Frontiers in Science and Engineering - International Journal Volume 2- Number 1- 2012

Frontiers in Science and Engineering
International Journal

Edited by The Hassan II Academy of Science and Technology of Morocco

Contents

Dépôt légal : 2012 PE 0007
ISSN : 2028 - 7615

ACADEMY Press MA

Email : fse@academiesciences.ma
www.academiesciences.ma/fse/

ACADEMY Press MA

Volume 2 - Number 1 - 2012

Editorial board

Frontiers in Science and Engineering
International Journal

Edited by The Hassan II Academy of Science and Technology of Morocco

Contents

Dépôt légal : 2012 PE 0007
ISSN : 2028 - 7615

ACADEMY Press MA

Email : fse@academiesciences.ma
www.academiesciences.ma/fse/

Layout by : AGRI-BYS S.A.R.L (A.U)
Printed by : Imprimerie LAWNE
11, rue Dakar, 10040 - Rabat

WELCOME TO FSE

Frontiers in Science and Engineering, an International Journal edited by Hassan II Academy of Science and Technology uses author-supplied PDFs for all online and print publication.

The objective of this electronic journal is to provide a platform of exchange of high quality research papers in science and engineering. Though it is rather of wide and broad spectrum, it is organized in a transparent and simple interactive manner so that readers can focus on their direct interest.

All papers are submitted to the normal peer-review process. Publication criteria are based on : i) Novelty of the problem or methodology and problem solving, ii) Salience of the approach and solution technique, iii) Technical correctness and outputs, iv) Clarity and organization.

Papers are first reviewed by the Executive Board Director who receives the paper and, if relevant in terms of the overall requirements, it is then proposed to one of the most appropriate associate editor on the field who will select 2 to 3 expert reviewers. We are right now in the process of completing our international Editorial Board. Electronic printing will allow considerable time savings for submission delays which will be reduced drastically to less than three to six months. Prospective authors are therefore invited to submit their contribution for assessment while subjected to similar quality criteria review used in paper journals.

Authors are notified of acceptance, need for revision or rejection of the paper. It may be noted that papers once rejected cannot be resubmitted. All the details concerning the submission process are described in another section.
This electronic journal is intended to provide :

- the announcement of significant new results,
- the state of the art or review articles for the development of science and technology,
- the publication of proceedings of the Academy or scientific events sponsored by the Academy,
- the publication of special thematic issues.

So that the scientific community can :

- promptly report their work to the scientific community,
- contribute to knowledge sharing and dissemination of new results.

The journal covers the established disciplines, interdisciplinary and emerging ones. Articles should be a contribution to fundamental and applied aspects, or original notes indicating a significant discovery or a significant result.

The topics of this multidisciplinary journal covers amongst others :
Materials Science, Mathematics, Physics, Chemistry, Computer sciences, Energy, Earth Science, Biology, Biotechnology, Life Sciences, Medical Science, Agriculture, Geosciences, Environment, Water, Engineering and Complex Systems, Science education, Strategic and economic studies, and all related modeling, simulation and optimization issues, etc. ...

Once, a certain number of papers in a specific thematic, is reached, the Academy might edit a special paper issue in parallel to the electronic version.

INTRODUCTION

Dear colleagues, scientists and potential authors

This is the second issue of the paper version of our electronic multidisciplinary International Journal edited by Hassan II Academy of Science and Technology of Morocco covering :

- Mathematics, Applied Mathematics, Computer Sciences
- Physics, Chemistry, Engineering Sciences
- Life Sciences (Medical, Health, Agriculture, Biology, Genetics)
- Earth, Water and Oceans, Environmental Sciences
- Strategic Studies and Economic Development

through its publications management system ACADEMY Press MA.

While the electronic version is organized into the themes organized within Frascati (Field of Science and Technology Classification in the Frascati Manual (OECD) and can be browsed as an hypertext allowing high resolution graphics and animations, this issue is a direct print version of the papers presented online in February 2012 globally.

We hope that our adoption of the Open Access model for our newly launched e-journal "Frontiers in science and Engineering" will stand with your help and support through your contributions so that the publications can be widely disseminated.

