz space on Ω and $E_M(Q_T)$ the inhomogeneous Musielak-Orlicz space on Q_T . Then there embeddings map

$$E_M(Q_T) \subseteq L^1(0, T; E_M(\Omega)). \quad (19)$$

Lemma 3.5: Let $Q_T = \Omega \times (0, T)$, let M a Musielak-Orlicz function, $W^1E_M(\Omega)$ the Musielak-Orlicz-Sobolev space on Ω and $W^1E_M(Q_T)$ the inhomogeneous Musielak-Orlicz-Sobolev space on Q_T . Then the following embeddings

$$W^1E_M(Q_T) \subset L^1(0, T; W^1E_M(\Omega)) \quad (20)$$

$$W^{-1}E_{\bar{M}}(Q_T) \subset L^1((0, T); W^{-1}E_{\bar{M}}(\Omega)) \quad (21)$$

are continuous

Proof 3.6: Let $u \in W^1 E_M(Q_T)$, we have $u \in E_M(Q_T)$ and $D_x^\alpha u \in E_M(Q_T)$. By the previous lemma, we get

$$\int_0^T \|u\|_{M,\Omega} dt \leq (T+1)\|u\|_{M,Q_T}, \quad (22)$$

and

$$\int_0^T \|D_x^\alpha u\|_{M,\Omega} dt \leq (T+1)\|D_x^\alpha u\|_{M,Q_T} \quad \text{for all } |\alpha| \leq 1, \quad (23)$$

which implies

$$\int_0^T \|u\|_{L^1(0,T;W^1 E_M(\Omega))} dt \leq (T+1)\|u\|_{W^{1,x} E(Q_T)}. \quad (24)$$

Consequently (20) is proved.

Using the same Technics we will prove (21). Since every $f \in W^{-1,x} E_{\overline{M}}(Q_T)$ reads as

$$f = \sum_{|\alpha| \leq 1} D_x^\alpha g_\alpha \quad \text{where } g_\alpha \in E_{\overline{M}}(Q_T)$$

and

$$\|f\|_{W^{-1,x} L_{\overline{M}}(Q_T)} = \sum_{|\alpha| \leq 1} \|g_\alpha\|_{\overline{M},Q_T}.$$

This gives

$$\int_0^T \sum_{|\alpha| \leq 1} \|g_\alpha(t)\|_{\overline{M},\Omega} \leq (1+T)\|f\|_{W^{-1,x} L_{\overline{M}}(Q_T)},$$

by definition of the quotient norm of $W^{-1} L_{\overline{M}}(\Omega)$ we have

$$\|f(t)\|_{W^{-1} L_{\overline{M}}(\Omega)} \leq \sum_{|\alpha| \leq 1} \|g_\alpha(t)\|_{\overline{M},\Omega},$$

and then

$$\int_0^T \|f(t)\|_{W^{-1} L_{\overline{M}}(\Omega)} dt \leq (T+1)\|f\|_{W^{-1,x} L_{\overline{M}}(Q_T)}.$$

This gives the desired result.

Theorem 3.7: [2] Let M be a Musielak function. If F is bounded in $W_0^{1,x} L_M(Q_T)$ and $\frac{\partial f}{\partial t} : f \in F$ is bounded in $W^{-1,x} L_{\overline{M}}(Q_T)$, then F is relatively compact in $L^1(Q)$.

Lemma 3.8: [40] Let B be a Banach space. If $f \in \mathcal{D}'(]0, T[; B)$ is such that $\frac{\partial f}{\partial t} \in L^1(0, T; B)$ then $f \in C(]0, T[; B)$ and for all $h > 0$ we have $\|\tau_h(f) - f\|_{L^1(0,T;B)} \leq h \|\frac{\partial f}{\partial t}\|_{L^1(0,T;B)}$.

Remark 3.9: By the Theorem 3.4, if $F \subset L^1(0, T; B)$ is such that $\left\{ \frac{\partial f}{\partial t} : f \in F \right\}$ is bounded in $L^1(0, T; B)$ then $\|\tau_h(f) - f\|_{L^1(0,T;B)} \rightarrow 0$ as $h \rightarrow 0$ uniformly with respect to $f \in F$.

Corollary 3.10: Let M be a Musielak-Orlicz function. Let (u_n) be a sequence of $W^{1,x} L_M(Q_T)$ such that

$$u_n \rightharpoonup u \quad \text{weakly in } W^{1,x} L_M(Q_T) \quad \text{for } \sigma(II L_M, II E_{\overline{M}})$$

and

$$\frac{\partial u_n}{\partial t} = h_n + k_n \quad \text{in } \mathcal{D}'(Q_T)$$

with (h_n) bounded in $W^{-1,x} L_{\overline{M}}(Q_T)$ and (k_n) bounded in the space $L^1(Q_T)$ of measures on Q_T . Then

$$u_n \rightarrow u \quad \text{strongly in } L_{loc}^1(Q_T).$$

If further $u_n \in W_0^{1,x} L_M(Q_T)$ then $u_n \rightarrow u$ in $L^1(Q_T)$.

Proof 3.11: The proof is easily adapted from that given in [12] by using Theorem 3.7 and Remark 3.9 instead of lemma [40].

IV. Existence result

Let Ω be a bounded Lipschitz domain in \mathbb{R}^N ($N \geq 2$), $T > 0$ and set $Q_T = \Omega \times [0, T]$. We denote $Q_\tau = \Omega \times [0, \tau]$. Let M and P two Musielak-Orlicz functions such that $P \ll M$ and their conjugate respectively \overline{M} and \overline{P} satisfy (9) and (10). Consider a second-order partial differential operator

$$A : D(A) \subset W^{1,x} L_M(Q_T) \rightarrow W^{-1,x} L_{\overline{M}}(Q_T)$$

in divergence form

$$A(u) = -\text{div}(a(x, t, u, \nabla u))$$

where

$a : \Omega \times \mathbb{R} \times \mathbb{R}^N \rightarrow \mathbb{R}^N$ is a Carathéodory function satisfying

for almost every $(x, t) \in Q_T$ and all $s \in \mathbb{R}$, $\xi \neq \xi' \in \mathbb{R}^N$

$$|a(x, t, s, \xi)| \leq \beta(c_1(x, t) + \overline{M}_x^{-1} P(x, k_1 |s| + \overline{M}_x^{-1} M(x, k_1 |\xi|)) \quad (26)$$

$$[a(x, t, s, \xi) - a(x, t, s, \xi')][\xi - \xi'] > 0 \quad (27)$$

$$a(x, t, s, \xi) \xi \geq \alpha[M(x, |s|) + M(x, |\xi|)] \quad (28)$$

with $c_1(x, t) \in E_{\overline{M}}(Q_T)$, $c(x, t) \geq 0$ and $\alpha, \beta, k > 0$. The function ϕ is a Carathéodory function satisfying the following conditions

$$|\Phi(x, t, s)| \leq \gamma(x, t) \overline{P}_x^{-1} P(x, |s|), \quad (29)$$

with $\gamma \in L^\infty(Q_T)$

$$f \in L^1(Q_T) \quad (30)$$

$$u_0 \in L^1(\Omega). \quad (31)$$

Lemma 4.1: Under assumptions (25)- (28), let (z_n) be a sequence in $W_0^{1,x} L_M(Q_T)$ such that,

- (i) $z_n \rightharpoonup z$ in $W_0^{1,x} L_M(Q_T)$ for $\sigma(II L_M(Q_T), II E_{\overline{M}})$
- (ii) $(a(x, t, z_n, \nabla z_n))_n$ is bounded in $(L_M(Q_T))^N$
- (iii) $\int_{Q_T} [a(x, t, z_n, \nabla z_n) - a(x, t, z_n, \nabla z \chi_s)][\nabla z_n - \nabla z \chi_s] dx dt \rightarrow 0.$

as n and s tend to ∞ , and where χ is the characteristic function of

$$Q_s = \{(x, t) \in Q_T; |\nabla z| \leq s\}$$

Then,

$$\nabla z_n \rightarrow \nabla z \quad \text{a.e. in } Q_T, \quad (32)$$

$$\lim_{n \rightarrow \infty} \int_{Q_T} [a(x, t, z_n, \nabla z_n) \nabla z_n] dx dt = \int_{Q_T} [a(x, t, z, \nabla z) \nabla z] dx dt \quad (34)$$

$$M(x, |\nabla z_n|) \rightarrow M(x, |\nabla z|) \text{ strongly in } L^1(Q_T) \quad (35)$$

Proof 4.2: We proceed as in the case of Orlicz spaces (see [1]), we get the desired result.

V. DEFINITION OF AN ENTROPY SOLUTION.

The definition of an entropy solution for problem (1) can be stated as follows.

Definition 5.1:

A measurable function $u : \Omega \times (0, T) \rightarrow \mathbb{R}$ is called entropy solution of (1) if u belongs to $L^\infty(0, T; L^1(\Omega))$, $T_K(u)$ belongs to $D(A) \cap W_0^{1,x} L_M(Q_T)$ for every $K > 0$, $\Theta_K(u(\cdot, t))$ belongs to $L^1(\Omega)$ for every $t \in [0, T]$ and for every $K > 0$ and u satisfies :

$$\begin{aligned} & \int_{\Omega} \Theta_K(u - v) dx + \langle \frac{\partial v}{\partial t}, T_K(u - v) \rangle_{Q_\tau} \\ & + \int_{Q_\tau} a(x, t, T_K(u), \nabla T_K(u)) \nabla T_K(u - v) dx dt \\ & + \int_{Q_\tau} \Phi(x, t, u) \nabla T_K(u - v) dx dt \\ & \leq \int_{Q_\tau} f T_K(u - v) dx dt + \int_{\Omega} \Theta_K(u_0 - v(0)) dx, \end{aligned} \quad (36)$$

and

$$u(x, 0) = u_0(x) \text{ for a.e } x \in \Omega, \quad (37)$$

for every $\tau \in [0, T]$, $K > 0$ and for all $v \in W_0^{1,x} L_M(Q_T) \cap L^\infty(Q_T)$ such that $\frac{\partial v}{\partial t}$ belongs to $W^{-1,x} L_{\overline{M}}(Q_T) + L^1(Q_T)$ (recall that $\Theta_K(r) = \int_0^r T_K(r) dr$ is the primitive of the usual truncation T_K).

This section is devoted to establish the following existence theorem

Theorem 5.2: Assume that the hypotheses (25)-(29) are satisfied, then there exists at least one solution of problem (1) in the sens of Definition (36).

Proof 5.3:

Step1 : Approximation problem.

Let f_n and u_{0n} regular functions in $L^1(Q_T)$ (resp $L^1(\Omega)$) such that:

$$f_n \rightarrow f \text{ in } L^1(Q_T) \text{ and } \|f_n\|_{L^1} \leq \|f\|_{L^1} \quad (38)$$

and

$$\|u_{0n}\|_{L^1} \leq \|u_0\|_{L^1} \text{ and } u_{0n} \rightarrow u_0 \text{ in } L^1(\Omega), \quad (39)$$

as n tends to $+\infty$.

Now, we consider the following regularized problem

$$\begin{cases} \frac{\partial u_n}{\partial t} - \text{div} \left(a(x, t, u_n, \nabla u_n) \right) - \text{div}(\Phi(x, t, u_n)) = f_n \text{ in } Q_T \\ u_n(x, 0) = u_{0n}(x) \text{ in } \Omega \\ u_n = 0 \text{ on } \partial\Omega \times (0, T), \end{cases} \quad (40)$$

The problem (40) can be written as follows

$$\begin{cases} \frac{\partial u_n}{\partial t} - \text{div} \left(F_n(x, t, u_n, \nabla u_n) \right) = f_n \text{ in } Q_T \\ u_n(x, 0) = u_{0n}(x) \text{ in } \Omega \\ u_n = 0 \text{ on } \partial\Omega \times (0, T), \end{cases} \quad (41)$$

with $F_n(x, t, u_n, \nabla u_n) = a(x, t, u_n, \nabla u_n) + (\Phi(x, t, u_n))$.

Note that F_n satisfies the assumptions (A_1) , (A_2) and (A_3) as in [27].

Indeed, using (26), (27) and (29) we deduce that F_n satisfies (A_1) , (A_2) , it remains to prove (A_3) . Let $u_n \in W_0^{1,x} L_M(Q_T)$ by (29) and Young inequality we obtain

$$\begin{aligned} |\Phi(x, t, u_n) \nabla u_n| & \leq |\gamma(x, t)| (P(x, |u_n|) + P(x, |\nabla u_n|)) \\ & \leq C_\gamma (P(x, |u_n|) + P(x, |\nabla u_n|)). \end{aligned} \quad (42)$$

$P \ll M$, then we have for all $\varepsilon > 0$ there exists t_0 that

$$P(x, t) \leq M(x, \varepsilon t) \text{ for all } t \geq t_0, \text{ a.e. } x \in \Omega. \quad (43)$$

Let

$$E_1 = \{(x, t) \in Q_T; |u_n(x, t)| \geq t_0\}$$

$$\text{and } E_2 = \{(x, t) \in Q_T; |\nabla u_n(x, t)| \geq t_0\}$$

Case 1 : if $(x, t) \in E_1 \cap E_2$

In virtue of (42) and (43), we have

$$|\Phi(x, t, u_n) \nabla u_n| \leq C_\gamma (M(x, \varepsilon |u_n|) + M(x, \varepsilon |\nabla u_n|)). \quad (44)$$

Without loss of generality, we can assume that $\varepsilon = \frac{\alpha}{2C_\gamma + \alpha}$ which is $\varepsilon \leq 1$, then by convexity of the function $M(x, \cdot)$, one has

$$\begin{aligned} |\Phi(x, t, u_n) \nabla u_n| & \leq C_\gamma \varepsilon (M(x, |u_n|) + M(x, |\nabla u_n|)) \\ & \leq \frac{\alpha}{2} (M(x, |u_n|) + M(x, |\nabla u_n|)), \end{aligned} \quad (45)$$

which implies

$$\Phi(x, t, u_n) \nabla u_n \geq -\frac{\alpha}{2} (M(x, |u_n|) + M(x, |\nabla u_n|)). \quad (46)$$

From (28) and (46), we have

$$F_n(x, t, u_n, \nabla u_n) \cdot \nabla u_n \geq \frac{\alpha}{2} M(x, |\nabla u_n|). \quad (47)$$

Case 2 : if $(x, t) \in E_1^c \cap E_2^c$

We have

$$|\Phi(x, t, u_n) \nabla u_n| \leq C_\gamma (P(x, |u_n|) + P(x, |\nabla u_n|)) \quad (48)$$

Using the Remark 2.7, we obtain

$$P(x, |u_n|) \leq \text{ess sup}_{x \in \Omega} P(x, t_0) < R_1 < \infty \quad (49)$$

and

$$P(x, |\nabla u_n|) \leq \text{ess sup}_{x \in \Omega} P(x, t_0) < R_2 < \infty. \quad (50)$$

From (49) and (50) we get

$$|\Phi(x, t, u_n)\nabla u_n| \leq C_0. \tag{51}$$

By (28) and (51) we deduce

$$F_n(x, t, u_n, \nabla u_n) \cdot \nabla u_n \geq \alpha M(x, |\nabla u_n|) - C_0 \tag{52}$$

Case 3 :if $(x, t) \in E_1^c \cap E_2$.

In this case, by using Remark 2.7 and (43) we get :

$$|\Phi(x, t, u_n)\nabla u_n| \leq C_1 + C_\gamma M(x, r|\nabla u_n|). \tag{53}$$

We can assume again that $r = \frac{\alpha}{2C_\gamma + \alpha}$ which is $r \leq 1$, then by convexity of the function $M(x, \cdot)$, one has

$$\Phi(x, t, u_n)\nabla u_n \geq -\frac{\alpha}{2}M(x, |\nabla u_n|) - C_1.$$

which implies by using (27)

$$\begin{aligned} F_n(x, t, u_n, \nabla u_n) \cdot \nabla u_n &\geq \frac{\alpha}{2}M(x, |\nabla u_n|) + \alpha M(x, |u_n|) - C_1 \\ &\geq \frac{\alpha}{2}M(x, |\nabla u_n|) - C_1. \end{aligned} \tag{54}$$

By the same way if $(x, t) \in E_1 \cap E_2^c$ we get

$$\begin{aligned} F_n(x, t, u_n, \nabla u_n) \cdot \nabla u_n &\geq \frac{\alpha}{2}M(x, |u_n|) + \alpha M(x, |\nabla u_n|) - C_2 \\ &\geq \alpha M(x, |\nabla u_n|) - C_2. \end{aligned} \tag{55}$$

Finally, from (47), (52) and (54) the assumption (A_3) in [27] is true.

Then there exists at least one solution u_n of (40), (the existence of u_n can be obtained from Galerkin solutions corresponding to the equation (40) as in [27], see Theorem 1 of [2] for more details).

Step 2 : A priori estimates.

Lemma 5.4: Suppose that the assumptions (25) - (29) are true and let u_n be a solution of the approximate problem (40). Then for all $K, n > 0$, we have

$$\int_{Q_T} M(x, |\nabla T_K(u_n)|) dx dt \leq CK. \tag{56}$$

Where C is a positive constant independent of n and K .

And

$$\lim_{K \rightarrow \infty} mes \{(x, t) \in Q_T; |u_n| > K\} = 0. \tag{57}$$

Proof 5.5: Let us note that in the following of this work we will set

$$\Theta_K(t) = \int_0^t T_K(s) ds \tag{58}$$

the primitive of the truncated function $T_K(s)$.

Taking $v = T_K(u_n)_{\chi(0,\tau)}$ as test function in the equation (40) we obtain

$$\begin{aligned} &\int_{\Omega} \Theta_K(u_n)(\tau) dx - \int_{\Omega} \Theta_K(u_{0n}) dx \\ &+ \int_{Q_\tau} a(x, t, u_n, \nabla u_n) \nabla T_K(u_n) dx dt \\ &+ \int_{Q_\tau} \Phi(x, t, u_n) \nabla T_K(u_n) dx dt \\ &= \int_{Q_\tau} f_n T_K(u_n) dx dt, \end{aligned} \tag{59}$$

since $\nabla T_K(u_n) = 0$ in set $\{(x, t) \in Q_T; |u_n(x, t)| > K\}$ which implies that

$$\begin{aligned} &\int_{\Omega} \Theta_K(u_n)(\tau) dx + \int_{Q_\tau} a(x, t, T_K(u_n), \nabla T_K(u_n)) \nabla T_K(u_n) dx dt \\ &+ \int_{Q_\tau} \Phi(x, t, T_K(u_n)) \nabla T_K(u_n) dx dt \\ &= \int_{Q_\tau} f_n T_K(u_n) dx dt + \int_{\Omega} \Theta_K(u_{0n}) dx. \end{aligned} \tag{60}$$

First, from (38) and (39) we have

$$\int_{Q_\tau} f_n T_K(u_n) dx dt + \int_{\Omega} \Theta_K(u_{0n}) dx \leq K(\|f\|_{1, Q_T} + \|u_0\|_{1, \Omega}) \equiv C_2 K, \tag{61}$$

where $C_2 = (\|f\|_{L^1(Q_T)} + \|u_0\|_{L^1(Q_T)})$.

Moreover, by the Young's inequality and the fact that $\gamma \in L^\infty(Q_T)$ we have

$$\begin{aligned} &\int_{Q_\tau} \Phi(x, t, T_K(u_n)) \nabla T_K(u_n) dx dt \leq C_\gamma \int_{Q_\tau} P(x, |T_K(u_n)|) dx dt \\ &+ C_\gamma \int_{Q_\tau} P(x, \nabla T_K(u_n)) dx dt \end{aligned} \tag{62}$$

where $C_\gamma = \|\gamma\|_{L^\infty(Q_T)}$.

From the Remark 2.7 and (43) we therefore get

$$\begin{aligned} &\int_{Q_\tau} P(x, T_K(u_n)) dx dt = \int_{\{(x,t) \in Q_\tau; |T_K(u_n)| \leq t_0\}} P(x, T_K(u_n)) dx dt \\ &+ \int_{\{(x,t) \in Q_\tau; |T_K(u_n)| \geq t_0\}} P(x, T_K(u_n)) dx dt \\ &\leq \int_{\{(x,t) \in Q_\tau; |T_K(u_n)| \leq t_0\}} \text{ess sup}_{x \in \Omega} P(x, |t_0|) dx dt \\ &+ \int_{\{(x,t) \in Q_\tau; |T_K(u_n)| \geq t_0\}} M(x, \varepsilon |T_K(u_n)|) dx dt \\ &\leq R_3 + \int_{Q_\tau} M(x, \varepsilon |T_K(u_n)|) dx dt. \end{aligned} \tag{63}$$

Using the same technics as above, one has

$$\int_{Q_\tau} P(x, |\nabla T_K(u_n)|) dx dt \leq R_4 + \int_{Q_\tau} M(x, \varepsilon |\nabla T_K(u_n)|) dx dt. \tag{64}$$

Hence

$$\begin{aligned} & \int_{Q_T} \Phi(x, t, T_K(u_n)) \nabla T_K(u_n) \, dx \, dt \\ & \leq C_\gamma(R_3 + R_4) + C_\gamma \int_{Q_T} M(x, \varepsilon |T_K(u_n)|) \, dx \, dt \quad (65) \\ & + C_\gamma \int_{Q_T} M(x, \varepsilon |\nabla T_K(u_n)|) \, dx \, dt, \end{aligned}$$

where R_3 and R_4 are constants not depending on K and n . By choosing $\varepsilon = \frac{\alpha}{2C_\gamma + \alpha}$ and convexity of the function M we get

$$\begin{aligned} & \int_{Q_T} \Phi(x, t, T_K(u_n)) \nabla T_K(u_n) \, dx \, dt \\ & \leq C_\gamma(R_3 + R_4) + \frac{\alpha}{2} \int_{Q_T} M(x, |T_K(u_n)|) \, dx \, dt \quad (66) \\ & + \frac{\alpha}{2} \int_{Q_T} M(x, |\nabla T_K(u_n)|) \, dx \, dt. \end{aligned}$$

From (28), (61) and (66) we deduce that

$$\int_{Q_T} M(x, |\nabla T_K(u_n)|) \, dx \, dt \leq CK \text{ for } K \geq 1. \quad (67)$$

Where C is a positive constant independent of K and n . We prove (57). Indeed, it result from (28) and (67) that

$$\text{meas}\{(x, t) \in Q_T; |u_n| > K\} \leq \frac{CK}{\inf_{x \in \Omega} M(x, K)}. \quad (68)$$

Let tending K to infinity. We deduce:

$$\lim_{K \rightarrow \infty} \text{meas}\{(x, t) \in Q_T; |u_n| > K\} = 0. \quad (69)$$

Then we conclude that there exists some $v_K \in W_0^{1,x}L_M(Q_T)$ such that

$$T_K(u_n) \rightharpoonup v_K \text{ weakly in } W_0^{1,x}L_M(Q_T) \text{ for } \sigma(\Pi L_M, \Pi E_{\overline{M}}). \quad (70)$$

Let $\varepsilon > 0$, since (57), (70) and the fact $T_K(u_n)$ is a Cauchy sequence in measure, there exists some $K_\varepsilon > 0$ such that $\text{meas}\{(x, t) \in Q_T; |u_n - u_m| > \lambda\}$ for all $n, m > N_0(K_\varepsilon, \lambda)$. This proves that $(u_n)_n$ is a Cauchy sequence in measure in Q_T thus converges almost everywhere to some measurable function u .

We conclude that there exists some $u \in W_0^{1,x}L_M(Q_T)$ such that

$$T_K(u_n) \rightharpoonup T_K(u) \text{ weakly in } W_0^{1,x}L_M(Q_T), \quad (71)$$

for $\sigma(\Pi L_M, \Pi E_{\overline{M}})$

Next, if we multiply the approximation equation (40) by $\theta'_K(t)$, where $\theta_K(\cdot)$ is a $C^2(\mathbb{R})$ nondecreasing function such that $\theta_K(t) = t$ for $|t| \leq \frac{K}{2}$ and $\theta_K(t) = K$ for $|t| \geq K$, we obtain

$$\begin{aligned} & \frac{\partial \theta_k(u_n)}{\partial t} = \text{div}\left(a(x, t, u_n, \nabla u_n) \theta'_k(u_n)\right) \\ & - a(x, t, u_n, \nabla u_n) \theta''_k(u_n) \nabla u_n \\ & + \text{div}\left(\theta'_k(u_n) \Phi(x, t, u_n)\right) \\ & - \Phi(x, t, u_n) \theta''_k(u_n) \nabla u_n + f_n \theta'_k(u_n), \end{aligned} \quad (72)$$

in the sense of distributions.

Due to(26) and the fact that $T_K(u_n)$ is bounded in $W_0^{1,x}L_M(Q_T)$, the term $-\text{div}\left(a(x, t, u_n, \nabla u_n) \theta'_K(u_n)\right) + a(x, t, u_n, \nabla u_n) \theta''_K(u_n) + f_n \theta'_K(u_n)$ is bounded in $W^{-1}L_{\overline{M}}(Q_T)$. Furthermore, we have $\text{supp}(\theta'_K)$ and $\text{supp}(\theta''_K)$ are both in $[-K, K]$, which gives

$$\begin{aligned} & \left| \int_{Q_T} \theta''_K(u_n) \Phi(x, t, u_n) \nabla u_n \, dx \, dt \right| \\ & \leq \|\theta''_K\|_{L^\infty} \int_{Q_T} |\Phi(x, t, T_K(u_n))| |T_K(u_n)| \, dx \, dt, \end{aligned}$$

by (29), $\gamma \in L^\infty(Q_T)$ and the Young's inequality it follows that

$$\begin{aligned} & \left| \int_{Q_T} \theta''_K(u_n) \Phi(x, t, u_n) \nabla u_n \, dx \, dt \right| \leq \|\theta''_K\|_{L^\infty} \|\gamma\|_{L^\infty(Q_T)} \\ & \times \left[\int_{Q_T} P(x, |\nabla T_K(u_n)|) \, dx \, dt + \int_{Q_T} P(x, |T_K(u_n)|) \, dx \, dt \right]. \end{aligned} \quad (73)$$

By applying the same Technics as in the proof of Lemma 5.4, we prove that $\theta''_K(u_n) \Phi(x, t, u_n) \nabla u_n$ is bounded in $L^1(Q_T)$. In the same way, we show that $\text{div}\left(\theta'_k(u_n) \Phi(x, t, u_n)\right)$ is bounded in $W^{-1,x}L_M(Q_T)$.

Hence all above implies that

$$\frac{\partial \theta_k(u_n)}{\partial t} \text{ is bounded in } W^{-1,x}L_{\overline{M}}(Q_T) + L^1(Q_T). \quad (74)$$

Proceeding as in [35] and using Corollary 3.10, we easily show that there exists a measurable function $u \in L^\infty(0, T; L^1(\Omega))$ such that for every $K > 0$

$$T_K(u_n) \rightharpoonup T_k(u) \text{ weakly in } W^{1,x}L_M(Q_T) \text{ for } \sigma(\Pi L_M, \Pi E_{\overline{M}}) \quad (75)$$

and

$$T_K(u_n) \rightarrow T_k(u) \text{ strongly in } L^1(Q_T) \text{ and a.e in } Q_T \quad (76)$$

Now, we prove the following lemma

Lemma 5.6: Let u_n be a solution of the approximate problem (40), then for all $K \geq 0$,

$$\left(a(x, t, T_K(u_n), \nabla T_K(u_n))\right)_n \text{ is bounded in } (L_{\overline{M}}(Q_T))^N. \quad (77)$$

Proof 5.7: Let $\varphi \in (E_M(Q_T))^N$ be arbitrary. In view of the monotonicity of a , one easily has

$$\left(a(x, t, u_n, \nabla u_n) - a(x, t, u_n, \varphi)\right) (\nabla u_n - \varphi) \geq 0. \quad (78)$$

Hence

$$\begin{aligned} & \int_{\{|u_n| \leq K\}} a(x, t, u_n, \nabla u_n) \varphi \, dx \, dt \quad (79) \\ & \leq \int_{\{|u_n| \leq K\}} a(x, t, u_n, \nabla u_n) \nabla u_n \, dx \, dt \\ & + \int_{\{|u_n| \leq K\}} a(x, t, u_n, \varphi) (\varphi - \nabla u_n) \, dx \, dt. \end{aligned}$$

Using (26) and since $T_K(u_n)$ is bounded in $W_0^{1,x}L_M(Q_T)$, one easily deduces that

$$\int_{Q_T} a(x, t, T_K(u_n), \nabla T_K(u_n)) \nabla T_K(u_n) dx dt \leq CK_1. \tag{80}$$

Combining the fact that $T_K(u_n)$ is bounded in $W_0^{1,x}L_M(Q_T)$, (79) and (80), we get

$$\int_{Q_T} a(x, t, T_K(u_n), \nabla T_K(u_n)) \varphi dx dt \leq CK_2. \tag{81}$$

Hence, thanks the Banach-Steinhaus Theorem, the sequence $(a(x, t, T_K(u_n), \nabla T_K(u_n)))_n$ is a bounded in $(L_{\overline{M}}(Q_T))^N$, thus up to a sub-sequence

$$a(x, t, T_K(u_n), \nabla T_K(u_n)) \rightharpoonup l_K \text{ in } (L_{\overline{M}}(Q_T))^N \tag{82}$$

for $\sigma(\Pi L_{\overline{M}}, \Pi E_M)$, for some $l_K \in (L_{\overline{M}}(Q_T))^N$.

Step 3 : Modular convergence of the gradient.

This step is devoted to introduce for $K \geq 0$ fixed, a time regularization $w_{\mu,j}^i$ of the function $T_K(u)$.

We first introduce two smooth sequences, namely, $(v_j) \subset \mathcal{D}(Q_T)$ such that $v_j \rightarrow u$ in $W_0^{1,x}L_M(Q_T)$ for the modular convergence and almost everywhere in Q_T , and $(\psi_i) \subset \mathcal{D}(\Omega)$ which converges strongly to u_0 in $L^2(\Omega)$ and such that $\|\psi_i\|_{L^2(\Omega)} \leq 2\|u_0\|_{L^2(\Omega)}$, for all $i \geq 1$. For a fixed positive real number K , we consider the truncation function at height K , T_K . Then, for every $K, \mu > 0$ and $i, j \in \mathbb{N}$, we introduce the function $w_{\mu,j}^i \in W_0^{1,x}L_M(Q_T)$ (to simplify the notation, we drop out the index K) defined as $w_{\mu,j}^i = T_K(v_j)_\mu + e^{-\mu t} T_K(\psi_i)$, where $T_K(v_j)_\mu$ is the mollification with respect to time of $T_K(v_j)$ given in (17). From Lemma (3.1), we know that

$$\frac{\partial w_{\mu,j}^i}{\partial t} = \mu(T_K(v_j) - w_{\mu,j}^i), w_{\mu,j}^i(\cdot, 0) = T_K(\psi_i), |w_{\mu,j}^i| \leq K \tag{83}$$

a.e in Q_T ,

$$w_{\mu,j}^i \rightarrow w_\mu^i \stackrel{\text{def}}{=} T_K(u)_\mu + e^{-\mu t} T_K(\psi_i) \text{ in } W_0^{1,x}L_M(Q_T), \tag{84}$$

for the modular convergence as $j \rightarrow \infty$.

$$T_K(u)_\mu + e^{-\mu t} T_K(\psi_i) \rightarrow T_K(u) \text{ in } W_0^{1,x}L_M(Q_T), \tag{85}$$

for the modular convergence as $\mu \rightarrow \infty$.

We will establish the following proposition.

Proposition 5.8: Let u_n be a solution of the approximate problem (25)-(29). Then, for any $K \geq 0$:

$$\nabla u_n \rightarrow \nabla u \text{ a.e. in } Q_T, \tag{86}$$

$$a(x, t, T_K(u_n), \nabla T_K(u_n)) \rightharpoonup a(x, t, T_K(u), \nabla T_K(u)) \tag{87}$$

weakly in $(L_{\overline{M}}(Q_T))^N$,

$$M(|\nabla T_K(u_n)|) \rightarrow M(|\nabla T_K(u)|) \text{ strongly in } L^1(Q_T), \tag{88}$$

as n tends to $+\infty$.

Let us consider the function h_m defined on \mathbb{R} by:

$$h_m(s) \begin{cases} 1 & \text{if } |s| \leq m \\ -|s| + m + 1 & \text{if } m \leq |s| \leq m + 1 \\ 0 & \text{if } |s| \geq m + 1, \end{cases}$$

for any $m \geq K$.

Using the admissible test function $\varphi_{n,j,m}^{\mu,i} = (T_K(u_n) - w_{i,j}^\mu) h_m(u_n)$ as test function in (40) leads to

$$\begin{aligned} & \langle \frac{\partial u_n}{\partial t}, \varphi_{n,j,m}^{\mu,i} \rangle + \int_{Q_T} a(x, t, u_n, \nabla u_n) \\ & \times (\nabla T_K(u_n) - \nabla w_{i,j}^\mu) h_m(u_n) dx dt \\ & + \int_{Q_T} a(x, t, u_n, \nabla u_n) (T_K(u_n) - w_{i,j}^\mu) \nabla u_n h'_m(u_n) dx dt \\ & + \int_{\{m \leq |u_n| \leq m+1\}} \Phi(x, t, u_n) \nabla u_n h'_m(u_n) (T_K(u_n) - w_{i,j}^\mu) dx dt \\ & + \int_{Q_T} \Phi(x, t, u_n) h_m(u_n) (\nabla T_K(u_n) - \nabla w_{i,j}^\mu) dx dt \\ & = \int_{Q_T} f_n \varphi_{n,j,m}^{\mu,i} dx dt. \end{aligned} \tag{89}$$

Denoting by $\epsilon(n, j, \mu, i)$ any quantity such that,

$$\lim_{i \rightarrow \infty} \lim_{\mu \rightarrow \infty} \lim_{j \rightarrow \infty} \lim_{n \rightarrow \infty} \epsilon(n, j, \mu, i) = 0.$$

By the definition of the sequence $w_{i,j}^\mu$, we can establish the following lemma.

Lemma 5.9: Let $\varphi_{n,j,m}^{\mu,i} = (T_K(u_n) - w_{i,j}^\mu) h_m(u_n)$, we have for any $K \geq 0$:

$$\langle \frac{\partial u_n}{\partial t}, \varphi_{n,j,m}^{\mu,i} \rangle \geq \epsilon(n, j, \mu, i), \tag{90}$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between $L^1(Q_T) + W^{-1,x}L_{\overline{M}}(Q_T)$ and $L^\infty(Q_T) \cap W_0^{1,x}L_M(Q_T)$.

Proof 5.10: Using the same techniques as in Orlicz space (see [6]), we can easily get the result.

Now, we turn to complete the proof of Proposition 5.8., we prove below the following results for any fixed $K \geq 0$.

$$\int_{Q_T} f_n \varphi_{n,j,m}^{\mu,i} dx dt = \epsilon(n, j, \mu). \tag{91}$$

$$\int_{Q_T} \Phi(x, t, u_n) h_m(u_n) (\nabla T_K(u_n) - \nabla w_{i,j}^\mu) dx dt = \epsilon(n, j, \mu), \tag{92}$$

$$\int_{\{m \leq |u_n| \leq m+1\}} \Phi(x, t, u_n) \nabla u_n h'_m(u_n) (T_K(u_n) - w_{i,j}^\mu) dx dt = \epsilon(n, j, \mu), \tag{93}$$

$$\int_{Q_T} a(x, t, u_n, \nabla u_n) \nabla u_n h'_m(u_n) (T_K(u_n) - w_{i,j}^\mu) dx dt \leq \epsilon(n, j, \mu, m). \tag{94}$$

$$\int_{Q_T} \left[a(x, t, T_K(u_n), \nabla T_K(u_n)) - a(x, t, T_K(u_n), \nabla T_K(u)) \chi_s \right] \text{Hence} \tag{95}$$

$$\times \left[\nabla T_K(u_n) - \nabla T_K(u) \chi_s \right] dx dt \leq \epsilon(n, j, \mu, m, s).$$

Proof of (91) : By the almost everywhere convergence of u_n , we have $(T_K(u_n) - w_{i,j}^\mu)h_m(u_n)$ converges to $(T_K(u) - w_{i,j}^\mu)h_m(u)$ in $L^\infty(Q_T)$ weak-* and then,

$$\int_{Q_T} f_n(T_K(u_n) - w_{i,j}^\mu)h_m(u_n) dx dt$$

$$\rightarrow \int_{Q_T} f(T_K(u) - w_{i,j}^\mu)h_m(u) dx dt.$$

So that,

$$(T_K(u) - w_{i,j}^\mu)h_m(u) \rightarrow (T_K(u) - T_K(u)_\mu - e^{-\mu t}T_K(\psi_i))$$

in $L^\infty(Q_T)$ weak-* as $j \rightarrow \infty$, and also

$$(T_K(u) - T_K(u)_\mu - e^{-\mu t}T_K(\psi_i)) \rightarrow 0$$

in $L^\infty(Q_T)$ weak-* as $\mu \rightarrow +\infty$. Then, we deduce that,

$$\int_{Q_T} f_n(T_K(u_n) - w_{i,j}^\mu)h_m(u_n) dx dt = \epsilon(n, j, \mu). \tag{96}$$

Proof of (92) and (93): For n large enough, we have

$$\Phi(x, t, u_n)h_m(u_n) = \Phi(x, t, T_{m+1}(u_n))h_m(T_{m+1}(u_n)) \tag{97}$$

a.e in Q_T .

In order to prove (92) and (93), we will apply Lemma 2.9, Let remark that $P \ll M \Leftrightarrow \bar{M} \ll \bar{P}$ (see [25]). Thus we need only to show that $\Phi(x, t, T_{m+1}(u_n))$ converge to $\Phi(x, t, T_{m+1}(u))$ with respect to the modular convergence in $(L_{\bar{P}}(Q_T))^N$ to get the desired result.

Indeed, we put $M_n = \bar{P}\left(x, \frac{\Phi(x, t, T_{m+1}(u_n)) - \Phi(x, t, T_{m+1}(u))}{\mu}\right)$. we have that Φ is a Carathéodory function and using the pointwise convergence of u_n we get that $\Phi(x, t, T_{m+1}(u_n)) \rightarrow \Phi(x, t, T_{m+1}(u))$ a.e in Q_T as $n \rightarrow \infty$, then since $\bar{P}(0) = 0$, one has

$$M_n = \bar{P}\left(x, \frac{\Phi(x, t, T_{m+1}(u_n)) - \Phi(x, t, T_{m+1}(u))}{\mu}\right) \rightarrow 0, \tag{98}$$

a.e in Q_T as $n \rightarrow \infty$.

By the convexity of \bar{P} , for μ and n large enough and by (29), we obtain

$$M_n = \bar{P}\left(x, \frac{\Phi(x, t, T_{m+1}(u_n)) - \Phi(x, t, T_{m+1}(u))}{\mu}\right)$$

$$\leq \frac{C_\gamma}{\mu} \bar{P}\left(x, \bar{P}^{-1} P(x, |T_{m+1}(u_n)|)\right)$$

$$+ \frac{C_\gamma}{\mu} \bar{P}\left(x, \bar{P}^{-1} P(x, |T_{m+1}(u)|)\right) \tag{99}$$

$$\leq \frac{2C_\gamma}{\mu} \text{ess sup}_{x \in \Omega} P(x, m+1) = C_m \text{ a.e. in } Q_T.$$

By Remark 2.7 we have $C_m \in L^1(Q_T)$. Then, using (98), (99) and by Lebesgue's dominated convergence theorem, we obtain

$$\int_{Q_T} M_n dx \rightarrow 0 \text{ as } n \text{ goes to infinity.} \tag{100}$$

$$\Phi(x, t, T_{m+1}(u_n)) \rightarrow \Phi(x, t, T_{m+1}(u)) \tag{101}$$

with respect to the modular convergence in $L_{\bar{P}}(Q_T)$ as $n \rightarrow +\infty$. By applying Lemma 2.9, we obtain $\Phi(x, t, T_{m+1}(u_n)) \rightarrow \Phi(x, t, T_{m+1}(u))$ in $(E_{\bar{M}}(Q_T))^N$.

Then by virtue of, $\nabla T_K(u_n) \rightharpoonup \nabla T_K(u)$ weakly in $(L_M(Q_T))^N$, then

$$\int_{Q_T} \Phi(x, t, u_n)h_m(u_n)(\nabla T_K(u_n) - \nabla w_{i,j}^\mu) dx dt$$

$$\rightarrow \int_{Q_T} \Phi(x, t, u)h_m(u)(\nabla T_K(u) - \nabla w_{i,j}^\mu) dx dt \tag{102}$$

as $n \rightarrow +\infty$.

In the other hand, by using the modular convergence of $w_{i,j}^\mu$ as $j \rightarrow +\infty$ and letting μ tends to infinity, we get (92).

Now we turn to prove (93).

First, remark for $n \geq m+1$ we have that

$$\nabla u_n h'_m(u_n) = \nabla T_{m+1}(u_n) \text{ a.e in } Q_T. \tag{103}$$

By the almost everywhere convergence of u_n , we have $(T_K(u_n) - w_{i,j}^\mu)$ converges to $(T_K(u) - w_{i,j}^\mu)$ in $L^\infty(Q_T)$ weak-* and since the sequence $(\Phi(x, t, T_{m+1}(u_n)))_n$ converges strongly in $E_{\bar{M}}(Q_T)$ then,

$$\Phi(x, t, T_{m+1}(u_n))(T_K(u_n) - w_{i,j}^\mu) \rightarrow \Phi(x, t, T_{m+1}(u))(T_K(u) - w_{i,j}^\mu)$$

converges strongly in $E_{\bar{M}}(Q_T)$ as n goes to $+\infty$.

Using again the fact that, $\nabla T_{m+1}(u_n) \rightharpoonup \nabla T_{m+1}(u)$ weakly in $(L_M(Q_T))^N$ as n tends to $+\infty$ we obtain

$$\int_{\{m \leq |u_n| \leq m+1\}} \Phi(x, t, u_n) \nabla u_n h'_m(u_n) (T_K(u_n) - w_{i,j}^\mu) dx dt$$

$$\rightarrow \int_{\{m \leq |u| \leq m+1\}} \Phi(x, t, u) \nabla u (T_K(u) - w_{i,j}^\mu) dx dt, \tag{104}$$

as n tends to $+\infty$.

By using the modular convergence of $w_{i,j}^\mu$ as $j \rightarrow +\infty$ and letting μ tends to infinity, we get (93).

Proof of (94): Concerning the third term of the right hand side of (89) we obtain that

$$\int_{\{m \leq |u_n| \leq m+1\}} a(x, t, u_n, \nabla u_n) \nabla u_n h'_m(u_n) (T_K(u_n) - w_{i,j}^\mu) dx dt \tag{105}$$

$$\leq 2K \int_{\{m \leq |u_n| \leq m+1\}} a(x, t, u_n, \nabla u_n) \nabla u_n dx dt.$$

Then by (77), we deduce that,

$$\int_{\{m \leq |u_n| \leq m+1\}} a(x, t, u_n, \nabla u_n) \nabla u_n h'_m(u_n) (T_K(u_n) - w_{i,j}^\mu) dx dt \tag{106}$$

$\leq \epsilon(n, \mu, m)$, which is the desired results.

Proof of (95): By means of (89)-(94), we obtain

$$\int_{Q_T} a(x, t, u_n, \nabla u_n) (\nabla T_K(u_n) - \nabla w_{i,j}^\mu) h_m(u_n) dx dt \leq \epsilon(n, \mu, m). \tag{107}$$

Using the same techniques as [24], we obtain

$$\begin{aligned} & \lim_{s \rightarrow \infty} \lim_{n \rightarrow \infty} \int_{Q_T} \left[a(x, t, T_K(u_n), \nabla T_K(u_n)) \right. \\ & \quad \left. - a(x, t, T_K(u_n), \nabla T_K(u) \chi_s) \right] \\ & \quad \times \left[\nabla T_K(u_n) - \nabla T_K(u) \chi_s \right] dx dt = 0. \end{aligned} \tag{108}$$

This implies by the Lemma 4.1., the desired statement and hence the proof of Proposition 5.8. is achieved.

Step 4 : Passing to the limit

Let $v \in W^{1,x}L_M(Q_T) \cap L^\infty$ such that $\frac{\partial v}{\partial t}$ belongs to $W^{-1,x}L_M(Q_T) + L^1(Q_T)$, there exists a prolongation $\bar{v} = v$ on Q_T , $\bar{v} \in W^{1,x}L_M(\Omega \times \mathbb{R}) \cap L^1(\Omega \times \mathbb{R}) \cap L^\infty(\Omega \times \mathbb{R})$, and

$$\frac{\partial v}{\partial t} \in W^{-1,x}L_M(\Omega \times \mathbb{R}) + L^1(\Omega \times \mathbb{R}).$$

There exists also a sequence $(\omega_j) \subset \mathcal{D}(\Omega \times \mathbb{R})$ such that

$$\begin{aligned} \omega_j & \rightarrow \bar{v} \text{ in } W_0^{1,x}L_M(\Omega \times \mathbb{R}), \text{ and} \\ \frac{\partial \omega_j}{\partial t} & \rightarrow \frac{\partial \bar{v}}{\partial t} \text{ in } W^{-1,x}L_M(\Omega \times \mathbb{R}) + L^1(\Omega \times \mathbb{R}). \end{aligned} \tag{109}$$

for the modular convergence and $\|\omega_j\|_{\infty, Q_T} \leq (N + 2)\|v\|_{\infty, Q_T}$ (see [2]).

Now, let us take $T_K(u_n - \omega_j) \chi_{(0,\tau)}$ as a test function in (40), thus for every $\tau \in [0, T]$, we get

$$\begin{aligned} & \left\langle \frac{\partial u_n}{\partial t}, T_K(u_n - \omega_j) \right\rangle_{Q_\tau} \\ & + \int_{Q_\tau} a(x, t, T_{\hat{K}}(u_n), \nabla T_{\hat{K}}(u_n)) \nabla T_K(u_n - \omega_j) dx dt \\ & + \int_{Q_\tau} \Phi(x, t, T_{\hat{K}}(u_n)) \nabla T_K(u_n - \omega_j) dx dt \\ & = \int_{Q_\tau} f_n T_K(u_n - \omega_j) dx dt, \end{aligned} \tag{110}$$

where $\hat{K} = K + C\|v\|_{\infty, Q_T}$, which implies

$$\begin{aligned} & \left\langle \frac{\partial u_n}{\partial t}, T_K(u_n - \omega_j) \right\rangle_{Q_\tau} \\ & + \int_{Q_\tau \cap \{|u_n - \omega_j| \leq K\}} a(x, t, T_{\hat{K}}(u_n), \nabla T_{\hat{K}}(u_n)) \nabla u_n dx \\ & - \int_{Q_\tau \cap \{|u_n - \omega_j| \leq K\}} a(x, t, T_{\hat{K}}(u_n), \nabla T_{\hat{K}}(u_n)) \nabla \omega_j dx \\ & + \int_{Q_\tau} \Phi(x, t, T_{\hat{K}}(u_n)) \nabla T_K(u_n - \omega_j) dx dt \\ & = \int_{Q_\tau} f_n T_K(u_n - \omega_j) dx dt. \end{aligned} \tag{111}$$

By Fatou's lemma and the fact that

$$a(x, t, T_{\hat{K}}(u_n), \nabla T_{\hat{K}}(u_n)) \rightarrow a(x, t, T_{\hat{K}}(u), \nabla T_{\hat{K}}(u))$$

weakly in $(L_M(Q_T))^N$ for $\sigma(\Pi L_M, \Pi E_M)$, one easily sees that

$$\begin{aligned} & \int_{Q_\tau \cap \{|u_n - \omega_j| \leq K\}} a(x, t, T_{\hat{K}}(u_n), \nabla T_{\hat{K}}(u_n)) \nabla u_n dx \\ & - \int_{Q_\tau \cap \{|u_n - \omega_j| \leq K\}} a(x, t, T_{\hat{K}}(u_n), \nabla T_{\hat{K}}(u_n)) \nabla \omega_j dx \\ & \geq \int_{Q_\tau \cap \{|u - \omega_j| \leq K\}} a(x, t, T_{\hat{K}}(u), \nabla T_{\hat{K}}(u)) \nabla u dx \\ & - \int_{Q_\tau \cap \{|u - \omega_j| \leq K\}} a(x, t, T_{\hat{K}}(u), \nabla T_{\hat{K}}(u)) \nabla \omega_j dx. \end{aligned} \tag{112}$$

As in (98), we obtain $\Phi(x, t, T_{\hat{K}}(u_n)) \rightarrow \Phi(x, t, T_{\hat{K}}(u))$ in $E_M(Q_T)$ as $n \rightarrow +\infty$ and using the fact that $\nabla T_K(u_n - \omega_j) \rightarrow \nabla T_K(u - \omega_j)$ in $L_M(Q_T)$, as $n \rightarrow +\infty$, we can easy see that

$$\begin{aligned} & \int_{Q_\tau} \Phi(x, t, T_{\hat{K}}(u_n)) \nabla T_K(u_n - \omega_j) dx dt \\ & \rightarrow \int_{Q_\tau} \Phi(x, t, T_{\hat{K}}(u)) \nabla T_K(u - \omega_j) dx dt. \end{aligned} \tag{113}$$

Since $T_K(u_n - \omega_j) \rightarrow T_K(u - \omega_j)$ weakly* in L^∞ as $n \rightarrow +\infty$, we have

$$\int_{Q_\tau} f_n T_K(u_n - \omega_j) dx dt \rightarrow \int_{Q_\tau} f T_K(u - \omega_j) dx dt.$$

Turn now to see the first term of (110),

$$\begin{aligned} \left\langle \frac{\partial u_n}{\partial t}, T_K(u_n - \omega_j) \right\rangle_{Q_\tau} & = \int_{\Omega} \Theta_K(u_n - \omega_j) dx \\ & + \left\langle \frac{\partial \omega_j}{\partial t}, T_K(u_n - \omega_j) \right\rangle_{Q_\tau} \\ & - \int_{\Omega} \Theta_K(u_{n0} - \omega_j(0)) dx. \end{aligned} \tag{114}$$

First, let see that $u_n \rightarrow u$ in $C([0, T]; L^1(\Omega))$ (see [19]). Moreover, since $\Theta_K(u_n - \omega_j)(\tau) \leq K|u_n(\tau)| + K|\omega_j(\tau)|$, we have by Lebesgue Theorem

$$\int_{\Omega} \Theta_K(u_n - \omega_j)(\tau) dx \rightarrow \int_{\Omega} \Theta_K(u - \omega_j)(\tau) dx,$$

as $n \rightarrow +\infty$. Then, we can pass to the limit in (114) as $n \rightarrow +\infty$ we obtain

$$\begin{aligned} \lim_{n \rightarrow +\infty} \left\langle \frac{\partial u_n}{\partial t}, T_K(u_n - \omega_j) \right\rangle_{Q_\tau} & = \int_{\Omega} \Theta_K(u - \omega_j) dx \\ & + \left\langle \frac{\partial \omega_j}{\partial t}, T_K(u - \omega_j) \right\rangle_{Q_\tau} \\ & - \int_{\Omega} \Theta_K(u_0 - \omega_j(0)) dx. \end{aligned} \tag{115}$$

Now, let n goes to infinity in (110), we get

$$\begin{aligned}
 & \int_{\Omega} \Theta_K(u - \omega_j) dx + \left\langle \frac{\partial \omega_j}{\partial t}, T_K(u - \omega_j) \right\rangle_{Q_\tau} \\
 & + \int_{Q_\tau} a(x, t, u, \nabla u) \nabla T_K(u - \omega_j) dx dt \\
 & + \int_{Q_\tau} \Phi(x, t, u) \nabla T_K(u - \omega_j) dx dt \\
 & \leq \int_{Q_\tau} f T_K(u - \omega_j) dx dt \\
 & + \int_{\Omega} \Theta_K(u_0 - \omega_j(0)) dx.
 \end{aligned} \tag{116}$$

By (109), as j tends to $+\infty$ we have

$$\left\langle \frac{\partial \omega_j}{\partial t}, T_K(u - \omega_j) \right\rangle_{Q_\tau} \rightarrow \left\langle \frac{\partial v}{\partial t}, T_K(u - v) \right\rangle_{Q_\tau}.$$

Moreover, for every $\tau \in [0, T]$, we have $\|\omega_j - v(\tau)\|_{L^1(\Omega)} \rightarrow 0$ as $j \rightarrow +\infty$. Therefore, we pass now to the limit as $j \rightarrow +\infty$ in (116), we get

$$\begin{aligned}
 & \int_{\Omega} \Theta_K(u - v) dx + \left\langle \frac{\partial v}{\partial t}, T_K(u - v) \right\rangle_{Q_\tau} \\
 & + \int_{Q_\tau} a(x, t, u, \nabla u) \nabla T_K(u - v) dx dt \\
 & + \int_{Q_\tau} \Phi(x, t, u) \nabla T_K(u - v) dx dt \\
 & \leq \int_{Q_\tau} f T_K(u - v) dx dt + \int_{\Omega} \Theta_K(u_0 - v(0)) dx.
 \end{aligned} \tag{117}$$

The proof of Theorem 5.2 is complete.