The Editors

Propane oxidative dehydrogenation over unpromoted and Nb promoted NiO loaded calcium-hydroxyapatite catalysts

Mariam Khachani, Mohamed Kacimi and Mahfoud Ziyad*

Université Mohammed V – Agdal, Faculté des Sciences, Laboratoire de Physico-chimie des Matériaux et Catalyse, Département de Chimie, Avenue Ibn Battouta, B.P. 1014, Rabat, Morocco.

Abstract

Nickel loaded calcium-hydroxyapatite (xNi/CaHAp) and niobium promoted xNi/CaHAp catalysts were synthesized and characterised by X-ray, FTIR, U.V-visible-NIR spectroscopies and temperature programmed reduction (H_2-TPR). X-ray diffraction patterns of xNi/CaHAp showed that for nickel loadings above 5 wt.% Ni, diffraction lines belonging to bulk NiO start to appear. The average size of those nickel oxide particles is equal to 16, 23, and 35 nm for x = 10, 15 and 20 Wt.%, respectively. U.V-visible and TPR showed that the loaded nickel is hosted by octahedral and pseudo-octahedral sites.

Calcium-hydroxyapatite loaded with different amounts of nickel and niobium promoted xNi/CaHAp catalysts were tested in oxidative dehydrogenation propane (ODH). The best performance was achieved with a nickel loading of x = 10%. The corresponding stationary conversion of propane is equal to 22% with a propylene yield of 13%.

Addition of niobium to xNi/CaHAp decreases the global conversion but enhances the propylene selective and the stability of the catalyst with time on stream. The best results were obtained with 0.15 wt.% Nb. The propylene yield at stead state reaches 14% at 425°C.

Keywords: calcium-hydroxyapatite, nickel loaded hydroxyapatite, niobium-nickel loaded hydroxyapatite, propane oxidative dehydrogenation.

*To whom correspondence should be addressed Tel: +212661909377

E-mail address: ziyad@fsr.ac.ma

1. Introduction

The increasing demand of propylene and the numerous disadvantages of the actually employed process (steam cracking of naphtha or propane) justify the intensive work carried out in order to develop alternative ways for the production this gas (1-3). The high cost of olefins has also prompted the development of new catalysts for oxidative dehydrogenation (ODH) of propane into propylene. The oxygen in the reaction mixture favours the continuous removal of carbon deposits from the catalyst surface, stabilizes the reaction rates over a wide temperature range, and leads to the formation of water rather than hydrogen (4,5). Various catalytic systems were investigated including molten alkali salts (6), bulk mixed metal oxides (7-10), heteropolyacid catalysts and molecular sieve-based catalysts (11-15). It was observed that propane is more reactive than methane and ethane. It can be oxidized at temperatures as low as 300°C. Beside the oxidative dehydrogenation, the reaction between oxygen and hydrocarbons produce oxygenated compounds or lead to total oxidation suggesting that the selectivity to propylene usually remains moderate (16). In that perspective search for systems promoting selectively oxidative dehydrogenation is a priority, particularly because the conversions and the yields claimed in the literature do not seem to be encouraging for industrial applications.

Phosphate-based catalysts were studied in many reaction types. Loaded with nickel, chromium, iron or palladium, calcium-hydroxyapatite $Ca_{10}(PO_4)_6(OH)_2$ exhibits good activity in methane oxidation and in direct synthesis of methyl isobutyl ketone (17-18). In ethane ODH, cobalt exchanged (Co^{2+}/Ca^{2+}) hydroxyapatite $Ca_{10-x}Co_x(PO_4)_2(OH)_2$ showed an ethylene yield of 22% at 550°C. This good performance was attributed to isolated Co^{2+} sites on the apatite surface and to lattice oxygen mobility induced by cobalt incorporation in the phosphate network. On the other hand, the performance of CaHAp as a carrier is essentially due to its ion-exchange and basic properties (19,20).