REFERENCES

[1] L. Aharouch, E. Azroul and M. Rhoudaf, *Existence of solutions for unilateral problems in L^1 involving lower order terms in divergence form in Orlicz spaces*. J. Appl. Anal. 13 (2007), no.151-181.

[2] M. L. Ahmed Oubeid, A. Benkirane, and M. Sidi El Vally, *Strongly Nonlinear Parabolic Problems in Musielak-Orlicz-Sobolev Spaces*, v. 33 1 (2015): pp 193-225.

[3] Y. Ahmida and A. Youssfi, *Poincaré-type inequalities in Musielak Spaces*. Annales Academiæ Scientiarum Fennicæ Mathematica, 44, (2019) 1041-1054.

[4] Y. Ahmida, I. Chlebicka, P. Gwiazda, A. Youssfi, *Gossez's approximation theorems in Musielak-Orlicz-Sobolev spaces*. J. Funct. Anal. 275 (9), (2018) 2538-2571.

[5] M. Ait khellou; *Sur certains problèmes non linéaires elliptiques dans les espaces de Musielak-Orlicz*. These (2015).

[6] E. Azroul, H. Redwane and M. Rhoudaf; *Existence of a renormalized solution for a class of nonlinear parabolic equations in Orlicz Spaces*. Port. Math. 66, no. 1, 29-63, (2009).

[7] A. Benkirane and A. Elmahi. An existence theorem for a strongly nonlinear elliptic problem in Orlicz spaces. Nonlinear Anal., 36(1, Ser. A Theory Methods):11-24, 1999.

[8] A. Benkirane and M. Sidi El Vally (Ould Mohamedhen Val): Some approximation properties in Musielak-Orlicz-Sobolev spaces, Thai.J. Math., Vol. 10, N2, pp. 371-381 (2012).

[9] A. Benkirane and M. Sidi El Vally (Ould Mohamedhen val): *Variational inequalities in Musielak-Orlicz-Sobolev spaces*, Bull. Belg. Math. Soc. Simon Stevin. Vol. 21, N 5, pp. 787-811 (2014).

[10] A. Benkirane, J. Douieb, and M. Ould Mohamedhen Val. *An approximation theorem in Musielak-Orlicz-Sobolev spaces*. Comment. Math., 51(1):109-120, 2011.

[11] P. Bénilan, L. Boccardo, T. Gallouët, R. Gariepy, M. Pierre and J.-L. Vazquez, *An L^1 -theory of existence and uniqueness of solutions of nonlinear elliptic equations*, Ann. Scuola Norm. Sup. Pisa, 22, (1995), 241-273.

[12] L. Boccardo and F. Murat, *almost everywhere convergence of the gradients of solution to elliptic and parabolic equations*, Nonlinear Analysis, Theory, Methods and Applications, Vol. 19, No. 6. pp. 581-597, 1992.

[13] L. Boccardo, D. Giachetti, J.-I. Diaz and F. Murat, *Existence and regularity of renormalized solutions for some elliptic problems involving derivation of nonlinear terms*, J. Differential Equations, 106, (1993), 215-237.

[14] Y. Chen, S. Levine and M. Rao *Variable exponent, linear growth functionals in image restoration*. SIAM J. Appl. Math., 66, 1383-1406 (2006).

[15] R.-J. DiPerna, P.-L. Lions; *On the Cauchy problem for Boltzmann equations : Global existence and weak stability*, Ann. Math., 130, (1989), 321-366.

[16] T. Donaldson. *Inhomogeneous Orlicz-Sobolev spaces and nonlinear parabolic initial value problems*. J. Differential Equations, 16:201-256, 1974.

[17] A. Elmahi and D. Meskine. *Parabolic equations in Orlicz spaces*. J. London Math. Soc. (2), 72(2):410-428, 2005.

[18] A. Elmahi and D. Meskine. *Strongly nonlinear parabolic equations with natural growth terms in Orlicz spaces*. Nonlinear Anal., Theory Methods Appl., 60(1):1-35, 2005.

[19] A. Elmahi and D. Meskine, *Strongly nonlinear parabolic equations with natural growth terms in Orlicz spaces*, Nonlinear Analysis. Theory, Methods and Applications, 60, (2005), pp. 1-35.

[20] J.P. Gossez, *Nonlinear elliptic boundary value problems for equation with rapidly or slowly increasing coefficients*, Trans. Amer. Math. Soc, 190, (1974) PP:217-237.

[21] J.-P. Gossez, *Nonlinear elliptic boundary value problems for equations with rapidly or slowly increasing coefficients*, Trans. Amer. Math. Soc., 190, (1974), pp. 163-205.

[22] P. Gwiazda, P. Wittbold, A. Wroblewska-Kaminska and A. Zimmermann, *Renormalized solutions to nonlinear parabolic problems in generalized Musielak-Orlicz spaces* ELSEVIER, (2015).

[23] P. Gwiazda, I. Skrzypczak, A. Zatorska-Goldstein, *Existence of renormalized solutions to elliptic equation in Musielak-Orlicz space* J. Differential Equations 264 (2018) 341-377.

[24] S. Hadj Nassar, H. Moussa and M. Rhoudaf, *Renormalized Solution for a nonlinear parabolic problems with noncoercivity in divergence form in Orlicz Spaces*, Applied Mathematics and Computation 249 (2014) 253-264.

[25] A. KUFNER, O. JHON, B. OPIC, *Function spaces* Academia, Praha, 1977.

[26] O. Kováčik, J. Rákosník; *On spaces $L^{p(x)}$ and $W^{k,p(x)}$* , J. Czechoslovak. Math. 41(1991), 592-618.

[27] R. Landes and V. Mustonen, *A strongly nonlinear parabolic initial-boundary value problem*, Ark. Mat. 25 (1987), 2940.

[28] C. Leone and A. Porretta *Entropy solutions for nonlinear elliptic equation in L^1* , Nonlinear Analysis. Theory. Methods and Applications, Vol. 32, No. 3, pp. 325-34, 1998.

[29] P.-L. Lions, *Mathematical Topics in Fluid Mechanics, Vol. 1: Incompressible models*, Oxford Univ. Press, (1996).

[30] H. Moussa, F. Ortegón Gallego and M. Rhoudaf, *Capacity Solution to a Coupled System of Parabolic-Elliptic Equations in Orlicz-Sobolev Spaces*. NoDEA 25:14 (2018) 1-37.

[31] F. Murat, *Soluciones renormalizadas de EDP elípticas no lineales*, Cours à l'Université de Séville, Publication R93023, Laboratoire d'Analyse Numérique, Paris VI, (1993).

[32] J. Musielak; *Modular spaces and Orlicz spaces ;Lecture Notes in Math.* 1034 (1983).

[33] F. Ortegón Gallego, M. Rhoudaf and H. Sabiki, *On a nonlinear parabolic-elliptic system in Musielak-Orlicz spaces*. EJDE 2018, No. 121 (2018) 1-36.

[34] H. Nakano, *Modular Semi-Ordered Linear Spaces*. Maruzen Co., Ltd., Tokyo, 1950.

[35] Porretta, A.; *Existence results for strongly nonlinear parabolic equations via strong convergence of truncations*. Ann. Mat. Pura Appl. (IV) 177, 143172 (1999)

[36] A. Prignet, *Existence and uniqueness of entropy solutions of parabolic problems with L^1 data*, Nonlin. Anal. TMA 28 (1997), pp. 1943-1954.

[37] K.R. Rajagopal and M. Růžička, *Mathematical modeling of electrorheological materials*, Contin. Mech. Thermodyn. 13 (2001) 59-78.

[38] M. Růžička, *Electrorheological fluids: modeling and mathematical theory.*, Lecture Notes in Mathematics, Springer, Berlin, 2000.

- [39] P. Perona and J. Malik; *Scale-space and edge detection using anisotropic diffusion*, IEEE Trans. Pattern Anal. Machine Intell., 12 (1990), pp. 629-639.
- [40] J. Simon, *Compact sets in the space $L^1(0, T; B)$* , Ann. Mat. Pura. Appl. 146 (1987) 65-96.
- [41] V. Zhikov, Averaging of functionals of the calculus of variations and elasticity theory. *Math. USSR Izvestiya*, 29(1), 33-66 (1987).
- [42] M. Tienari, A degree theory for a class of mappings of monotone type in Orlicz-Sobolev spaces, *Ann. Acad. Scientiarum Fennice Helsinki*(1994).

Optimization of the Substation Location within an Offshore Wind Farm using Particle Swarm and Genetic Algorithms

CHAKIB EL MOKHI 

National School of Applied Sciences
Ibn Tofail University,
Kenitra, Morocco
chakib.elmokhi@uit.ac.ma

ADNANE ADDAIM 

National School of Applied Sciences
Ibn Tofail University,
Kenitra, Morocco
adnane.addaim@uit.ac.ma

Abstract—This paper presents the optimization of the substation location of a wind farm using metaheuristic algorithms. Using this method, the total length of the cables can be reduced, which has a direct impact on the electrical power losses. The cost of the electrical system is estimated to be 15-30% of the total investment cost. By optimizing the substation location, these high costs can be avoided during the design phase of the wind farm. The optimization methods are tested on both a regular and an irregular wind farm. To verify the reliability of the optimization methods, a real existing offshore wind farm is used. The results of the proposed optimization show that substation location optimization is an effective way to design the wind farm layout with low cost.

Index Terms—Wind farm cable layout, substation, power collection system, power losses, metaheuristic optimization.

I. INTRODUCTION

Interest in renewable energies has grown worldwide, especially in wind energy. However, installation costs are high for both onshore and offshore wind farms. In particular, the cost of electrical cabling for the collection and transmission system, based on offshore wind experience, is estimated to be 15–30% of the total investment cost [1]. Therefore, the application of optimization techniques in the development phase of such an electrical system can lead to significant economic gains by finding the optimal location of the substation and the optimal routing of the grid. This minimizes not only the construction costs, but also the other costs associated with power losses and system maintenance.

Several studies have addressed the optimal locations of wind turbines considering the wake effects [2] [3] [4] to improve energy efficiency and minimize energy losses. However, very few works are devoted to the optimization of substation location, optimization of electrical cabling, and selection of proper cable cross-sections for the electrical system. This article focuses on the optimization methods to find the best solution for the substation location.

The energy generated by each wind turbine is first collected via the electrical lines and transmitted to the substation. At this

substation, the voltage is converted to the desired transmission voltage levels and can then be fed into the main grid.

This article is organized as follows. Section II presents the metaheuristic optimization algorithms such as Genetic Algorithm (GA) and Particle Swarm Optimization (PSO), which are used to optimize the substation location. Section III presents the optimization results on a regular and irregular power collection system with some assumptions. In Section IV, the optimization process is performed on a real existing offshore wind farm. The conclusions and future work are presented in Section V.

II. OPTIMIZATION ALGORITHMS

A. Genetic Algorithm

Genetic algorithms, also called evolutionary algorithms, are inspired by Charles Darwin's concept of natural selection. GA belong to a family of bio-inspired populations based on heuristic optimization algorithms that borrow ideas from natural evolution. The concept of genetic algorithm was proposed in 1975 by John Holland of the University of Michigan to describe adaptive systems [5]. It involves the optimization of individuals competing against each other for resources. Some individuals are better suited for this purpose and therefore have a greater chance of survival and can pass on their traits to their offspring, while others do not. These traits are defined in the hereditary material. A single parameter describing a trait is called a genome. A whole set of such genomes is a chromosome. It is important to define a coding of the information so that the genomes can always be assigned to the same property. There are three criteria for further development:

- **Selection:** According to certain defined criteria, the most suitable individuals are selected for reproduction.
- **Crossing:** The new genome consists of a part of the parent **A** and a part of the parent **B**. The genetic information is thus recombined.
- **Mutation:** Errors that occur lead to a change in the genetic information.

When optimizing the substation site with GA, the proposed solutions within the population are formulated in such a way that the coding can be viewed as a genome that defines each solution. The evaluation function is used to determine the fitness of each solution. In this case, the fitness is the position of the substation to all turbines, where the sum of all distances should take the smallest value. If this is the case, the optimal location of the substation has been found.

The flowchart in Fig. 1 shows the main steps of a GA. First, single pairs of individuals are selected from the population to generate new proposed solutions. Then, the pair undergoes what is called crossover [6]. In this crossover, the two parents are combined to generate two new children that receive 50% of their genome from each parent. To ensure that the GA does not stick to a local solution, a mutation operator is used to randomly change the child solutions. These steps are repeated until the solutions converge or the remaining population no longer has sufficient differences.

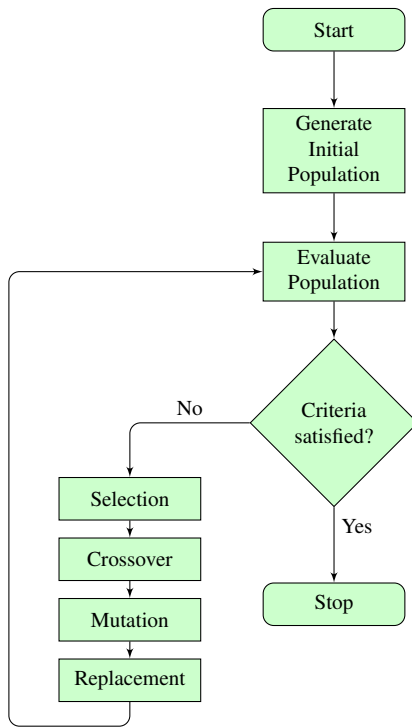


Fig. 1. Flowchart of the Genetic Algorithm

position in the search space and by its velocity. At each instant, all particles will adjust their positions and velocities, that is, their trajectories with respect to their best positions, to the particle with the best position in the swarm and to their current position. In reality, each particle is influenced not only by its own experience, but also by that of the other particles. In PSO, the position of each particle in a given iteration is related to its previous position as expressed in Equation 1:

$$x_i = x_{i-1} + v_i \tag{1}$$

where the velocity v_i is given by:

$$v_i = C_1 v_{i+1} + C_2(p - x_{i-1}) + C_3(g - x_i - 1) + C_4 r \tag{2}$$

where C_1 , C_2 , C_3 , and C_4 are coefficients representing the weights of the different contributors determined by tuning PSO to the problem at hand; p is the best position of the particle in question, g is the best position of the swarm, and $r \in [0, 1]$ is a random number [7] [8].

The main approach of PSO is shown in Fig. 2

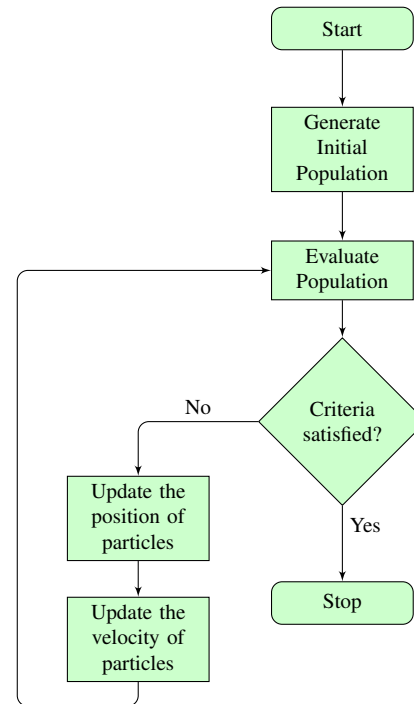


Fig. 2. Flowchart of the Particle Swarm Optimization Algorithm

B. Particle Swarm

Particle Swarm Optimization (PSO) is a metaheuristic optimization that originated in the USA in 1995. It was invented by Russel Eberhart (electrical engineer) and James Kennedy (social psychologist) [4] [7]. Like the GA, this algorithm is modeled after a biological system, although unlike the GA, it is not based on an evolutionary process, but on the behavior of birds and fish living in groups, i.e., their behavior. The algorithm starts with a random initialization of the particle swarm in the search space. Each particle is modeled by its

III. OPTIMIZATION OF THE SUBSTATION LOCATION

A. Regular Wind Farm Layout

To assess the two algorithms GA and PSO, we first consider some assumptions. The case is a regular wind farm array consisting of 64 turbines. The distance between all turbines is 1km each. The substation is located outside the turbines and has 8 cable entries. The cables are laid in a string structure. The calculated cable length between the arrays for the power

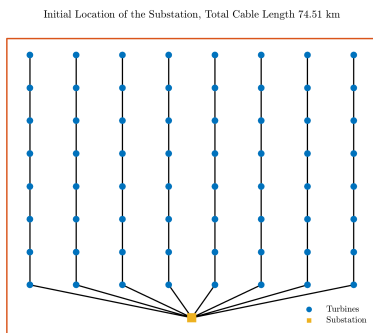


Fig. 3. Initial regular wind farm layout

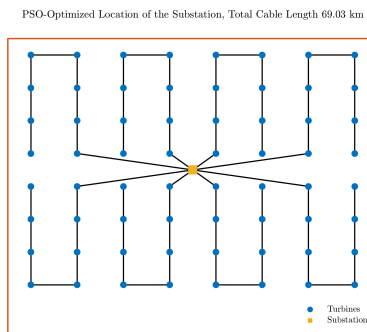


Fig. 5. Optimal substation location with PSO algorithm

collection system is approximately $74.5km$. Fig. 3 shows the above assumptions.

After minimizing the objective function with the GA algorithm by minimizing the Euclidean distance from the substation to each of the turbines, we found that the total length of the inter-array cable is $69km$. Fig. 4 shows the optimal location of the substation and the calculated total cable length.

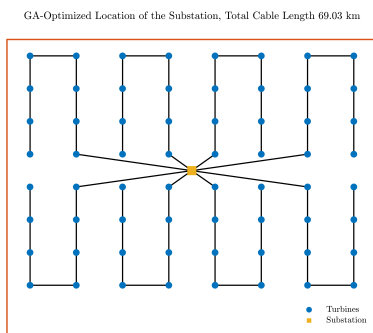


Fig. 4. Optimal substation location with GA algorithm

Using the PSO algorithm, the result for the total cable length of the collection system is also $69km$, similar to the GA system. However, the computation time of the optimization with PSO is much shorter than with the GA algorithm. Fig. 5 shows the result of the PSO algorithm.

Both algorithms, GA and PSO, give the same results for a regular wind farm layout. The saved cable length is $5.48km$, which is 7.36% of the total cable savings.

B. Irregular Wind Farm Layout

In this case, the turbines are generated randomly. The minimum distance between the turbines is $1km$. The substation is located outside the turbine field and has 5 cable entries with 6 turbines each. The total cable length for the collection system is $49.24km$. The initial state is shown in Fig. 6.

Applying the GA and PSO algorithms to optimize the substation location when the string structure of the wind farm layout is irregular also gives very good results. The saved cable length is about $4.10km$, which is equivalent to 8.32% cable length savings. Figures 7 and 8 show the results for GA and PSO algorithms, respectively.

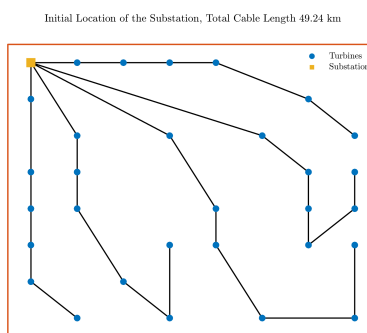


Fig. 6. Initial irregular wind farm layout

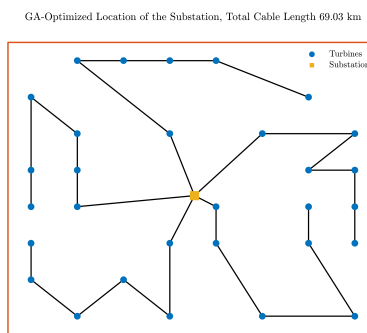


Fig. 7. Optimal substation location with GA algorithm

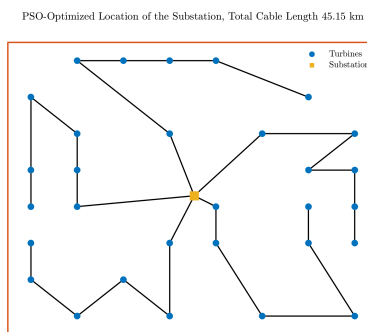


Fig. 8. Optimal substation location with PSO algorithm

IV. REAL CASE STUDY - HORNS REV 1

“Horns Rev 1” was built on an area of $20km^2$ off the coast of Denmark. The offshore wind farm consists of 80 turbines. Table 1 shows more information about the offshore wind farm and Fig. 9 shows the power curve characteristics for the wind turbine.

TABLE I
 WIND FARM SPECIFICATION [9]

Name	Information
Number of turbines	80
Wind farm capacity	160MW
Wind turbine capacity	2MW
Annual energy production	600GWh
Turbine model	V80 – 2MW
Manufacturer	Vestas
Coordinates	55.529722, 7.906111

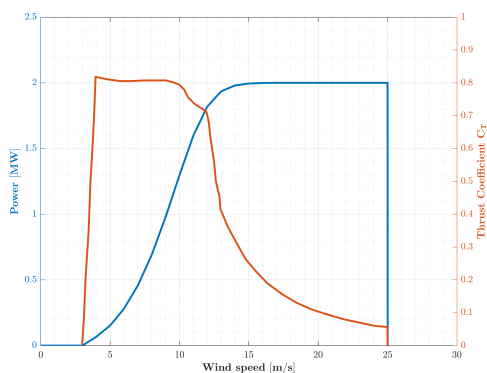


Fig. 9. Power curve of the wind turbine Vestas V80 – 2MW

The initial substation location and the initial layout is shown in Fig. 10.

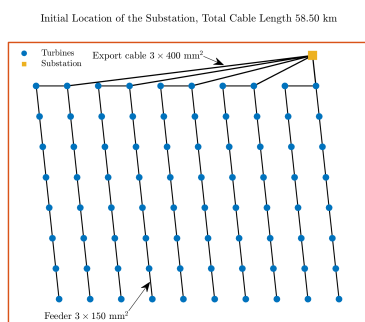


Fig. 10. Wind farm layout of Horns Rev 1

For this case study, it is assumed that the total cable length for the power collection system is $58.50km$. In reality, the length for the cable topology is $63km$ [10]. The distance between the turbines is $560m$, which is 7 times the diameter of the turbines. The substation is located outside the turbine field and has 5 cable entries.

After optimization with the two proposed algorithms GA and PSO, interesting results are provided. It is assumed that the substation has 6 cable entries to avoid the crossing between the submarine cables. The layout has 6 cable branches, 4 with 16 turbines each and 2 with 8 turbines each. Figures 11 and 12 show the proposed cable routing after the substation site optimization. The saved cable length for the power collection system is very interesting and is $12.63km$, which is a saving of 21.60% of the total cable length.

GA-Optimized Location of the Substation, Total Cable Length 45.86 km

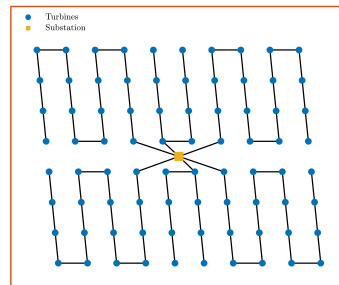


Fig. 11. Optimal substation location with GA

PSO-Optimized Location of the Substation, Total Cable Length 45.86 km

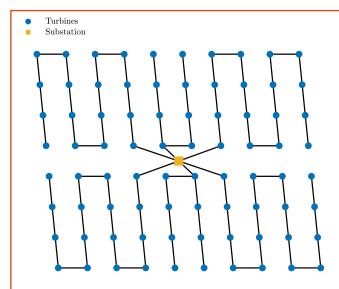


Fig. 12. Optimal substation location with PSO

V. CONCLUSIONS AND FUTURE WORK

This paper focuses on substation siting optimization using metaheuristic algorithms. These algorithms have been tested on both regular and irregular wind farms and show good results in terms of saving cable length for the power collection system. The length of the cable has a direct impact on the power loss and the total investment cost. That is, the longer the cable, the higher the power losses and the total investment cost. The reliability of the GA and PSO optimization methods was verified using a real existing offshore wind farm. The results were satisfactory. However, to further verify the reliability, this work needs to be extended to optimization and selection of the appropriate cable cross-section.

REFERENCES

[1] Fischetti, M.; Pisinger, D. Optimizing wind farm cable routing considering power losses. *Eur. J. Oper. Res.* **2018**, *270*, 917–930, doi:10.1016/j.ejor.2017.07.061.

- [2] Hassoine, M.A.; Lahlou, F.; Addaim, A.; Ait Madi, A. Wind Farm Layout Optimization using Real Coded Multi-population Genetic Algorithm. In Proceedings of the 2019 International Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS 2019), Fez, Morocco, 2019, pp. 1-5.
- [3] Hassoine, M.A.; Lahlou, F.; Addaim, A.; Ait Madi, A. A Layout Investigation of Large Wind Farm in Akhfennir using Real Coded Genetic Algorithm. In Proceedings of the Third International Conference on Computing and Wireless Communication Systems (ICCWCS 2019), Kenitra, Morocco, 24–25 April 2019.
- [4] Gao, X.; Yang, H.; Lu, L.; Koo, P. Wind turbine layout optimization using multi-population genetic algorithm and a case study in Hong Kong offshore. *J. Wind Eng. Ind. Aerodyn.* **2015**, *139*, 89–99, doi:10.1016/j.jweia.2015.01.018.
- [5] Hachimi, H. *Hybridations D'Algorithmes Métaheuristiques en Optimisation Globale et leurs Applications*; HAL Archives: Lyon, France, 2013.
- [6] Gass, S.I.; Fu, M.C. *Encyclopedia of Operations Research and Management Science*, 3rd ed.; Springer: Berlin, Germany, 2013; ISBN 978-1-4419-1154-4.
- [7] Kiranyaz, S.; Ince, T.; Gabbouj, M. *Multidimensional Particle Swarm Optimization for Machine Learning and Pattern Recognition*, 1st ed.; Springer: Berlin, Germany, 2014; ISBN 978-3-642-37846-1.
- [8] Ellaia, R.; Hachimi, H.; ELHami, A. A New Hybrid Genetic Algorithm and Particle Swarm Optimization. *Key Eng. Mater.* **2012**, *498*, 115–125, doi:10.4028/www.scientific.net/KEM.498.115.
- [9] Wikipedia contributors. (2019, September 22). Horns Rev Offshore Wind Farm. In Wikipedia, The Free Encyclopedia. Retrieved 18:00, November 12, 2019, from [https://en.wikipedia.org/w/index.php?title=Horns Rev Offshore Wind Farm&oldid=917114470](https://en.wikipedia.org/w/index.php?title=Horns_Rev_Offshore_Wind_Farm&oldid=917114470).
- [10] Maselis, A. (2016). Layout Optimization of Offshore Wind Farms affected by Wake effects, Cable topology and Support Structure variation.

Comparison between The Ring Flange and the new Quadra-Sector Flange- by modeling and analyzing of the mechanical stress qualification

Nour-Eddine M'HAOUACHE

Department of Engineering Sciences,
Laboratory of Complex Analysis and
Geometry– National School of Applied
Sciences (ENSA),BP 242, Av. de
l'université, 14 000Kenitra-Morocco

n.mhaouache@gmail.com

Hanaa HACHIMI

Systems Engineering Laboratory ,
Sultan Moulay Slimane University

hanaa.hachimi@usms.ma

hanaa.hachimi@usms.ac.ma

Abstract—IJOA Journal

This article deals with the economic aspect and the mechanical reliability of two variants of a structure.

One variant is known and marketed (Ring Flange); the other is innovative and marketed without warranty specifications (Quadra-Sector Flange). The raw material benefit is the compromise between the two variants, and the warranty expected by the customers remains the concern. The manufacture of the two structures is carried out in two operating modes and based on the same raw material; they also both work in the same conditions. The rationale behind this study is reassuring the customers, showing them both patterns by means of a thorough structural stress study based on analyzing our input data (external effects) and comparing the results of the two patterns.

Keywords—IJOA, Journal, Optimization

Flange, Stress, Black steel, Reliability

I. INTRODUCTION

The high production rate of galvanized steel pipes for drinking water supply and the construction of pumping stations require greater demand for black steel flanges [3,7] [3,8].The steel flange is essential connecting joints in industrial piping and remains very important for pipes assembly and disassembly.

It is a mechanical structure which is welded directly on the pipe or manufactured to be mounted on a more sophisticated mounting/dismounting equipment (simple dismounting joint, auto-stop dismantling joint, Gibault joint...) in order to ensure both coupling between two pipes and a dynamic sealing [7,6].

The high demand for these structures and the competitive price between suppliers pushed many of them to innovate a new structure (black steel flange), which functions under the same conditions as the other standard flange. This structure is made of four elements (Quadra-sectors), while the other is completely homogeneous (a single ring) [8, 63].

Many customers complain about the state of the finished product, others ask for a warranty of proper functioning. To better meet these demands we shall make a comparison between the two patterns (quadra-sector flange and ring flange) in order to distinguish the most robust flange structure, without stress.

II. MOTIVATION & METHODOLOGY

A. Motivation

In objective to give our client the justification of the reliable structure, and to assume that the quadra sector flange can make the same function, we have doing the present study.

B. Methodology

The flange is in direct contact with water, connected and mounted with another flange by bolts [2,10] (As shown in fig 1). In our example, the contact is modeled by a fluid-structure coupling: the fluid is water with a density of 1000 Kg / m^3 at an ambient temperature of 35° C to 37° C , and the structural element is steel modeled by 210 MPa Young's Modulus, with a density of 7.860 Kg / m^3 . Our operating environment is in most cases: underground pipes, pipe in conditioning workshops, drinking water supply for a terminal post, pipes in pumping station, or any water supply between two ends.

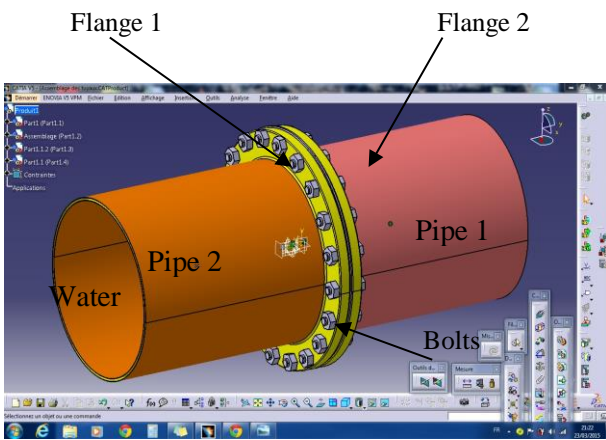


Fig 1: Flange mounting used for piping.

The fluid is water; it is incompressible (constant density): we are facing a Hydro-Elastic problem. Before starting the stress simulations we carried out a vibration analysis of the two patterns in order to observe the impact of the fluid on both structures. After that, we studied two structures using CATIA V5, and this showed us deterioration in the Quadra-sector structure, while the ring structure appeared without degradation. But this study remains limited because the input data are mechanical values and not consistent with CATIA V5 (because of the following: modeling water pressure as a load, the density is not included, and the large scale results). So we were pushed to focus on simulation, meshing and degradation analysis on ANSYS APDL, because its interfaces provide a better modeling and simulation of the fluid-structure coupling. In addition, the input parameters such as density, materials and environment are also included and do not require a large scale. The modeling on ANSYS APDL is justified by the importance of the structure and the final results. We went through the modeling of structures on ANSYS APDL in order to make a geometric and a material definition to the structure and to represent the operating environment. (As shown in figs 2 and 3).

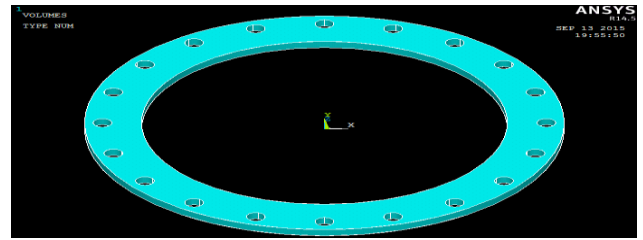


Fig 2: DN 500 PN 10 ring-flange.

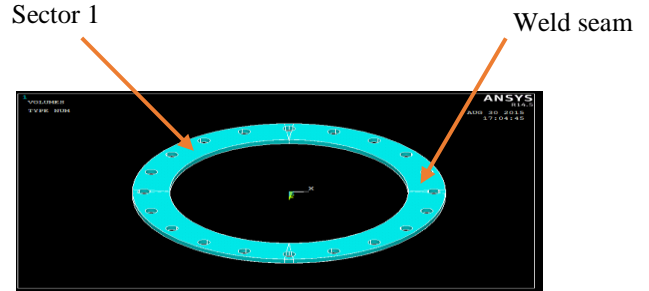


Fig 3: DN 500 PN 10 Quadra-sector flange.

We have a black steel structure, which is produced in manufacturing sites and by a specific preparation process, namely cutting, welding, and precision machining, drilling. Finally, to obtain the standard size and to ensure protection from moisture, a heat treatment is performed in galvanizing basins.

Since the steel ring flange structure has a material homogeneity and volume symmetry, we created the pattern by meshing to obtain a refined structure, represented in form of nodes. We used a tetrahedral meshing with 4 nodes 285 (As shown in fig 4), as this allows us to mesh the structure into more elements.

We have a black steel structure, which is produced in manufacturing sites and by a specific preparation process, namely cutting, welding, and precision machining, drilling. Finally, to obtain the standard size and to ensure protection from moisture, a heat treatment is performed in galvanizing basins.

Since the steel ring flange structure has a material homogeneity and volume symmetry, we created the pattern by meshing to obtain a refined structure, represented in form of nodes. We used a tetrahedral meshing with 4 nodes 285 (As shown in fig 4), as this allows us to mesh the structure into more elements.

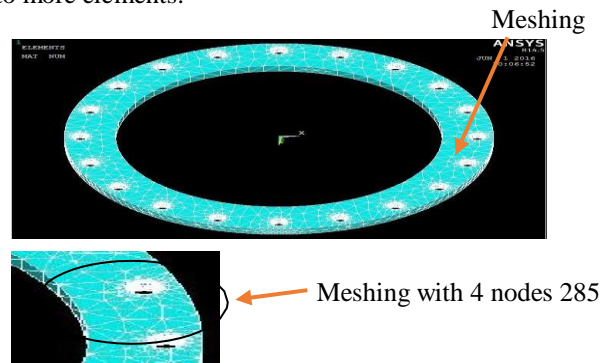


Fig 4: Meshing a DN 500 PN 10 ring-flange using ANSYS APDL

The quadra-sector flange has the same material features and functions under the same conditions as the ring flange, but its manufacture process is quite different. It has 4 elements (As shown in Fig 3), ground and welded with each others to form a flange similar to the ring flange (As shown in Fig 5). We focused the modeling to one single element (sector) (As shown in figs 6, 7, 8), because the structure has a material symmetry, and it has also a uniform pressure distribution (a stable and uniform pressure of 2.5 Br on a single sector). That is why we used the same mesh modeling as in the first structure (ring flange).

After meshing the two types of flanges, we moved to modeling and analysing the results. The findings that interest us are: the displacement of the structure under the pressure (on the inner diameter), and the Von Mises stress. We used them to evaluate the stress of the structure. Results analysis isn't limited only to maximum and minimum values, but also to average extreme values of several nodes as we shall see in what follows.

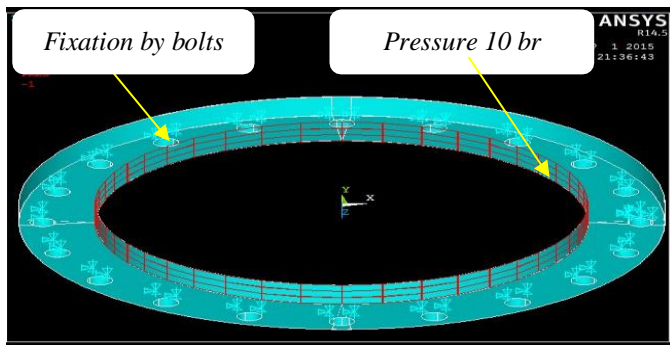
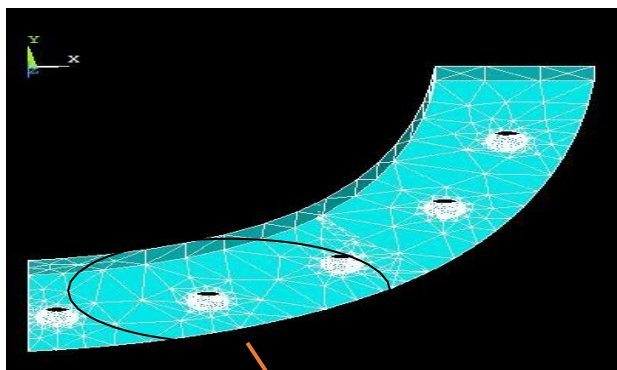


Fig 5: Boundary conditions for a DN 500 PN Quadra-sector flange.



Refining of a weld

Fig 6 & 7: Meshing of one sector of a DN 500 PN 10 Quadra-sector flange & Refining of a weld seal.

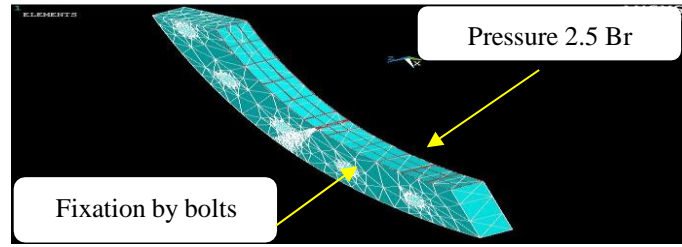


Fig 8: Boundary conditions for one flange sector.

III. RESULTS

Why a quadra-sector flange? For manufacturers the answer is related to the value of Quality/Cost: manufacturing a flange through more stages of production, but a cheaper cost of raw materials (a benefit of 78.29%). If we take a DN 500 PN10 flange (nominal diameter of 500 mm and an operating pressure of 10 br), the manufacture will require a sheet size of 720*720 mm², with a thickness of 40 mm, and a weight of 165.89 Kg to produce a 36 Kg flange. So we use only 21.70% of the metal sheet to produce the structure. It's too expensive for flange manufacturers! The production of flanges into sectors allows a significant benefit not only in raw materials, but also in product sales, in presence of a logical and motivating purchase-sale policy. Our concerns, however, are the claims and complaints of the customers about the heterogeneity of the structure and the appearance of welded surfaces on the flange. (As shown in fig 9).

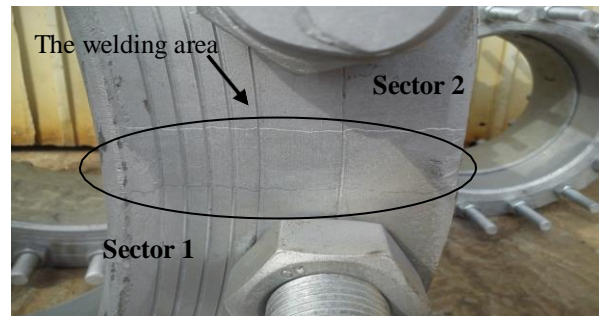


Fig 9: Appearance of the welded area after galvanizing

The following table gives the results of the vibration analysis of the hydro-elastic water-flange problem:

Points	water/ring-flange interaction	water/ Quadra-sector flange interaction
Point 1	43790.2	41051.2
Point 2	43857.3	45921.4
Point 3	43895.2	45933.5
Point 4	43988.2	45985.5
Point 5	44120.5	46133.6
Point 6	44132.7	46670.5

Table 1 : Natural vibration frequencies

Ring-flange pattern

After meshing and using boundary conditions, the flange was put under the conditions of the operating environment by coupling it to another flange (bolted with another flange), and applying a fluid (water) pressure of 10 Br to it. We got the following results in ANSYS APDL interface :

Displacement:

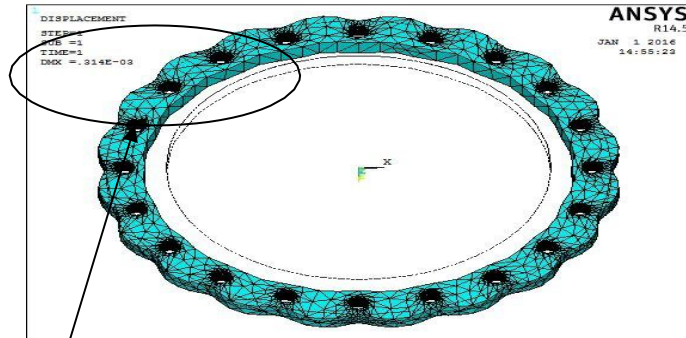


Fig 10: Displacement

The maximum value of displacement

We notice a slight displacement of the structure under the pressure applied, a displacement of $0.314 \cdot 10^{-3}$ mm. This value is the maximum, and justifies the vibration of the structure when the fluid flows.

The maximum nodal displacement

Displacement (nodal solution)



Fig 11: Displacement (nodal solution)

Node	Maximum value of displacement (mm)
396	$0.299 \cdot 10^{-3}$
5112	$-0.3007 \cdot 10^{-3}$
365	$-0.319 \cdot 10^{-3}$
386	$0.314 \cdot 10^{-3}$

Table 2 : maximum nodal displacements.

The advantage of this result is the nodal demonstration of the maximum displacement values. It helps us define the area having a maximum displacement, which causes no degradation of the structure (no critical area, no change in the structure shape, no cracking in our model).

Stress (Von Mises stress)

We illustrate in the following figure the stress of a DN 500 PN 10 ring-flange, by using the representation of Von Mises stress: the result of vibratory and cyclic stresses applied to the structure.

PLNSOL, S, EQV, 0,1 ANSYS control allows to visualize the stress value in all nodes. The figure shows a minimum value of 0.016144 MPa and a maximum value of 2.49406 MPa, but both have no impact on the structure (no crack initiation, and no degradation)

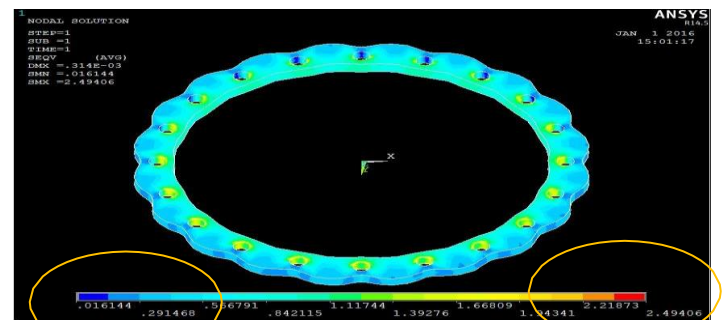


Fig 12: Stress (VON MISES stress)

Von Mises stress in nodal mode

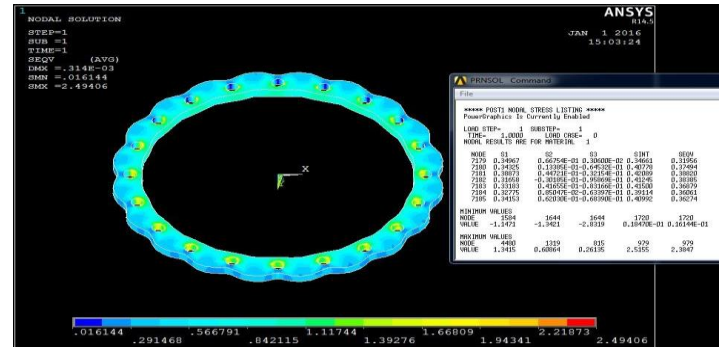


Fig 13: Stress (VON MISES stress)

By using the nodal solution, we got several maximum values of Von Mises stress:

Node	Maximum values of Von Mises stress(MPa)
4480	1.3415
1319	0.6086
815	0.2613
979	2.5155
979	2.3847

Table 3 VON MISES stress (nodal solution)

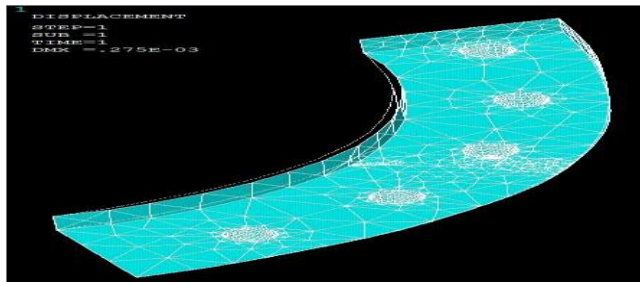
On table 3 the nodal mode allowed us to read some maximum values of Von Mises stress, for example the value 2.3847 MPa at the node 979, which is close to the maximum value 2.49406 MPa.

The most stressed and critical areas of the flange are the holes. That's logical, because the internal surface of the holes undergoes a fixation exerted by the joining bolts, and a pressure at the internal surface of the ring.

A slight inward movement is generated; this movement of the structure (internal ring) causes deformation of the internal surface of the holes (without degradation), and that justifies the stress on holes (see the red color, see fig.6).

Quadra-sector flange pattern

After meshing and using boundary conditions, the quadra-sector flange was put under the conditions of the operating environment by coupling it to another similar flange, and applying a fluid pressure of 0.25 Br to it. We got the following results in the same interface ANSYS:



We have a maximum average movement of 0.275×10^{-3} mm (0.275 micrometer) on the structure under the applied pressure. This value justifies the displacement of the structure when the fluid flows, and does not lead to any degradation under the effect of water pressure: there is no deformation or cracking on the structure [4, 31].

The maximum nodal displacement

Displacement(nodalsolution)

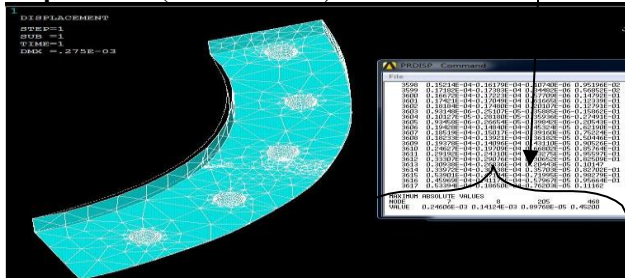


Fig 15: Displacement (nodal solution)

Node	Maximum value of displacement (mm)
07	0.246×10^{-3}
08	0.141×10^{-3}
205	-0.897×10^{-5}

Table 4 : maximum nodal displacements.

Now let's analyze the Von Mises stress in order to decide whether there is stress or not.

Stress (VON MISES stress)

We illustrate in this case the Von Mises stress of a sector of a DN 500 PN 10 flange. With the same PLNSOL, S, EQV 0.1 control, we can see clearly a minimum value of 0.381×10^{-4} MPa and a maximum value of 0.53980.2 MPa, (see figs 9 and 10).

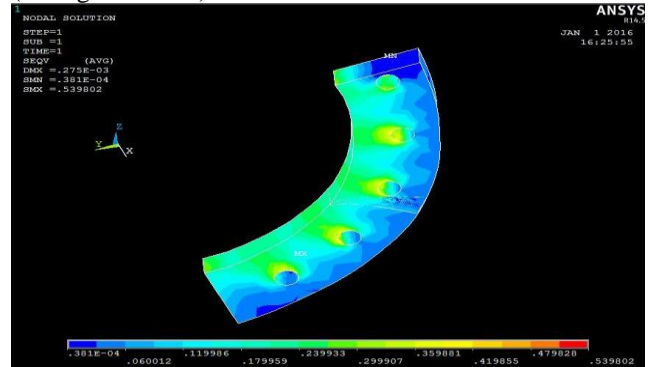


Fig 16: Stress (VON MISES stress)

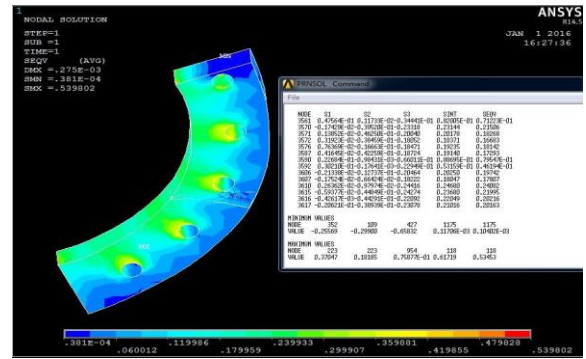


Fig 16: Stress (Maximum VON MISES stress)

As in the first case of the ring flange, the nodal solution gives us several maximum values of Von Mises stress:

Node	Maximum values of Von Mises stress (MPa)
223	0.3704
223	0.1818
954	0.7587×10^{-1}
118	0.6171
118	0.5345

Table 5: VON MISES stress (nodal solution)

We see on the table 5 above some maximum values of Von Mises stress. For example, the value 0.5345 Mpa at node 118, which is close to the average maximum value 0.5398 Mpa (see Fig 10).

The most stressed and critical areas of the sector are still the holes, but the degradation can be noticed on the weld seam,

which undergoes a deformation under the pressure of 0.25 Br. It's an internal deformation of the structure. See the figure below:

Weld seam before simulation

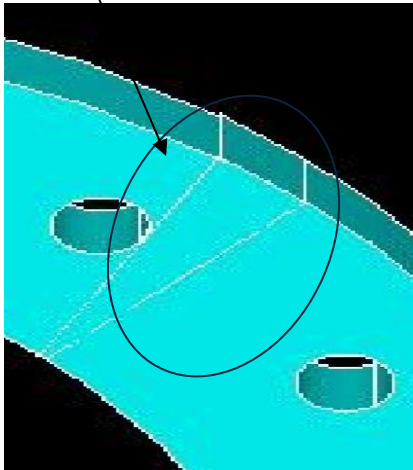


Fig 17: the weld seam before simulation

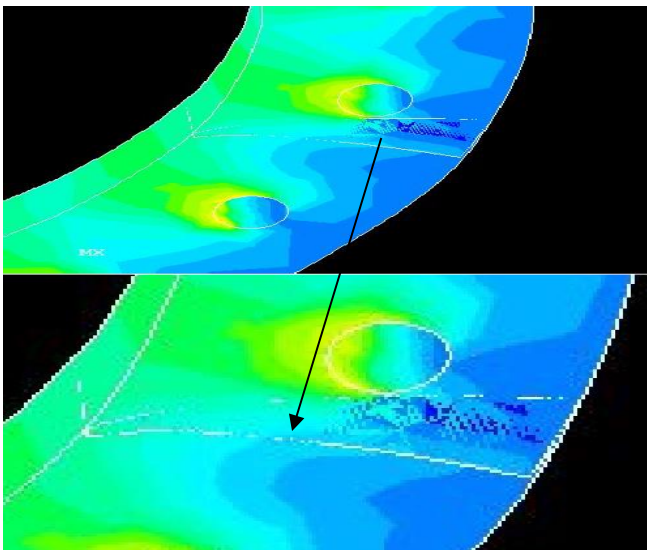


Fig 18: Deformation of the weld seam.

The simulation and analysis study of a sector (one element of the quadra-sector flange) helps us to detect a deformation of the weld seam, which is caused by pressure loads on the sector. Such a deformation causes crack propagation on the weld seam depending on the life time of the structure and the frequency of pressure.

We started our study with modeling and simulation on CATIA V5 interface. We considered the pressure as a mechanical load, and got automatically large scale results because of the transition from a pressure to a load (multiplying by surface unit), which is recommended by the simulation in CATIA V5. The stress analysis, based on the results we got, highlights the degradation of the second pattern (i.e. the quadra-sector flange). The lack of a good pressure modeling and the inadequacy of the parameters with the structure on CATIA V5 pushed us to work on ANSYS APDL interface, instead.

Firstly, we got a normal displacement, a maximum stress at the holes, without degradation or deformation of the ring flange structure. Based on this result, which is consistent with the result seen on CATIA V5, we can conclude that this structure isn't subject to degradation.

Secondly, we made use of the material and dynamic symmetry of the quadra-sector flange, in order to analyze just one element and then to draw conclusions about the quadra-sector flange type as a whole.

This study has shown significant deformation of the weld seam and a maximum stress in holes.

In a proper functioning under the conditions of the operating environment, a distorted weld joint, very close to a highly stressed area (fastening by bolting,) proved that the quadra-sector flange is indeed under stress and subject to degradation. Is this structure reliable or not? What are the qualitative and quantitative parameters that should be improved? The answers to these questions would be the subject of a reliability-based calculation in our next publication.

REFERENCES

- [1] Guide: Eléments de Machine, Gilbert Drouin –Polytechnique Montréal.
- [2] Guide: Directives concernant l'utilisation sûre des joints d'étanchéité. Publication ESA / FSA N°: 009/98.
- [3] Mémoire : Analyse des assemblages à brides boulonnées à face plate - Galai Hicham – L'école de technologie Supérieur- Quebec.
- [4] Thèse: Estimation des Lois de fiabilité en Mécanique par essais accélérés. Ouahiba TEBBI- Université d'ANGERS.
- [5] Thèse: Méthodologie d'analyse de fiabilité basée sur des technique d'optimisation et modèles sans maillage: Application aux systèmes mécaniques - Jhohan Enrique ROJAS-Université Fédérale de Uberlândia.
- [6] Thèse: Endommagement par fatigue et prédiction de la durée de vie des joints soudés Type caisson –Abdulkader Zalt- Université de Lorraine.
- [7] Guide Technique Professionnel pour l'inspection des tuyauteries en exploitation- DT 96 Janvier 2012.
- [8] Catalogue Brides AFNOR (Acier, Inox,..).

Solution of fuzzy differential equation by approximation of fuzzy number

Idris Bakhadach

Laboratory of Applied Mathematics
& Scientific Calculus,
Sultan Moulay Slimane University
idris.bakhadach@gmail.com

Hamid Sadiki

Laboratory of Applied Mathematics
& Scientific Calculus,
Sultan Moulay Slimane University
h.sadiki@usms.ma

Said Melliani

Laboratory of Applied Mathematics
& Scientific Calculus,
Sultan Moulay Slimane University
said.melliani@gmail.com

Lalla Saadia Chadli

Laboratory of Applied Mathematics
& Scientific Calculus,
Sultan Moulay Slimane University
sa.chadli@yahoo.fr

Abstract—In the present paper a definition of fuzzy algebra is presented, the condition of approximation of fuzzy number is proven. Finally the application to solve a fuzzy differential equation is given.

Index Terms—Fuzzy metric space, generalized fuzzy derivative, fuzzy algebra.

I. Introduction

Many scientific papers and many applications have proved that fuzzy set theory let us effectively model and transform imprecise information. It is not surprising that fuzzy numbers play an important role among all fuzzy sets since the predominant carrier of information are numbers. However, the crucial point in fuzzy modeling is to assign membership functions corresponding to fuzzy numbers that represent vague concepts and imprecise terms expressed often in a natural language. The representation does not only depend on the concept but also on the context in which it is used. But even for similar contexts, fuzzy numbers representing the same concept may vary considerably. Therefore, the problem of constructing meaningful membership functions is a difficult one and numerous methods for their construction have been described in the literature. All these methods may be classified into direct or indirect methods that involve one or multiple experts [13].

In practice, fuzzy intervals are often used to represent uncertain or incomplete information. An interesting problem is to approximate general fuzzy intervals by interval,

triangular, and trapezoidal fuzzy numbers, so as to simplify calculations. Recently, many scholars investigated these approximations of fuzzy numbers. According to the different aspects of distance, these researches can be grouped into two classes: the Euclidean distance class [12] and the non-Euclidean distance class [8]. The autor in [10] give a necessary and sufficient conditions of linear operators which are preserved by interval, triangular, symmetric triangular, trapezoidal, or symmetric trapezoidal approximations of fuzzy numbers. In [11] presented a new nearest trapezoidal approximation operator preserving expected interval. But there is no work that has presented a stable part by multiplication, which is our goal in this paper with the approximation to one of whose element in order for example to give a sens of the solution of a differential equation whose contains the product of two specific fuzzy numbers.

This paper is organized as follows: After this introduction we present some concepts concerning the fuzzy metric space in section 2. The fuzzy algebra is defined in Section 3. A method of approximation is is discussed in Section 4, and we presented an application in the las section.

II. preliminaries

In this section, we present some definitions and introduce the necessary notation, which will be used throughout the paper.

We denote E^1 the class of function defined as follows:

$E^1 = \{u : \mathbb{R} \rightarrow [0, 1], \quad u \text{ satisfies (1-4) below}\}$

- 1) u is normal, i.e. there is a $x_0 \in \mathbb{R}$ such that $u(x_0) = 1$;
- 2) u is a fuzzy convex set;
- 3) u is upper semi-continuous;
- 4) u closure of $\{x \in \mathbb{R}^1, \quad u(x) > 0\}$ is compact

For all $\alpha \in (0, 1]$ the α -cut of an element of E^1 is defined by

$$u^\alpha = \{x \in \mathbb{R}, \quad u(x) \geq \alpha\}$$

By the previous properties we can write

$$u^\alpha = [\underline{u}(\alpha), \bar{u}(\alpha)]$$

The multiplication by a scalar is defined as follows

$$\lambda u(x) = \begin{cases} u(\frac{x}{\lambda}), & \lambda \neq 0 \\ \tilde{0}, & \lambda = 0. \end{cases}$$

By the extension principal of Zadeh we have

$$\begin{aligned} (u + v)^\alpha &= u^\alpha + v^\alpha; \\ (\lambda u)^\alpha &= \lambda u^\alpha \end{aligned}$$

For all $u, v \in E^1$ and $\lambda \in \mathbb{R}$

The distance between two element of E^1 is given by (see [4])

$$d(u, v) = \sup_{\alpha \in (0, 1]} \max\{|\bar{u}(\alpha) - \bar{v}(\alpha)|, |\underline{u}(\alpha) - \underline{v}(\alpha)|\}$$

The metric space (E^1, d) is complete, separable and locally compact and the following properties for metric d are valid:

- 1) $d(u + v, u + w) = d(u, v)$;
- 2) $d(\lambda u, \lambda v) = |\lambda|d(u, v)$;
- 3) $d(u + v, w + z) \leq d(u, w) + d(v, z)$;

Remark II.1. The space (E^1, d) is a linear normed space with $\|u\| = d(u, 0)$.

Definition II.2. Let $u, v \in E^1$. We put $u^\alpha = [\underline{u}(\alpha), \bar{u}(\alpha)]$ and $v^\alpha = [\underline{v}(\alpha), \bar{v}(\alpha)]$. We define the product of u and v by

$$(u \odot v)^\alpha = \left[\min\{\underline{u}(\alpha)\underline{v}(\alpha), \underline{u}(\alpha)\bar{v}(\alpha), \bar{u}(\alpha)\underline{v}(\alpha), \bar{u}(\alpha)\bar{v}(\alpha)\}, \max\{\underline{u}(\alpha)\underline{v}(\alpha), \underline{u}(\alpha)\bar{v}(\alpha), \bar{u}(\alpha)\underline{v}(\alpha), \bar{u}(\alpha)\bar{v}(\alpha)\} \right]$$

Definition II.3. [6] The generalized Hukuhara difference of two fuzzy numbers $u, v \in E^1$ is defined as follows

$$u -_g v = w \Leftrightarrow \begin{cases} u = v + w \\ \text{or } v = u + (-1)w \end{cases}$$

In terms of α -levels we have

$(u -_g v)^\alpha = \left[\min\{\underline{u}(\alpha) - \underline{v}(\alpha), \bar{u}(\alpha) - \bar{v}(\alpha)\}, \max\{\underline{u}(\alpha) - \underline{v}(\alpha), \bar{u}(\alpha) - \bar{v}(\alpha)\} \right]$
 and the conditions for the existence of $w = u -_g v \in E^1$ are

- case (i) $\begin{cases} \underline{w}(\alpha) = \underline{u}(\alpha) - \underline{v}(\alpha) \text{ and } \bar{w}(\alpha) = \bar{u}(\alpha) - \bar{v}(\alpha) \\ \text{with } \underline{w}(\alpha) \text{ increasing, } \bar{w}(\alpha) \text{ decreasing, } \underline{w}(\alpha) \leq \bar{w}(\alpha) \end{cases}$
- case (ii) $\begin{cases} \underline{w}(\alpha) = \bar{u}(\alpha) - \bar{v}(\alpha) \text{ and } \bar{w}(\alpha) = \underline{u}(\alpha) - \underline{v}(\alpha) \\ \text{with } \underline{w}(\alpha) \text{ increasing, } \bar{w}(\alpha) \text{ decreasing, } \underline{w}(\alpha) \leq \bar{w}(\alpha) \end{cases}$

for all $\alpha \in [0, 1]$.

Proposition II.4. [6]

$$\|u -_g v\| = d(u, v)$$

Since $\|\cdot\|$ is a norm on E^1 and by the proposition (II.4) we have

Proposition II.5.

$$\|\lambda u -_g \mu u\| = |\lambda - \mu|\|u\|$$

Let $f : [a, b] \subset \mathbb{R} \rightarrow E^1$ a fuzzy-valued function. The α -level of f is given by

$$f(x, \alpha) = \left[\underline{f}(x, \alpha), \bar{f}(x, \alpha) \right], \quad \forall x \in [a, b], \quad \forall \alpha \in [0, 1].$$

Definition II.6. [6] Let $x_0 \in (a, b)$ and h be such that $x_0 + h \in (a, b)$, then the generalized Hukuhara derivative of a fuzzy value function $f : (a, b) \rightarrow E^1$ at x_0 is defined as

$$\lim_{h \rightarrow 0} \left\| \frac{f(x_0 + h) -_g f(x_0)}{h} -_g f'_{gH}(x_0) \right\| = 0 \quad (\text{II.3})$$

If $f_{gH}(x_0) \in E^1$ satisfying (II.3) exists, we say that f is generalized Hukuhara differentiable (gH-differentiable for short) at x_0 .

Definition II.7. [6] Let $f : [a, b] \rightarrow E^1$ and $x_0 \in (a, b)$, with $\underline{f}(x, \alpha)$ and $\bar{f}(x, \alpha)$ both differentiable at x_0 .

We say that

- 1) f is $[(i) - gH]$ -differentiable at x_0 if

$$f'_{i,gH}(x_0) = \left[\underline{f}'(x, \alpha), \bar{f}'(x, \alpha) \right] \quad (\text{II.4})$$

- 2) f is $[(ii) - gH]$ -differentiable at x_0 if

$$f'_{ii,gH}(x_0) = \left[\bar{f}'(x, \alpha), \underline{f}'(x, \alpha) \right] \quad (\text{II.5})$$

Theorem II.8. Let $f : J \subset \mathbb{R} \rightarrow E^1$ and $g : J \rightarrow \mathbb{R}$ and $x \in J$. Suppose that $g(x)$ is differentiable function at x and the fuzzy-valued function $f(x)$ is gH-differentiable at x . So

$$(fg)'_{gH} = (f'g)_{gH} + (fg')_{gH}$$

Proof. Using (II.5), for h enough small we get

$$\begin{aligned} & \left\| \frac{f(x+h)g(x+h) -_g f(x)g(x)}{h} -_g ((f'(x)g(x))_{gH} + (f(x)g'(x))_{gH}) \right\| \\ &= \left\| \frac{f(x+h)g(x+h) -_g f(x)g(x+h) + f(x)g(x+h)}{h} \right. \\ & \quad \left. -_g \frac{f(x)g(x)}{h} -_g ((f'(x)g(x))_{gH} + (f(x)g'(x))_{gH}) \right\| \\ &= \left\| \frac{(f(x+h) -_g f(x))g(x+h) + f(x)(g(x+h) -_g g(x))}{h} \right. \\ & \quad \left. -_g ((f'(x)g(x))_{gH} + (f(x)g'(x))_{gH}) \right\| \\ &\leq \left\| \frac{(f(x+h) -_g f(x))g(x+h)}{h} -_g ((f'(x)g(x))_{gH}) \right\| \\ & \quad + \left\| \frac{f(x)(g(x+h) -_g g(x))}{h} -_g ((f(x)g'(x))_{gH}) \right\| \\ &\leq \left\| \frac{(f(x+h) -_g f(x))g(x+h)}{h} -_g ((f'(x)g(x))_{gH}) \right\| \\ & \quad + \left\| f(x) \frac{(g(x+h) -_g g(x))}{h} -_g ((f(x)g'(x))_{gH}) \right\| \end{aligned}$$

which complet the proof by passing to limit. \square

Definition II.9. [9] Let $f : [a, b] \rightarrow E^1$. We say that $f(x)$ is fuzzy Riemann integrable to $I \in E^1$ if for any $\epsilon > 0$, there exists $\delta > 0$ such that for any division $P = \{[u, v]; \xi\}$ with the norms $\Delta(P) < \delta$, we have

$$d \left(\sum_p^* (v - u)f(\xi), I \right) < \epsilon$$

where \sum_p^* denote the fuzzy summation. We choose to write $I = \int_a^b f(x)dx$.

Theorem II.10. [6] If f is gH -differentiable with no switching point in the interval $[a, b]$ then we have

$$\int_a^b f(t)dt = f(b) -_g f(a)$$

III. Fuzzy algebra

In this section consider \mathcal{P} the set of all $\mu \in E^1$ such that

$$\mu^\alpha = [-a^\alpha, a^\alpha], \quad \forall \alpha \in [0, 1]$$

where $a \in [0, 1]$

Lemma III.1. $(\mathcal{P}, +)$ is a group.

Proof. First $\{0\} = [0, 0] \in \mathcal{P}$.

Now let $\mu, \nu \in \mathcal{P}$ putting $\mu^\alpha = [-a^\alpha, a^\alpha]$ and $\nu^\alpha = [-b^\alpha, b^\alpha]$.

Using case (i) we get

$$\begin{aligned} \mu^\alpha -_g \nu^\alpha &= [-a^\alpha - b^\alpha, a^\alpha + b^\alpha] \\ &= [-c^\alpha, c^\alpha] \end{aligned}$$

where $c = (a^\alpha + b^\alpha)^{\frac{1}{\alpha}}$.

So $\mu -_g \nu \in \mathcal{P}$, thus $(\mathcal{P}, +)$ is a group. \square

Lemma III.2. The product \odot is stable on \mathcal{P} .

Proof. Let $\mu, \nu \in \mathcal{P}$ putting $\mu^\alpha = [-a^\alpha, a^\alpha]$ and $\nu^\alpha = [-b^\alpha, b^\alpha]$.

We have

$$(\mu \odot \nu)^\alpha = [-(ab)^\alpha, (ab)^\alpha]$$

so $\mu \odot \nu \in \mathcal{P}$. \square

Theorem III.3. The quadruplet $(\mathcal{P}, +, \odot, \cdot)$ is an algebra.

Proof. First prove that the triplet $(\mathcal{P}, +, \cdot)$ is a vector space.

By lemmas III.1 and III.2 $(\mathcal{P}, +)$ is a group ($-_g$ is the inverse of $+$) and stable by \cdot .

Now let $\lambda \in \mathbb{R}^+$ and $\gamma \in \mathbb{R}^-$ and $\mu \in \mathcal{P}$, we have

$$(\lambda + \gamma) \cdot \mu = \lambda \cdot \mu -_g (-\gamma) \cdot \mu$$

so (\mathcal{P}, \cdot) is a vector space which implies the result. \square

Definition III.4. The quadruplet $(\mathcal{P}, +, \odot, \cdot)$ is said a fuzzy algebra.

Definition III.5. Let $u, v \in E^1$. We put $u^\alpha = [\underline{u}(\alpha), \bar{u}(\alpha)]$ and $v^\alpha = [\underline{v}(\alpha), \bar{v}(\alpha)]$. we define

$$d_2(u, v) = \left[\int_0^1 (\bar{u}(\alpha) - \bar{v}(\alpha))^2 d\alpha + \int_0^1 (\underline{u}(\alpha) - \underline{v}(\alpha))^2 d\alpha \right]^{\frac{1}{2}}$$

By Brezis [7] $L^2(\mathbb{R})$ is a complet space it is easy to deduce the following proposition.

Proposition III.6. (E^1, d_2) is a complet metric space.

Now we define the following map

$$\langle \cdot, \cdot \rangle : \begin{cases} \mathcal{P} \times \mathcal{P} \rightarrow \mathbb{R} \\ (\mu, \nu) \rightarrow \langle \mu, \nu \rangle = \int_0^1 \underline{\mu}(\alpha)\underline{\nu}(\alpha) + \bar{\mu}(\alpha)\bar{\nu}(\alpha) d\alpha \end{cases}$$

Proposition III.7. The map $\langle \cdot, \cdot \rangle$ define an inner product on \mathcal{P} .

Proof. The linearity of integral show the bilinearity of the map, also the symmetry is clear, we have $\langle \mu, \mu \rangle \geq 0$.

If $\langle \mu, \mu \rangle = 0$ then $\underline{\mu}(\alpha) = \bar{\mu}(\alpha) = 0$, a.e. on $[0, 1]$ we get $\mu = 0$. \square

By III.6 we have

Proposition III.8. $(E^1, \langle \cdot, \cdot \rangle)$ is a Hilbert space.

Note that the norm associated to the inner product $\langle \cdot, \cdot \rangle$ is defined as follows

$$\|\mu\|_2 = \sqrt{\int_0^1 \underline{\mu}^2(\alpha) + \bar{\mu}^2(\alpha) d\alpha}, \quad \forall \mu \in E^1.$$

Proposition III.9. \mathcal{P} is a convex closed subspace of E^1 , by respect the d_2 metric.

Proof. Let $\mu \in \bar{\mathcal{P}}$, then there exist $\mu_n \in \mathcal{P}$ such that $d_2(\mu_n, \mu) \rightarrow 0$, By Brezis [7] theorem 4.9 page 94 there existe a subsequence μ_{n_k} of μ_n converge to μ , if we put $\mu_n^\alpha = [-a_n^\alpha, a_n^\alpha]$ for all $\alpha \in (0, 1]$ then a_{n_k} converge to $a \in \mathbb{R}^+$, which implies that $\mu^\alpha = [-a^\alpha, a^\alpha]$, for all

$\alpha \in (0, 1]$, by consequent $\mu \in \mathcal{P}$.

Now let $t \in [0, 1]$ and $\mu \in \mathcal{P}$, we get

$$t[-a^\alpha, a^\alpha] + (1-t)[-b^\alpha, b^\alpha] = [-c^\alpha, c^\alpha]$$

where $c = (ta^\alpha + (1-t)b^\alpha)^{\frac{1}{\alpha}}$.

Hence \mathcal{P} is convex. \square

IV. Approximation

In this section we consider the set $E^1(a)$ defined as $E^1_- = \left\{ \mu \in E^1, \text{supp}\mu = [-1, 1], \text{ and } \mu^1 = [-a, a], \text{ for certain } a \in [0, 1] \right\}$.

Proposition IV.1. The set E^1_- is stable by the product \odot .

Proof. Let $\mu, \nu \in E^1_-(a)$. First note that $\text{supp}\mu \subset \text{supp}\nu \odot \nu$. In fact:

Let $x \in \text{supp}\mu$, then there is $(x_n)_n \subset \mathbb{R}$ such that $x_n \rightarrow x$ as $n \rightarrow \infty$, and $\mu(x_n) > 0$ which implies that $\mu(x) \geq 0$, we have two possibility if $\mu(x)\nu(x) > 0$ then $x \in \text{supp}(\mu\nu)$ if not, $\mu(x)\nu(x) = 0$, since $\text{supp}(\nu) = \text{supp}(\mu) = [-1, 1]$ then $x = -1$ or $x = 1$ but these two values are in $\text{supp}(\mu\nu)$.

Conversely: by the same it become that $\text{supp}(\mu \odot \nu) \subset \text{supp}(\mu)$.

Thus

$$\text{supp}(\mu \odot \nu) = [-1, 1]$$

By the definition of the product of two fuzzy number we have

$$(\mu \odot \nu)^1 = [-ab, ab]$$

where $\mu^\alpha = [-a^\alpha, a^\alpha]$ and $\nu^\alpha = [-b^\alpha, b^\alpha]$.

So

$$\mu \odot \nu \in \mathcal{P}.$$

\square

and

$$|\bar{\mu}(\alpha) - \bar{\nu}(\alpha)| \leq \epsilon$$

so

$$d_2(\mu, \nu) \leq \epsilon.$$

Conversely:

Suppose that $d_2\mu, \nu \leq \epsilon$, then $\bar{\mu}(\alpha) = \bar{\nu}(\alpha)$ and $\underline{\mu}(\alpha) = \underline{\nu}(\alpha)$ a.e.

By the continuity of $\alpha \rightarrow \bar{\mu}(\alpha) - \bar{\nu}(\alpha)$, we get $\bar{\mu}(\alpha) = \bar{\nu}(\alpha)$ and $\underline{\mu}(\alpha) = \underline{\nu}(\alpha)$, for all $\alpha \in [0, 1]$. Using the idea of the previous part of this demonstration we have

$$|\mu(x) - \nu(x)| \leq \epsilon, \quad \forall x \in \mathbb{R}.$$

Which complet the proof. \square

Proposition IV.4. Any operation based on the extension principale of Zadeh preserve by the previous approximation.

Proof. Let μ, ν, μ' and ν' four fuzzy number such that

$$\mu \approx \mu' \quad \text{and} \quad \nu \approx \nu'.$$

We put $\xi = \mu * \nu$ and $\varsigma = \mu' * \nu'$.

Our goal is to prove that

$$\xi \approx \varsigma.$$

We can write

$$\xi(z) = \sup_{z=x*y} \min \{ \mu(x), \nu(x) \}$$

and

$$\varsigma(z) = \sup_{z=x*y} \min \{ \mu'(x), \nu'(x) \}$$

thus

$$\begin{aligned} |\xi(z) - \varsigma(z)| &= \left| \sup_{z=x*y} \min \{ \mu(x), \nu(x) \} - \sup_{z=x*y} \min \{ \mu'(x), \nu'(x) \} \right| \\ &\leq \left| \sup_{z=x*y} \min \{ \mu'(x) + \epsilon, \nu'(x) + \epsilon \} - \sup_{z=x*y} \min \{ \mu'(x), \nu'(x) \} \right| \\ &\leq \left| \sup_{z=x*y} \min \{ \mu'(x), \nu'(x) \} + \epsilon - \sup_{z=x*y} \min \{ \mu'(x), \nu'(x) \} \right| \\ &\leq \epsilon. \end{aligned}$$

which complet the proof. \square

Theorem IV.5. Let for all $\mu \in E^1$, there exist $\nu \in \mathcal{P}$ such that

$$\mu \approx \nu.$$

Proof. Since $(E^1, \langle \cdot, \cdot \rangle)$ is a Hilbert space and \mathcal{P} is a subset convex closed subset of E^1 then by theorem 5.2 in [7] there existe ν such that

$$d_2(\mu, \nu) \leq \epsilon$$

Definition IV.2. Two fuzzy numbers μ and ν are approximately equal if and only if given a sufficiently small, we find that:

$$|\mu(x) - \nu(x)| \leq \epsilon, \quad \forall x \in \mathbb{R}.$$

we write $\mu \approx \nu$.

Proposition IV.3. Two fuzzy numbers μ and ν of \mathcal{P} are approximately equal if and only if given a sufficiently small, we find that:

$$d_2(u, v) \leq \epsilon$$

Proof. We put $\mu^\alpha = [-a^\alpha, a^\alpha]$ and $\nu^\alpha = [-b^\alpha, b^\alpha]$, for all $\alpha \in [0, 1]$. Where $a, b \in [0, 1]$.

Since μ is increasing on $[-1, a]$ we get $\bar{\mu}(\alpha) = \mu^{-1}(x_\alpha)$, also μ decreasing on $[a, 1]$ then $\underline{\mu}(\alpha) = \mu^{-1}(y_\alpha)$, this two functions are continuous at $\alpha \in [0, 1]$.

If μ and ν are approximately equal, i.e.

$$|\mu(x) - \nu(x)| \leq \epsilon$$

we put $c = \min\{a, b\}$ thus

$$|\mu^{-1}(x) - \nu^{-1}(x)| \leq \epsilon$$

which implies $\mu \approx \nu$. □

Theorem IV.6. Let $\mu \in E_-^1$ and $\nu \in \mathcal{P}$ such that $\nu^\alpha = [-a^\alpha, a^\alpha]$.

A sufficient condition for $\mu \approx \nu$ is:

$$\max_{x \in J} \max \{ |\mu(x - D) - \mu(x)|, |\mu(x + D) - \mu(x)| \} \leq \epsilon$$

where

$$\begin{cases} D = \max_{\alpha \in (0,1]} |a^\alpha - \bar{\mu}(\alpha)|, & \text{if } J = [a, 1] \\ D = \max_{\alpha \in (0,1]} |a^\alpha + \underline{\mu}(\alpha)|, & \text{if } J = [-1, a] \end{cases}$$

Proof. For all $x \in [-1, a]$ it is clear that

$$x - D \leq \mu^{-1}(x) \leq x + D$$

since μ is an increasing function on $[-1, a]$ then

$$\mu(x - D) \leq x \leq \mu(x + D)$$

So that $\forall x \in [-1, a]$, if $\alpha = \mu(x)$ and $\alpha' = \nu(x)$, we get

$$\begin{aligned} |\alpha - \alpha'| &= |\mu(x) - \nu(x)| \\ &\leq \max \{ |\mu(x - D) - \mu(x)|, |\mu(x + D) - \mu(x)| \}. \end{aligned}$$

in the same on $[a, 1]$ we find the same result. □

Theorem IV.7. Let $u, v : I \rightarrow E^1$ two derivative functions, where I is an interval of \mathbb{R} .

If $u(t) \approx v(t)$ for all $t \in I$, then $u'(t) \approx v'(t)$ for all $t \in I$.

Proof. If u and v are $[(i) - gH]$ -differentiable or u and v are $[(ii) - gH]$ -differentiable

$$u'_{i,gH}(x) = [\underline{u}'(x, \alpha), \bar{u}'(x, \alpha)]$$

and

$$v'_{i,gH}(x) = [\underline{v}'(x, \alpha), \bar{v}'(x, \alpha)]$$

or

$$u'_{i,gH}(x) = [\bar{u}'(x, \alpha), \underline{u}'(x, \alpha)]$$

and

$$v'_{i,gH}(x) = [\bar{v}'(x, \alpha), \underline{v}'(x, \alpha)]$$

By the proposition IV.3 we have $u(t) \approx v(t)$.

Now if u is $[(i) - gH]$ -differentiable and v is $[(ii) - gH]$ -differentiable then

$$u'_{i,gH}(x) = [\underline{u}'(x, \alpha), \bar{u}'(x, \alpha)]$$

and

$$v'_{i,gH}(x) = [\bar{v}'(x, \alpha), \underline{v}'(x, \alpha)]$$

this time also by the proposition IV.3 we have $u(t) \approx v(t)$. □

It is easy to note that

Corollary IV.8. If $u(t) \approx v(t)$ then $\int u(t) \approx \int v(t)$.

V. Application

In this section consider the following equation

$$y'(t) = \tilde{a}y(t), \quad t \leq 0 \tag{V.1}$$

with $y \in E_-^1$ and $\tilde{a} \in E_-^1$.

By approximation there exists $z, \tilde{b} \in \mathcal{P}$ such that

$$z \approx y \quad \text{and} \quad \tilde{b} \approx \tilde{a}.$$

Now solving the following equation in \mathcal{P}

$$z' = \tilde{b}z$$

by putting $b^\alpha = [-b^\alpha, b^\alpha]$ and $z(t) = [-(\eta(t))^\alpha, (\eta(t))^\alpha]$. we get

$$(\eta'(t))^\alpha = b^\alpha(\eta(t))^\alpha, \quad \forall \alpha \in [0, 1].$$

Which implies that

$$\eta(t) = ce^{bt}$$

where c is a constant.

Since $t \leq 0$, then $z \in E_-^1$. Using theorem IV.6 each fuzzy element μ verified the condition is a solution.

Remark V.1. By this method the solution of V.1 is not unique.

VI. Conclusions

This study makes it possible to give a meaning to the multiplication of two fuzzy numbers which makes us solve certain differential equations with uncertain initial values.

References

- [1] R.P. Agarwal, S. Arshad, D. O'Regan, V. Lupulescu, Fuzzy fractional integral equations under compactness type condition, *Fract. Calc. Appl. Anal.* 15 (2012) 572–590.
- [2] R.P. Agarwal, V. Lakshmikantham, J.J. Nieto, On the concept of solution for fractional differential equations with uncertainty, *Nonlinear Anal.* 72 (2010) 2859–2862.
- [3] T. Allahviranloo, Z. Gouyandeh, A. Armand, A. Hasanoglu, On fuzzy solutions for heat equation based on generalized hukuhara differentiability, *Fuzzy Sets Syst.* 265 (2015) 1–23.
- [4] G.A. Anastassiou, *Fuzzy Mathematics: Approximation Theory. Studies in Fuzziness and Soft Computing*, Springer, Berlin Heidelberg, 2010.
- [5] A. Armand, Z. Gouyandeh, Fuzzy fractional integro-differential equations under generalized Caputo differentiability. *Annals of Fuzzy Mathematics and Informatics* Volume 10, No. 5, (November 2015), pp. 789798.
- [6] B. Bede, L. Stefanini, Generalized differentiability of fuzzy-valued functions, *Fuzzy Sets Syst.* 230(0) (2013) 119–141.
- [7] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer New York Dordrecht Heidelberg London, ISBN 978-0-387-70913-0.
- [8] S. Chanas, On the interval approximation of a fuzzy number, *Fuzzy Sets and Systems* 122(2001) 353–356.
- [9] M. Congxin, W. Ming, : Embedding problem of fuzzy number space: part III. *Fuzzy Sets Syst.* 46(2), 281–286 (1992)
- [10] Chi-Tsuen Yeh, Existence of interval, triangular, and trapezoidal approximations of fuzzy numbers under a general condition, *Fuzzy Sets and Systems*, Article in press.
- [11] P. Grzegorzewski, Edyta Mrowka, Trapezoidal approximations of fuzzy numbers, *Fuzzy Sets and Systems* 153 (2005) 115–135.
- [12] P. Grzegorzewski, Nearest interval approximation of a fuzzy number, *Fuzzy Sets and Systems* 130(2002) 321–330.

- [13] G.J. Klir, B.Yuan, Fuzzy sets and fuzzy logic, Theory and Applications, Prentice Hall, Englewood Cliffs, 1995.
- [14] Zadeh L. A., Fuzzy sets, Information and Control, 8, (1965), 338–353.

A New Model based on global hybridation of machine learning techniques for “Customer Churn Prediction In Banking Sector”

F. I. Khamlichi

DEVOTEAM

fahd.idrissikhamlichi@devoteam.com

H. Hachimi

USMS of Beni Mellal

hanaa.hachimi@usms.ma

Abstract— We present a new model based on a global hybridation of the most popular machine learning methods applied to the challenging problem of customer churning prediction in Banking Sector. In the first phase of our experiments, all models were applied and evaluated using cross-validation on a popular, public domain dataset. In the second phase, we describe our model and show the performance improvement. In order to determine the most efficient parameter combinations we performed a various simulation for each method and for a wide range of parameters. Our results demonstrate clear superiority of the proposed model against the popular existing ML models.

Keywords— Churn prediction, machine learning techniques, boosting algorithm, banking sector.

I. INTRODUCTION

The regulatory framework within which financial institutions and insurance firms operate require their interaction with customers to be tracked, recorded, stored in Customer Relationship Management (CRM) databases, and then data mine the information in a way that increases customer relations, average revenue per unit (ARPU) and decrease churn rate.

According to scientific research, churn has an equal or greater impact on Customer Lifetime Value (CLTV) when compared to one of the most regarded Key Performance Indicator (KPI's) such as Average Revenue Per User (ARPU).

As one of the biggest destructors of enterprise value, it has become one of the top issues for the banking industry.

The banking industry needs to intensify campaign to deliver a more efficient, customer focused and innovative offerings to reconnect with their customers. The problem of churn analysis is not peculiar to the banking industry.

In this paper we considered the churn detection problem in the Banking sector.

II. MOTIVATION & METHODOLOGY

A. Motivation

Churning is an important problem that has been studied across several areas of interest, such as mobile and telephony, insurance, and healthcare.

B. Methodology

Our paper aims to provide a new approach that allows a combination of existing models to yield even better results. The new approach is based on using a hybridization between a list of chosen models and build a model that chooses the most appropriate model for a new observation at inference time.

III. EVALUATION MEASURES

In order to evaluate classifiers performance in churn prediction for different schemes with their appropriate parameters, we use the measures of precision, recall, accuracy and F-measure, calculated from the contents of the confusion matrix.

True positive and false positive cases are denoted as TP and FP, respectively, while true negative and false negative cases are denoted as TN and FN.

These metrics are particularly important for the problem at hand (Churn Detection). They are the most significant metrics especially from the end-user point of view (Bank marketing teams and managers). A Bank usually assigns costs and benefits to clients that are predicted to stay and allows budgets to keep customers that are expected to churn.

The balance between the cost of a false positive and a false negative is thereby end-user dependent. The ‘best’ model is also by the same reason case dependent.

However, for benchmarking purposes, we have chosen the accuracy as the main metric to choose the best model and also to compare our overall results with the previous works.

IV. RELATED WORKS

Our work is inspired by the strategy of F. I. Khamlichi and R. Aboulaich in « global face recognition strategy based on multiple recognition systems combination » published in the international journal of imaging and robotics (issn 2231–525x), volume 6, number a11, autumn (2011). They extend multiple recognition systems for global face recognition. The idea behind this strategy is to combine the models based on the facial complexion as an input and basic criterion to specify

which model to apply in order to obtain the best results in facial recognition.

To our knowledge, the other paper that also learn about the combined systems is this of F. I. Khamlichi, R. Aboulaich and A. E. El-Mrhari in global decision strategy based on combining multiple regression methods published in the international journal of statistics & statistics & economics ((ijse) (issn: 0975-556x), volume 8, number s12, 2012).

V. OUR ALGORITHM

As mentioned in the introduction, our approach is based on choosing the best combination of known models for the specific problem and apply the best model at the instance level.

The intuition behind this approach is better illustrated by an example. A model with a task of classification of cats and dogs into their sub-races performs less than two models, one trained on cats and another trained on dogs (The difference of performances has been proven for face recognition in one of our previous works [8] « Global face Recognition Strategy based on multiple recognition systems combination »).

Our approach applied to this problem would provide an overall model that first identifies if the input picture is of a cat or a dog. The second stage will be to apply the best model for the identified ‘class’.

In the case of Churn Detection, we have to identify a way to split our dataset so we have a similar split between cats and dogs. The feature would ideally split our observations (customers) into few sub-populations that have different behaviours with regards to the chosen feature and also to our problem.

Such feature would show a variation of the target in a span basis (ie one class is predominant for observations having this feature within the same span). When the dataset is split based on this features’ spans, the classification hopefully becomes a simpler problem.

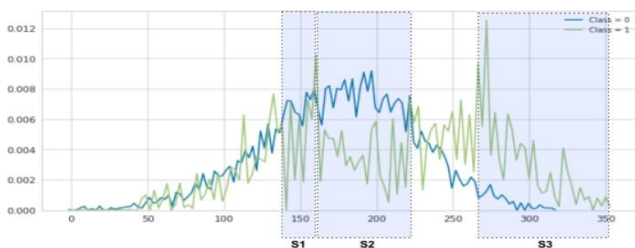


Fig. 1: Illustration: Examples of spans where one class is relatively predominant.

The number of possible features and their spans are very high for our dataset. This makes the identification via evaluation of every possibility computationally prohibiting. We have therefore decided to count on our understanding of the problem.

In our case, the problem of customers churn for Banks, we identified that the best assumption we can make is that the behavior of customers depends greatly on their age.

Visualizing the dataset shows that (3) spans is a good choice for the split. The next step is to split our data into subsets having the **Age** within three (3) defined spans. The resulting subsets are called clusters C1, C2 and C3.

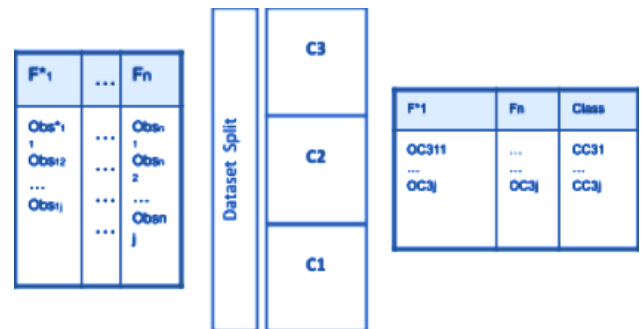


Fig. 2: Dataset split.

These clusters and their corresponding spans are as described in the following table.

Cluster	Span	Size
C1	0 - 30	3356
C2	30 - 45	3334
C3	45 - 100	3310

Table. 1: Dataset split results.

Once the data set is split, it is again split into training and validation sets (80% / 20%). The dataset is now ready for training and evaluation. We generally train n models M1, M2, ..., Mn, on each one of the subsets C1, C2 and C3.

Each model is then evaluated on the corresponding cluster and we compare the resulting performances: Gc1, Gc2, .. Gcn for each cluster c and identify the best model for the cluster c. The final results are the best identified models for each cluster: Mc1, Mc2 and Mc3 (Mci being the best model for cluster ‘ci’ where in general $i \in [1, \text{number of clusters}]$).

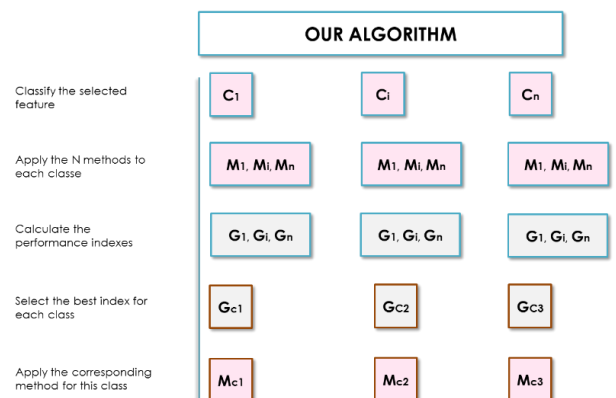


Fig. 3: The algorithm produces as many best models as the data split clusters.

At The inference time, each new observation is first placed into the corresponding cluster. We then use the best model for the given cluster to make the prediction.

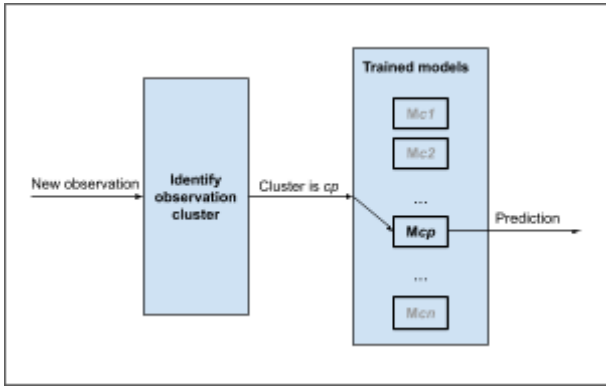


Fig. 4: Two stages inference.

VI. SIMULATION

A. Simulation Setup

Our main goal is to compare different models of classification on churn prediction and build the best one with a better prediction. In this sense, we implemented our simulation in three main steps using Python language following its power in the field of data science. First, we implemented different classification models throughout the dataset, the main models are first Logistic Regression, Support Vector Machine (SVM), Decision Tree Classification, Random Forest and XGBoost.

The next step is to select the influencing feature to divide it into classes and apply the model to each class in order to choose the perfect one, then merge the result into one to build a single model that groups the models previously built (according to the number of classes devised).

B. Results

To evaluate the performance of the tested classifiers, we worked with a churn dataset from the UCI Machine Learning Repository, this version contains 10000 samples and 21 features plus the variable churn, presented in Table.2.

```

Data columns (total 14 columns):
#   Column              Non-Null Count  Dtype
---  -
0   RowNumber           10000 non-null   int64
1   CustomerId          10000 non-null   int64
2   Surname             10000 non-null   object
3   CreditScore         10000 non-null   int64
4   Geography          10000 non-null   object
5   Gender              10000 non-null   object
6   Age                 10000 non-null   int64
7   Tenure              10000 non-null   int64
8   Balance             10000 non-null   float64
9   NumOfProducts      10000 non-null   int64
10  HasCrCard           10000 non-null   int64
11  IsActiveMember     10000 non-null   int64
12  EstimatedSalary    10000 non-null   float64
13  Exited              10000 non-null   int64
dtypes: float64(2), int64(9), object(3)

```

C. Classifiers performance with hybrid method

Next, we divide the data set into three classes according to the feature that hopefully has the most impact on the target (in our case: Age) using the library named `qcut` and apply the

models built previously in each class to choose those with the best results and merge them into one over the entire dataset, respecting the model used for each class.

Based on the results obtained in the Table.3 we choose to work the first class with SVM-POLY classifier with two parameters, C with 0.5, and gamma with 10.

For the second class we worked with the Logistic Regression LR classifier with penalty and C as the main parameters of this model, penalty =1 and C=0.1.(Table.4)

In the case of the third class we used the Decision Tree Classifier using `random_state`, `min_samples_split`, `min_samples_leaf` and `max_features` having successively those values `random_state=123`, `min_samples_split= 10`, `min_samples_leaf= 1`, `max_features=sqrt`.(Table.5).

SVM (POLY)	Accuracy	False Positive	True Negative	False Negative	True Positive
C1	99%	1	138	1	28
C2	97%	2	137	3	24
C3	98%	1	148	2	16

Table.3: Metrics obtained using SVM(POLY) model on the different clusters.

Logistic Regression	Accuracy	False Positive	True Negative	False Negative	True Positive
C1	98%	1	140	2	25
C2	99%	0	140	1	25
C3	97%	3	141	1	22

Table.4: Metrics obtained using Logistic Regression model on the different clusters.

Decision Tree	Accuracy	False Positive	True Negative	False Negative	True Positive
C1	98%	1	141	3	25
C2	98%	2	150	1	14
C3	99%	1	143	0	23

Table.5: Accuracy, False Positive, True Negative, False Negative and True Positive obtained using Decision Tree model on the different clusters.

In this table (Table.6) we summarize the best results chosen for each cluster.

	Logistic Regression	SVM (POLY)	Decision Tree
C1	98%	99%	98%
C2	99%	97%	98%
C3	97%	98%	99%

Table.6: Best Accuracy results for each cluster by model

5.4. Performance comparison – Discussion

Using the mechanism described in “Fig. 4: Two stages inference”, we evaluated the mechanism overall performance and compared it to the results from separate models. For the single models, the evaluation has been held on the whole dataset without separation based on any feature. Table.4 presents the resulting performance with a visible increase, especially the F-measure metric, with the new approach.

Classifiers	Accuracy	F-measure
	Single model	
Random Forest	95%	79%
XGBoost	95%	80%
Support Vector Machine	93%	73%
Decision Tree	90%	64%
Logistic Regression	86%	26%
K-Nearest Neighbors	81%	28%
Combined models approach		
Hybrid model	99%	97%

Table.7: Performance comparison for models’ classifiers with and without hybridization.

VII. CONCLUSION AND SUMMARY

Our simulations were performed using six of the most popular classification methods for the churn prediction problem of Bank customers based on a publicly available dataset. First of all, all methods were tested without the use of boosting under different settings. The two top performing methods in our test were XGBoost classifier and Random Forest classifier, both methods achieved accuracy 95% and F-measure 80% approximately. The Support Vector Machine

classifier (POLY kernel) obtained accuracy of 93% and 73% about F-measure. Decision Tree method achieved the accuracy of 90% and F-measure 64%. Logistic Regression and K-Nearest Neighbors methods fail short with accuracy 86% and 81% also an F-measure of about 26% and 28% respectively.

At a later stage, we considered to work with the entire dataset by dividing it into an impacting feature (classes) and implementing the algorithms on each slice then merging the results into a single algorithm. In our case we choose to work mainly with Support Vector Machine-POLY, Logistic Regression and Decision Tree. Overall, this method was the best way of classification with accuracy of almost 99% and F-measure over 97%.

This work has brought to light a new way of exploiting several algorithms in one according to the problem and the data we have, and still the best way to use the most popular algorithms and exploit them into one. Also, to use a larger and more detailed dataset can maximize the statistical significance of the result of our algorithm.

REFERENCES

- [1] The Chartered Institute of Marketing, Cost of customer acquisition versus customer retention (2010).
- [2] S. A. Qureshi, A. S. Rehman, A. M. Qamar, A. Kamal, A. Rehman, Telecommunication subscribers’ churn prediction model using machine learning, in: Digital Information Management (ICDIM), 2013 Eighth International Conference on, IEEE, 2013, pp. 131–136.
- [3] K. Kim, C.-H. Jun, J. Lee, Improved churn prediction in telecommunication industry by analyzing a large network, Expert Systems with Applications.
- [4] C. Kirui, L. Hong, W. Cheruiyot, H. Kirui, Predicting customer churn in mobile telephony industry using probabilistic classifiers in data mining, International Journal of Computer Science Issues (IJCSI) 10 (2).
- [5] G. Kraljević, S. Gotovac, Modeling data mining applications for prediction of prepaid churn in telecommunication services, AUTOMATIKA: časopis za automatiku, mjerenje, elektroniku, računarstvo i komunikacije 51 (3) (2010) 275–283.
- [6] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research, 321–357.
- [7] Dudyal, A. K., & Ravi, V. (2008). Predicting credit card customer churn in banks using data mining. International Journal of Data Analysis Techniques and Strategies, 1(1), 4-28.
- [8] Farquod, M. A., Ravi, V., & Raju, S. B. (2009). Data mining using rules extracted from SVM: an application to churn prediction in bank credit cards. In Rough Sets, Fuzzy Sets, Data Mining and Granular Computing (pp. 390-397). Springer Berlin Heidelberg.
- [9] F. I. Khamlichi, « Analysis and identification of the feelings or the behavior of a person through web pages from social web sites, using Fuzzy Logic method », JICCLAP2018, Qatar 2018.
- [10] M. Kaicer, S. Mhir and F. I. Khamlichi, « Vital Prognosis of Intoxicated Patients, Modeling by SVM », IProcedia Computer Science, Volume 127, 2018, Pages 154-160.
- [11] F. I. Khamlichi, A. Boufrahi, « Mammography diagnostic using recognition system based on closed loop fuzzy logic controller », International Journal of Imaging and Robotics (ISSN 2231–525X), 2017.
- [12] F. I. Khamlichi, R. Aboulaich and A. E. El Mrhari, « Quantitative-trading strategy based on fuzzy logic model », International Journal of Mathematics & Computation (IJMC) (ISSN: 0974-5718), Volume 19, Number 2, 2013.
- [13] R. Aboulaich, F. I. Khamlichi and M. Kaicer, « Microcredit scoring using fuzzy logic model », International Journal of Statistics & Economics (IJSE) (ISSN: 0975-556X), Volume 10, Number 1, 2013.
- [14] F. I. Khamlichi, R. Aboulaich and A. E. El Mrhari, « Global Decision Strategy based on combining multiple regression method

Spread of the Corona virus Epidemic in Morocco, USA, France, Spain, Italy, Egypt, Lebanon and Tunis by the Heteroscedastic Model during the first wave

Dounia Bouchta

Faculty of Sciences, University
Abdelmalek Essaâdi

bdounia1@gmail.com

Redouan El khamlichi

Faculty of Sciences, University
Abdelmalek Essaâdi

redouan.elkhamlichi@outlook.com

ELLIAZIDI Sara

National School of Applied Sciences,
University Abdelmalek Essaâdi

ellyazidi.sara@gmail.com

Mohammed Lamarti Sefian

Normal Higher School, University
Abdelmalek Essaâdi

memedma@yahoo.fr

R.Temsamani Khalid

Faculty of Sciences, University
Abdelmalek Essaâdi

ktemsamani@uae.ac.ma

Abstract— A novel coronavirus (2019-nCoV) pneumonia hit a China in the end of 2019, and subsequently reached other countries and regions. In this paper, we present scenarios for the spread of the coronavirus epidemic by defining a coefficient called $DS = (\text{date } i / \text{date } i-1)$ that represents the daily spread of covid19.

We have demonstrated, in the first wave, by Applying the ARCH Models that the spread of the coronavirus knew a calm period by the end of April and beginning of May. However, a variation in mid-May showed first the increase in infected people with new cases which our model showed with 95% of confidence, and this will be followed by a last but larger variation to appear on July 4, 2020 and followed by a decline in COVID-19 infection propagation in the following weeks.

Keywords—coronavirus, ARCH Models, daily spread coefficient

INTRODUCTION

Respiratory symptoms, Fever, flanked by cough mostly dry, gastrointestinal symptoms such as nausea, vomiting and diarrhoea and myalgia or fatigue appearing during the course of illness are mostly considered to confirm the infection with the COVID-19 [1], [2]. The totality of initial COVID-19 cases in morocco were associated to Moroccans coming from abroad especially from Europe, however, a growing number of cases due to person-to-person transmission have been reported. Since the first case detected in Morocco on March 06, 2020, several measures have been quickly adopted by the Government to slow down the spread of COVID-19; measures such as the suspension of international flights, closure of schools, universities, cultural and sporting places, cafes, restaurants, non-essential businesses and mosques to limit the displacement of the population as much as possible were taken. Morocco declared a state of health emergency on March 20, 2020.

1. Methodology

1) The diffusions are random functions, which are widely used in physics, chemistry, biology, statistics and finance. it captures instant dynamics tainted with uncertainty.

Their very nature makes them a great modelling tool: Well beyond their descriptive interest, they lend themselves to quantitative uses. In this paper, we propose to model the new effects following a geometric Brownian movement (most usually in finance to model the share price evolution).

An autoregressive process is a process in which each value is described as a linear combination of the previous values plus a random component called a "noise". The number of previous values considered is called noise of the process.

2) The classical Autoregressive Moving Average model, based on AR model, trying to model the conditional expectation, are fundamentally linear and are based on weak constraints (constant mean, unconditional variance and constant conditional variance), have chosen which restrict their application in fields where the series with certain dynamic characteristics or constant for the conditional variance.

3) The variance σ_t at time t is connected to the value of the series y_{t-1} at time $t - 1$. A relatively large value of y_{t-1}^* gives a relatively large value of the variance at time t . This means that the value of y_t is less predictable at time $t - 1$ than at times after a relatively small value of y_{t-1}^*

A GARCH (p, q) model (generalized autoregressive conditionally heteroscedastic) uses values of the past squared observations and past variances itself to model the variance at time t . We give the model by:

$$y_t = \sigma_t \varepsilon_t$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i y_{t-i}^2 + \sum_{i=1}^q \beta_i \sigma_{t-i}^2$$

($\varepsilon_t, t \in \mathbb{Z}$) distributed normally $N(0, 1)$

The choice of model and its validation are based on the skewness, the kurtosis and taking some tests [6]; Jarque-

Bera test [7], Ljung-Box (portemanteau) test [8], heteroscedastic test [9].

2. Daily Spread Coefficient of Corona virus in the world after confinement

After two months of detecting the first case in Morocco, the number of patients increased exponentially. We have noticed, between the first day detected case and May 03, 2020, an average growth daily rate increase of 15% in Morocco (first day March 03), USA 14% per day (Jan 21), Spain 14% per day (Feb 01), Italy 13% per day (Jan 31), France 11% per day (Jan 25), while the world average is 10% per day (Dec 31, 1919), Egypt 12% per day (Feb 15), Lebanon 10% per day (Feb 15), Tunis 12% per day (March 03).

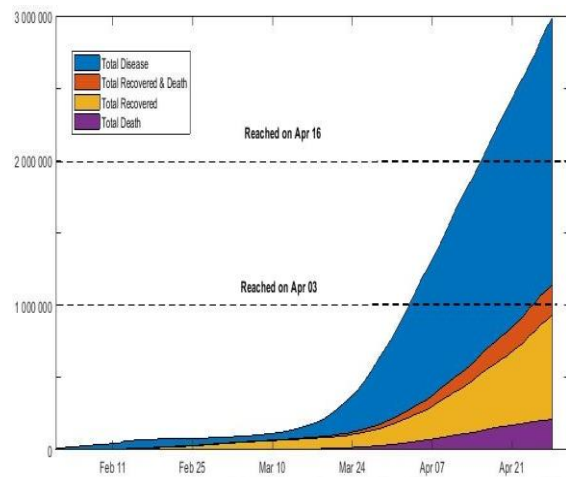


Fig 1. Global coronavirus cases in the world

Some viruses are highly contagious (spread easily), like measles, while other viruses do not spread as easily, the experts based on knowledge about other coronaviruses, believe that the virus is thought to spread mainly from person-to-person, between people who are in close contact with one another (within about 6 feet). Through respiratory droplets produced when an infected person coughs or sneezes. These droplets can land in the mouths or noses of people who are nearby or possibly be inhaled into the lungs. The spread might be possible before people show symptoms, but this is not thought to be the main way the virus spreads. while Patients are felt to be at highest risk of spreading the illness when they are most symptomatic [9]-[10].

The transmission of COVID-19 It may be possible through contaminated surfaces or fomites with subsequent contact with the eyes, nose, or mouth may also occur [2], [10], [11]. The COVID-19 seems spread easily and sustainably in people have been infected with the virus in an area, including some who are not sure how or where they became infected.

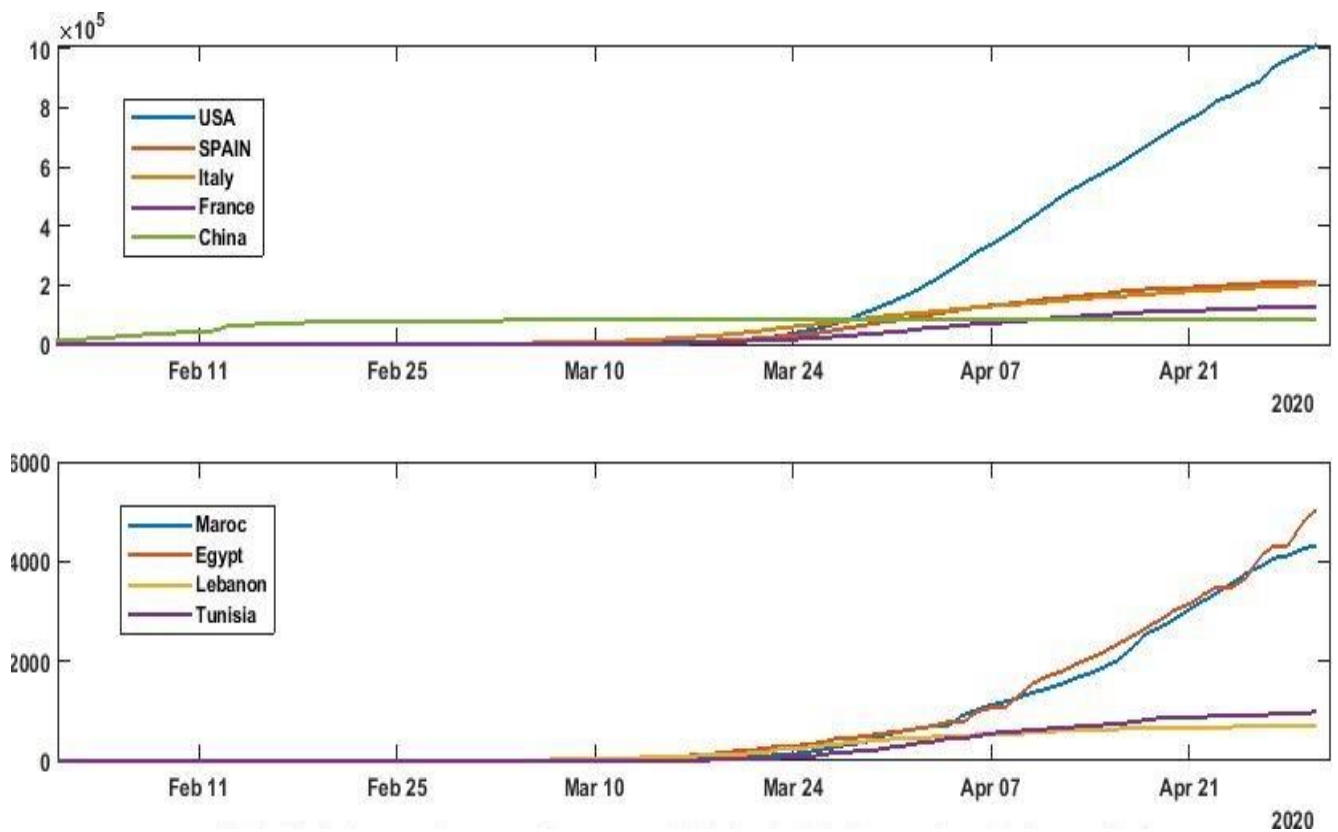
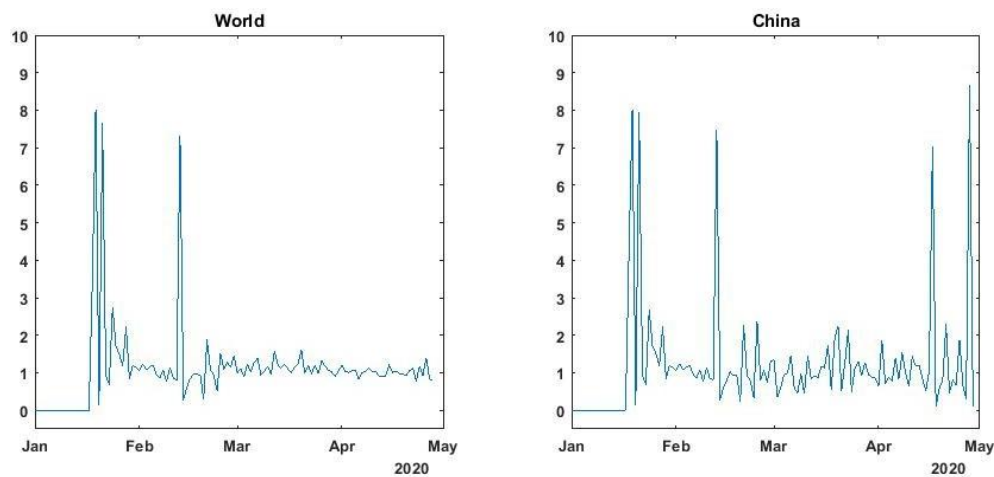
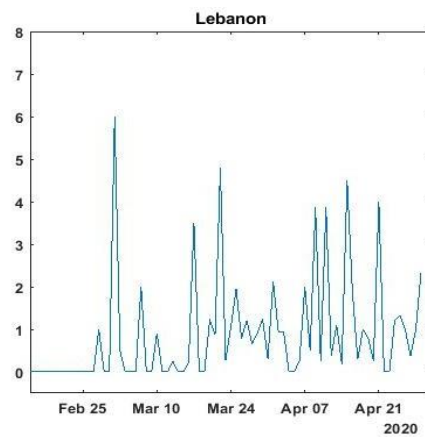
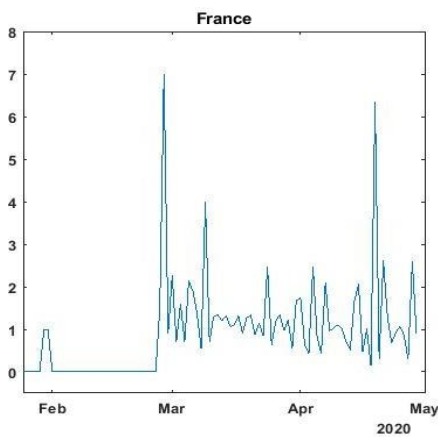
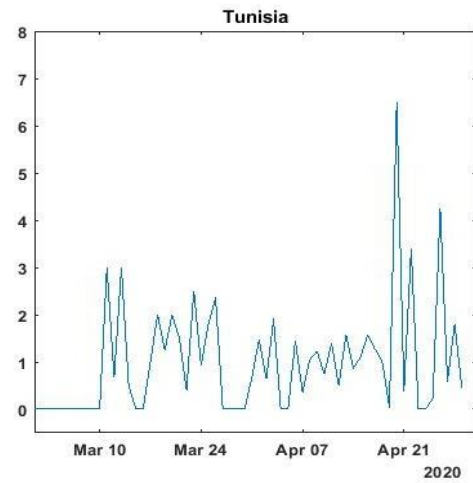
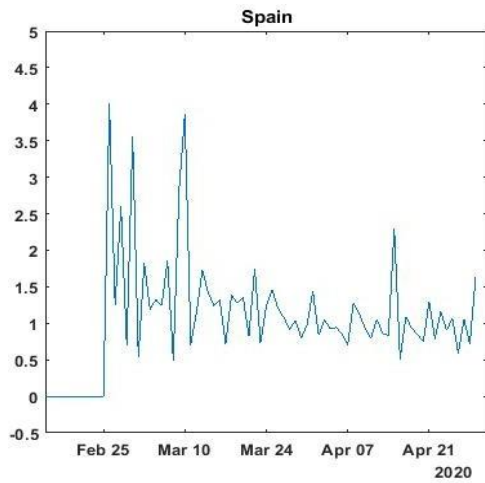
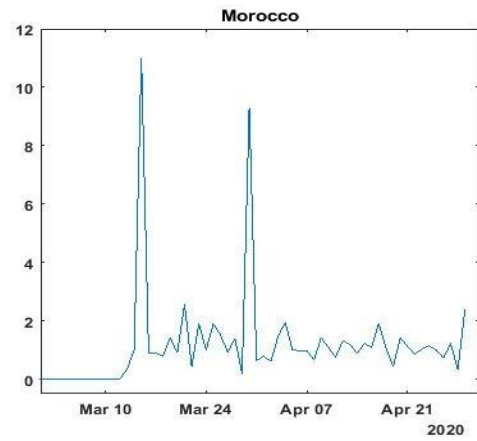
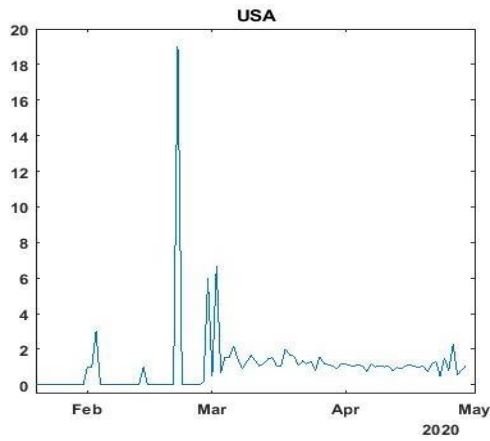


Fig 2. Global coronavirus cases in Morocco, USA, France, Spain, Italy, Egypt, Lebanon, Tunis, China.

We consider that a chronological series (X_t) results in the spread of COVID-19 in the different countries studied, we have defined the coefficient $DS = (\text{date } i / \text{date } i-1)$ represents the daily propagation or the daily spread of covid19 in the studied series, It reflects the “average” behaviour of each series. Fig 3. Shows The modelling of the daily proportion rate given by DS which can give a prediction of case, the figures show the existence of two waves of propagation, the first wave in the 1st 10 days, followed by a very important second wave in a month after the first.





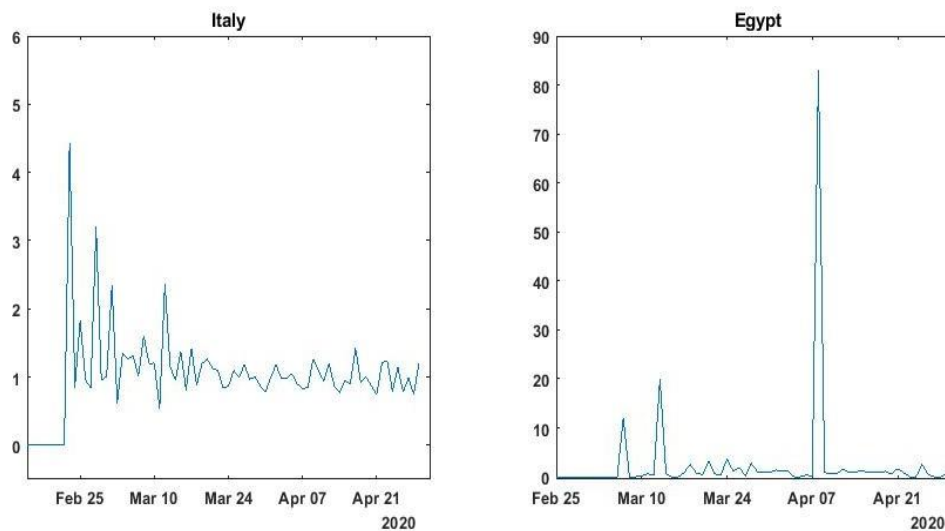


Fig 3. The modelling of the daily spread rate (DS) in the World, Morocco, USA, France, Spain, Italy, Egypt, Lebanon, Tunis and China.

3. 3. Scenarios for the Spread of the Corona virus epidemic in the USA, Morocco, Italy, France, Spain, Egypt, Lebanon and Tunis.

3.1 Ljung-Box Q-Test

By this test, we search that any of a group of autocorrelations of a time series are different from zero;

Null Hypothesis: The first autocorrelations of DS are

3.2 Engle's ARCH Test

In this test, we seek the non-attendance of homoscedasticity (contraire of Heteroscedasticity);

Null Hypothesis: The series exhibits no ARCH Effects.

We present results when the Null Hypothesis rejected is true.

jointly 0.

We present results when the Null Hypothesis rejected is true.

Tab1. Ljung-Box Q-Test

Tab2. Engle's ARCH Test

p+q give us the parameters of the model We will present two best GARCH model for that p+q using Akaike Information Criterion AIC and Bayesian information criterion in one hand and considering that the values follow a Gaussian or t distribution

Country	Test Statistic	Critical Value	Lags	DOF	Significance Level
Morocco	18.1654	16.919	15	9	5%
Tunis	25.2452	3.8415	1	1	5%
Egypt	4.5413	3.8415	14	1	5%
Lebanon	22.7886	3.8415	1	1	5%
Italy	5.6798	3.8415	3	1	5%
France	25.544	3.8415	1	1	5%
Spain	8.3383	3.8415	1	1	5%
USA	6.0391	3.8415	7	1	5%
World	7.597	3.8415	2	1	5%

Country	Test Statistic	Critical Value	p+q	Significance Level
Morocco	27.9811	24.9958	15	5%
Tunis	13.9744	3.8415	1	5%
Egypt	13.7833	11.0705	5	5%
Lebanon	9.3811	3.8415	1	5%
Italy	15.4012	9.4877	4	5%
France	36.5176	3.8415	1	5%
Spain	6.7769	3.8415	1	5%
USA	69.9164	33.9244	22	5%
World	20.4776	5.9915	2	5%

3.3 Models :

a- Gaussian distribution based maximum-likelihood estimation:

Tab 3. Model estimated parameter using Gaussian distribution

Parameter	Morocco	Tunis	Egypt	Lebanon	Italy	France	Spain	USA	World
Constant	0.070184	1.9538	0.15473	2.4592	0.0025836	0.43967	0.08569	0.1768	0.00016272
GARCH{1}					0.82001		0.96933		0.89252
GARCH{3}	0.0031843								
GARCH{4}	0.0019962								
ARCH{1}		0.51403		0.41506		1			0.088364
ARCH{2}			0.48639		0.075852			0.86879	
ARCH{3}	0.068107		0.51361						
ARCH{4}								0.0010096	
GARCH{5}	0.1069								
ARCH{7}								0.1302	
ARCH{10}	0.88792								
Offset	0.10636	0.16841	0.85848	0.23457	1.0037	0.14991	0.04082	1.5017	1.0428
AIC	151.2633	215.6739	182.3509	273.5607	62.424	303.2073	89.7999	356.2572	189.4955
BIC	162.2352	221.5277	190.3802	280.0373	71.3604	310.9314	96.414	II. 368.5289	200.2974

b- t-distribution based maximum likelihood estimation:

tab 4. Model estimated parameter using

t-distribution:

Parameter	Morocco	Tunis	Egypt	Lebanon	Italy	France	Spain	USA	World
Constant	2e-07	1.5242	0.5522	2.5586	0.0069823	0.22314	0.097588	0.019949	0.02617
GARCH{1}					0.79627				0.66599
GARCH{3}	0.82225							0.0047325	
GARCH{4}								0.0053119	
ARCH{1}		0.94583		1		1	0.86752	0.77358	0.11768
ARCH{2}			0.10931		0.076946			0.0067106	
ARCH{3}			0.89069					0.0048603	
ARCH{4}								0.0033048	
ARCH{5}	0.0096237								
ARCH{6}	0.040208							0.0010572	

ARCH{7}								0.062467	
ARCH{9}								0.059231	
ARCH{13}								0.0045191	
ARCH{17}								0.0015503	
ARCH{18}								0.072671	
DoF	2.208	3.7866	2.6615	2.7503	3.2057	2.7173	6.2549	3.1191	2.357
Offset	1.119	0.024345	0.79681	-0.0072848	0.99098	-0.017817	0.037495	1.0586	1.0526
AIC	111.8716	209.169	153.3543	264.7403	33.2396	259.2154	89.9538	189.5156	49.7467
BIC	123.3437	216.974	163.391	273.3759	44.4101	269.5143	98.7726	224.2779	63.2491

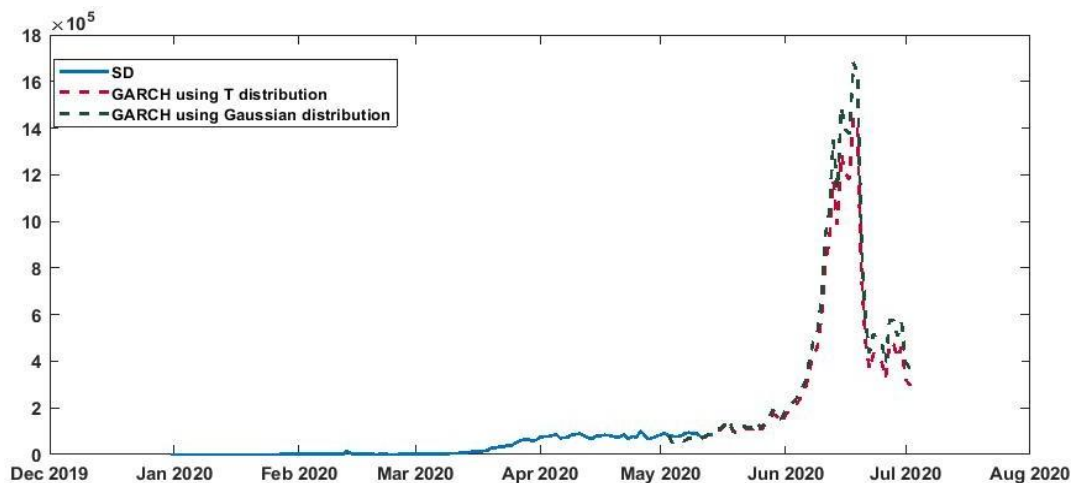
3.4 prevision for the spread of the coronavirus epidemic:

By Applying the GARCH(p,q) Conditional Variance Model with Offset with Gaussian distribution and t-distribution. Fig.4 present the graphs which show the evolution of the new cases until May 14, 2020 and the prevision of the studied countries using both of model estimated since May 03, 2020. We can notice, based on the two models, and although some countries, which in the Covid-19 spread tend to be extinguished,

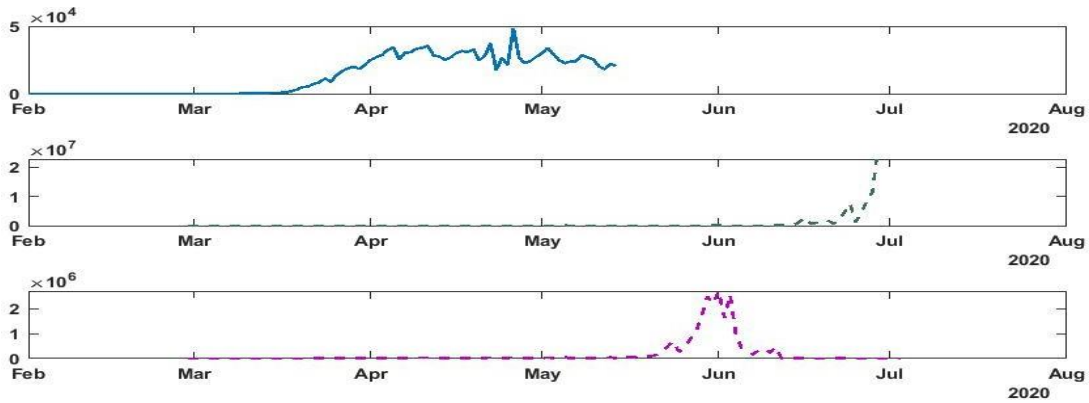
the month of June may have consequence on its spread. By the World, and with 95% of confidence an increasing data

will be held on July 4, 2020 followed by the decrease of the virus. This variation may be different between countries depending on the strategies adopted and strict compliance (France, Spain, Morocco) and also depending on the size of country's population (Tunis Lebanon).

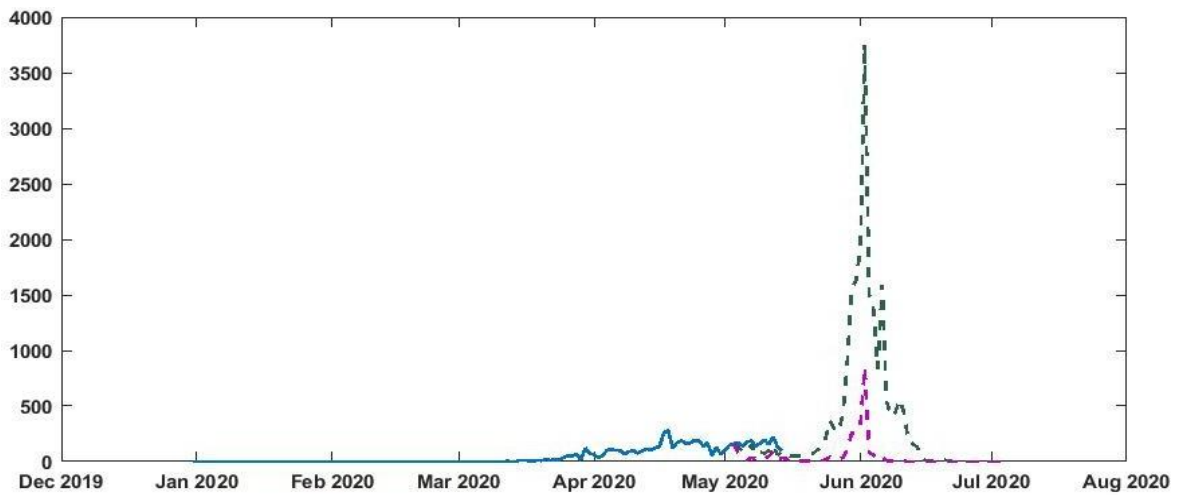
a- world



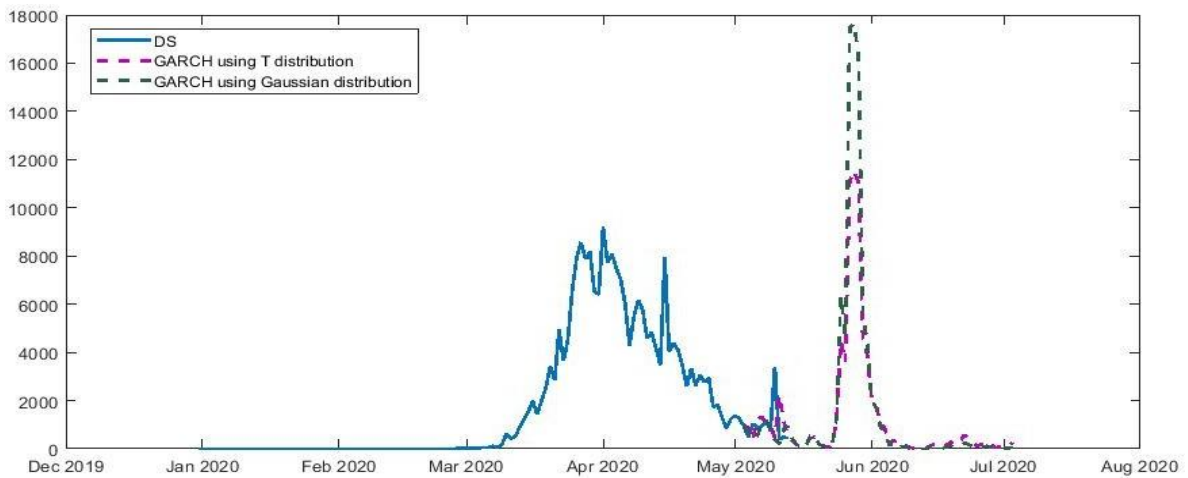
b- USA



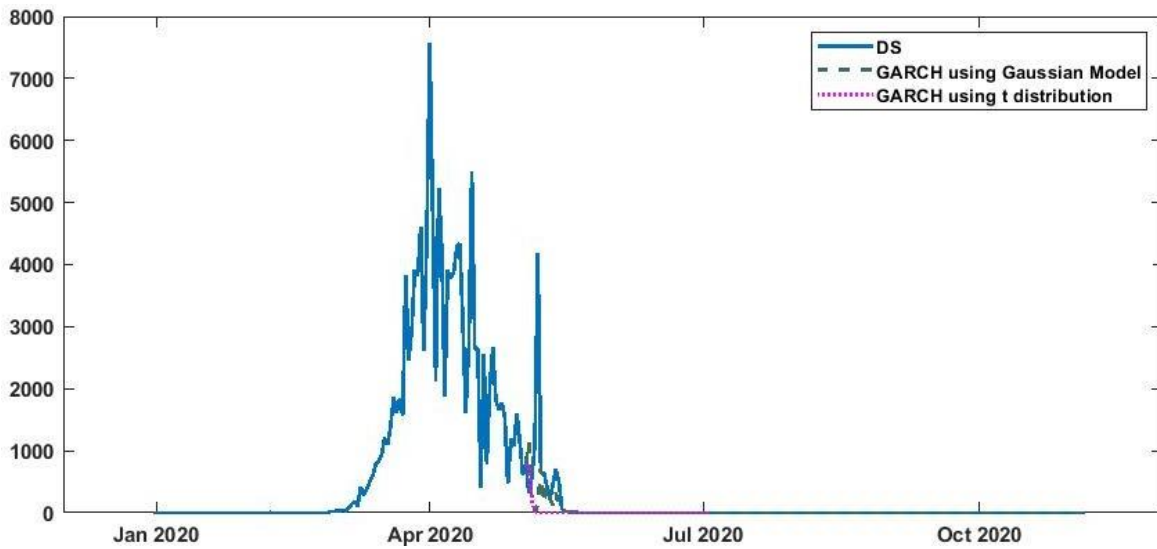
c- Morocco



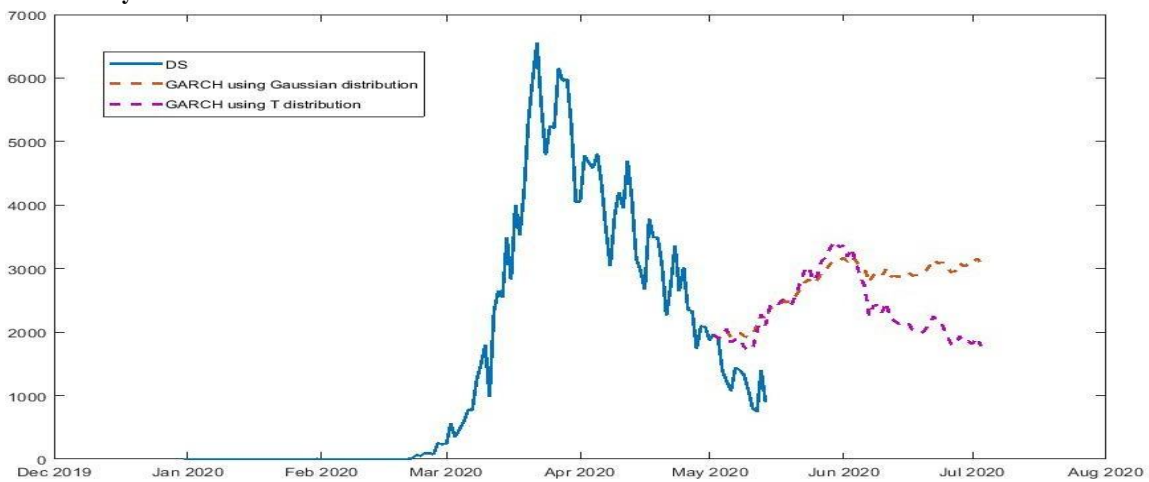
d- Spain



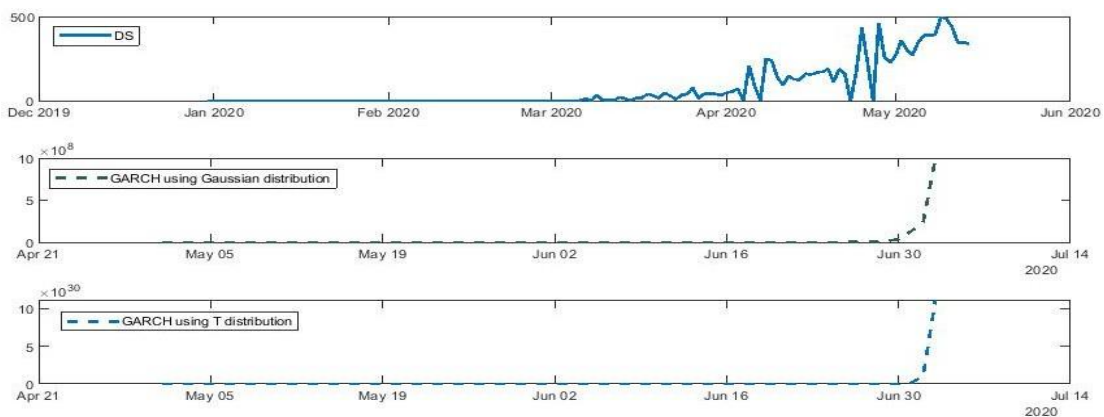
e- France



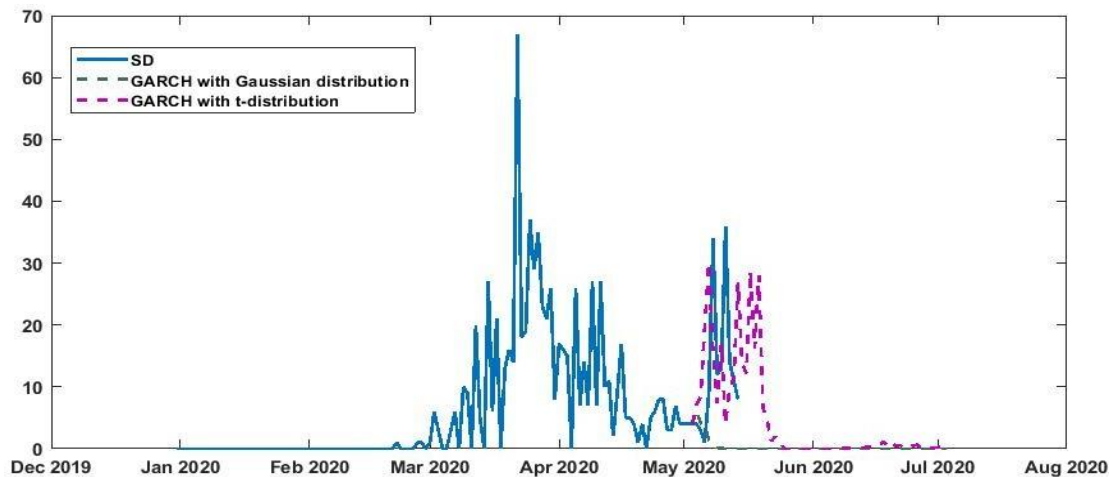
f- Italy



g- Egypt



h- Tunis



i- china

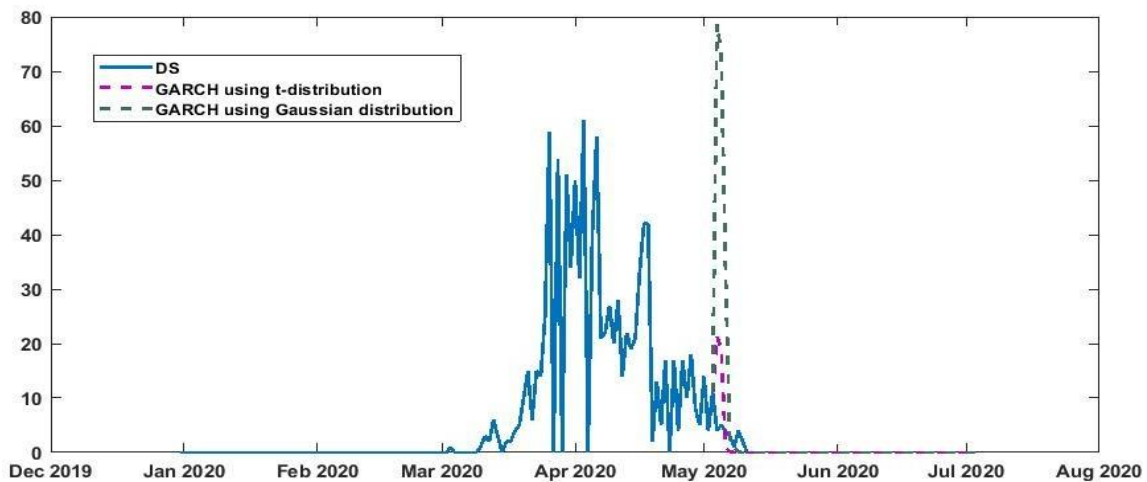


Fig 4. Scenarios for the spread of the coronavirus epidemic in the World, Morocco, USA, France, Spain, Italy, Egypt, Lebanon, Tunis and China.

CONCLUSION

Most of scenarios that try to model the COVID-19 evolution of new cases, are based on Gaussian distribution. We present in this study two scenarios of GARCH model using Gaussian and t- distribution. Both models anticipate an extinction of the new cases before the beginning of July provided strict respect of the measures taken (France, Spain, Morocco, and Tunis, Lebanon). However, global variation in countries like USA and Egypt need to take more severe measures to slowdown the spread of this virus. In fact, the failing in total extinction of the virus before the fall can cause a reactivation of this one especially if the measure of closing of the borders is lifted.

REFERENCES

[1] S. Chavez, B. Long, A. Koyfman, and S. Y. Liang, "Coronavirus Disease (COVID-19): A primer for emergency physicians," *Am. J. Emerg. Med.*, Mar. 2020, doi: 10.1016/j.ajem.2020.03.036.

[2] Q. Li *et al.*, "Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia," *N. Engl. J. Med.*, Jan. 2020, doi: 10.1056/nejmoa2001316.

[3] E. R., "Autoregressive Conditional Heteroscedasticity with estimates of the variance of united kingdom inflatio," *econometrica*, vol. 50, no. 4, pp. 987–1008, 1982.

[4] R. F. Engle, "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 1982, doi: 10.2307/1912773.

[5] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," *J. Econom.*, 1986, doi: 10.1016/0304-4076(86)90063-1.

[6] "Modèles ARCH et application financières," *Gourrieroux C*, p. 46, 1992.

[7] R. Sneyers, "Sur les tests de normalit'e," *Rev.*

- Stat. Appliquée*, vol. 22, pp. 29–36, 1974.
- [8] G. M. Ljung and G. E. P. Box, “On a measure of lack of fit in time series models,” *Biometrika*, 1978, doi: 10.1093/biomet/65.2.297.
- [9] B. R., “Econométrie,” no. 2002, p. 147.
- [10] R. S. Wax and M. D. Christian, “Practical recommendations for critical care and anesthesiology teams caring for novel coronavirus (2019-nCoV) patients,” *Canadian Journal of Anesthesia*. Springer, pp. 1–9, 12-Feb-2020, doi: 10.1007/s12630-020-01591-x.
- [11] M. Batista, “Estimation of the final size of the coronavirus epidemic by the logistic model,” *medRxiv*, no. March, p. 2020.02.16.20023606, 2020, doi: 10.1101/2020.02.16.20023606.
- [12] S. Zhao *et al.*, “Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak,” *Int. J. Infect. Dis.*, 2020, doi: 10.1016/j.ijid.2020.01.050.
- [13] “Covid-19. En graphiques comparés avec 3 pays, la réaction précoce du Maroc.” <https://www.medias24.com/covid-19-en-graphiques-compare-avec-3-pays-la-reaction-precoce-du-maroc-8940.html> (accessed May 21, 2020).
- [14] “La fin de l’épidémie de Covid-19 est encore lointaine en France, calcule l’Institut Pasteur - Science & Vie.” <https://www.science-et-vie.com/corps-et-sante/selon-les-calculs-de-linstitut-pasteur-lepidemie-de-covid-19-en-france-est-loin-55370> (accessed May 21, 2020).
- [15] R. R. Assessment, “Covid-19-Rapid-Risk-Assessment-Coronavirus-Disease-2019-Eighth-Update-8-April-2020,” vol. 2019, no. April, 2020.

AUTHORS INDEX

Abdelilah Jraifi
Adnane Addaim
Aouatif Amine
Aziz Chihab
Aziz Darouichi
Belaid Bouikhalene
Chakib Elmokhi
Dounia Bouchta
Fahd Idrissi Khamlichi
Khalid Tamsamani
Hajar Sabiki
Hamid Sadiki
Hamza EL-Houari
Hanaa Hachimi
Hassan Faouzi
Hicham Moussa
Houssam Benbrahim
Idris Bakhadach
Ilias Elmouki
Imad El Hassak
Lalla Saadia Chadli
Maria Zemzami
Mohamed Fakir
Mohamed Gouskir
Mohammed Boutalline
Mohammed Kaicer
Mohammed Lamarti Sefian
Nabil Hmina
Norelislam EL Hami
Nour-Eddine M'haouache
Radoine Hamzaoui
Redouan El khamlichi
Saadia Chadli
Said Melliani
Sara Elliazidi
Sara L'ghdaiche
Younna Elhissi

ABOUT THE EDITOR IN CHIEF

Prof. Dr. Hanaa Hachimi, Ph.D in Applied Mathematics & Computer Science and a Ph.D in Mechanics & Systems Reliability, Secretary General of Sultan Moulay Slimane University in Beni Mellal. President of the Moroccan Society of Engineering Sciences and Technology (MSEST). I am Associate Professor at the Sultan Moulay Slimane University (USMS) of Beni Mellal, Morocco. I am Director of the Systems Engineering Laboratory (LGS) and IEEE Senior Member, precisely I am affiliated at the Big Data, Optimization, Service and Security (BOSS) team at USMS. I am Lecture and Keynote Speaker of the courses: Optimization & Operational Research, Graph Theory, Statistics, Probability, Reliability and Scientific Computing. I am Member of the Moroccan Society of Applied Mathematics (SM2A).

Previously Associate Professor at Ibn Tofail University, National School of Applied Sciences of Kenitra, Morocco.

for more information, visit our website : <http://www.hanaahachimi.com/>

EDITORIAL BOARD

Editorial bord:

- Prof. Dr. Abou El Majd Badr (FS, Rabat, Morocco)
Prof. Dr. Addaim Adnane (EMI, Morocco)
Prof. Dr. Amine Abdellah (USMS, Beni Mellal, Morocco)
Prof. Dr. Amlan Chakrabarti (Director A.K. Choudhury School Of I.T., University Of Calcutta, India)
Prof. Dr. Assif Safaa (ENSA, El Jadida, Morocco)
Prof. Dr. Bakhadach Idris (USMS, Beni Mellal, Morocco)
Prof. Dr. Belhouideg Soufiane (FP, Beni Mellal, Morocco)
Prof. Dr. Ben Maissa Yann (INPT, Ieee Senior Member, Rabat, Morocco)
Prof. Dr. Benterki Djamel (Setif University, Algeria)
Prof. Dr. Bouloiz Hafida (ENSA, Agadir, Morocco)
Prof. Dr. Boutalline Mohammed (USMS, Beni Mellal, Morocco)
Prof. Dr. Chadli Lalla Saadia (FST, Beni Mellal, Morocco)
Prof. Dr. Chaoui Habiba (ENSA, Kenitra, Morocco)Prof. Dr. Saidi Rajaa (Insea, Rabat, Morocco)
Prof. Dr. Darouichi Aziz (FST, Marrakech, Morocco)
Prof. Dr. Driss Mentagui (Faculte Des Sciences, UIT, Kenitra, Morocco)
Prof. Dr. El Abbadi Laila (ENSA, Kenitra, Morocco)
Prof. Dr. El Hami Abdelkhalek (INSA, Rouen, France)
Prof. Dr. El Hissi Youmna (USMS, Beni Mellal, Morocco)
Prof. Dr. El Mokhi Chakib (UIT, Kenitra, Morocco)
Prof. Dr. Ellaia Rachid (EMI, Rabat, Morocco)
Prof. Dr. Farouk Yalaoui (Directeur Du Laboratoire D'optimisation Des Systemes De Production (Losi), UTT, Troyes, France)
Prof. Dr. G. Suseendran (Vistas, India)
Prof. Dr. Hanaa Hachimi (USMS, Morocco (Ieee Senior Member (Ieee Senior Member))
Prof. Dr. Hammadi Nait Charif (Universite De Bournemouth, Royaume-Uni)
Prof. Dr. Hmina Nabil (USMS, Morocco)
Prof. Dr. Ibrahim Rosziati (Uthm University, Malaysia)
Prof. Dr. Ing. Andrei Victor Sandu (Gheorghe Asachi Technical University Of Iasi, Romania)
Prof. Dr. Jensen Nils (Ostfalia, Wolfenbüttel, Germany)
Prof. Dr. Jihane Farahat (President Of Egyptian Center Of Innovation And Invention, Cairo, Egypt)
Prof. Dr. Jraifi Abdelilah (ENSA, Safi, Morocco)
Prof. Dr. Kaicer Mohammed (FS, Kenitra, Morocco)
Prof. Dr. Lakhout Abderrahim (Usherbrooke, Canada)
Prof. Dr. Madini-Zouine Zhour (ENSA, Kenitra, Morocco)
Prof. Dr. Maslouhi Mustapha (ENSA, Kenitra, Morocco)
Prof. Dr. Masulli Francesco (University Of Genova, Italy)
Prof. Dr. Mehar Chand (Guru Kashi University bathinda, India)
Prof. Dr. Mejri Mohammed (Ulaval, Quebec, Canada)
Prof. Dr. Melliani Said (FST, Béni Mellal, Morocco)

Prof. Dr. Mraoua Mohammed (HEC, Montreal, Canada)
Prof. Dr. Nayyar Anand (Duy Tan University, Vietnam)
Prof. Dr. Obaid Abdelatif (UQAM, Canada)
Prof. Dr. Oscar Castillo (Tijuana Institute Technology, Mexico)
Prof. Dr. Petrica Vizureanu (Gheorghe Asachi Technical University Of Iasi, Romania)
Prof. Dr. Ribaud Marina (University of Genoa, Italy)
Prof. Dr. Rokia Missaoui (Universite Du Quebec En Outaouais, Canada)
Prof. Dr. Rovetta Stefano (Unig, Genova, Italy)
Prof. Dr. El Ghazali Talbi (University Of Lille, France)
Prof. Dr. Rovetta Stefano (Unige University, Genova, Italy)
Prof. Dr. Rui Lopes (Electrical Engineering Department Fct Nova And Uninova - Cts, Portugal)
Prof. Dr. Semlali Naoual (EMI, Rabat, Morocco)
Prof. Dr. Soulaymani Abdelmajid (FSK, Kenitra, Morocco)
Prof. Dr. Xin She Yang (National Physical Laboratory, oxford University, United Kingdom)
Prof. Dr. Zhoure Madini (Ibn Tofail University, Morocco)
Prof. Dr. Zouine Youness (ENSA, Kenitra, Morocco)



International Journal
on Optimization and Applications