Savary *et al*. have studied propane ODH over V-P-O/TiO_2 at 400°C. They achieved a conversion of 9% and a propylene selectivity of 56% (21). X. Zhang *et al*. reported that silver doping of Mo-P-O enhances its catalytic performance because the number of active sites and the reducibility of the catalyst were increased (22). The catalytic performance of Zr-P-O catalysts in ethane ODH was improved by doping it with chromium. The observed improvements were attributed to the modification of the redox properties of the catalysts (23). It was also established that Mg-V-O is quite active in propane ODH (24). Aaddane *et al*.

showed that Mg-Co-PO_4 which has structural similarities with Mg-VO_4 exhibits a good behaviour in ethane and propane ODH (25).

Nickel oxide based catalysts were extensively studied in many processes, such as ODH reactions and natural gas reforming (26-30). In ethane ODH, they exhibit as reported by X. Zhang *et al*. very promising results (31-34). Promoted with niobium, nickel oxides Ni–Nb–O exhibit high activity and selectivity (90%) in ethane ODH, resulting in an overall ethylene yield of 46% at 400°C (35-37). In an earlier report, it was shown that Ni loaded hydroxyapatite is very efficient in methane dry reforming (17).

The present work is devoted to the study of the catalytic behaviour of nickel and niobium loaded calcium-hydroxyapatite (CaHAP) in the propane ODH. The choice of the CaHAp as a carrier was motivated by its basic properties. The selective ODH of light alkanes requires a catalyst that ensures an easy desorption of the olefin molecule before it undergoes total oxidation. Therefore, systems with low surface acidity but with basic properties should display weak bonding with the olefin. FTIR, Diffuse reflectance spectroscopy and Temperature Programmed Reduction were used in order to characterise the catalysts and correlate their surface properties with the catalytic activity.

2. Experimental

2.1 Preparation of the catalysts

The CaHAP was prepared using an ammoniacal solution of $(NH_4)_2HPO_4$ (7.92 g of $(NH_4)_2HPO_4$ in 100 mL of distilled water and 70 mL of ammoniac, 25%) was poured in a second solution of calcium nitrates (23.6 g in 100 mL of ammoniac (20 ml, 25%)) while stirring at 80°C. The pH was maintained at 9 during the mixing process. After the addition (1h), the mixture was kept for 24h at 80°C. The recovered solid by filtration was washed with hot water and dried in rotating oven at 120°C.

The nickel loaded hydroxyapatite xNi/CaHAp was synthesized utilising the conventional impregnation method using nickel nitrates dissolved in ammonia in order to inhibit the exchange ability of the calcium-hydroxyapatite. The recovered solid was dried at 120°C and then calcined at 450°C under air.

The preparation of Nb modified xNi/CaHAp catalysts was carried out using a method that prevents the precipitation of agglomeration of nickel nitrate with niobium oxalate before they spread on the support. The solution of niobium was prepared by pouring one gram of hydrated niobia (Nb_2O_5, nH_2O) into 15 mL of warm distilled water and 20 mL H_2O_2 (30%) and adding slowly to the mixture 10 mL of ammonia (25%). Portions of this solution of niobium were

mixed with different amounts of ammoniacal solutions of nickel nitrates in order to get different Nb/Ni ratios. The resulting mixtures were heated in reflux for 1h and then dried. The resulting products were calcined successively at 300 and 450°C in air for 2 h. The samples were named $xNi_{(1-y)}Nb_y$/CaHAp, where x is the weight percentage of Ni and y the incorporated niobium. The catalysts contain 1 to 20 wt.% of nickel and $0 \leq y \leq .0.6$ of niobium.

2.2 Characterization techniques

Chemical analyses of the catalysts were carried out using atomic absorption for calcium and colorimetry for phosphorus. Nickel and niobium were determined by inductive coupling plasma atomic emission spectroscopy (ICP-AES).

X-ray diffraction patterns were recorded with a Siemens D5000 high-resolution diffractometer using Ni-filtered Cu Kα radiation. The data were collected at room temperature with a 0.028 step size in 2θ, from 2θ = 20 to 60°. Crystalline phases were identified by comparison with ICSD reference files.

The textural features and BET surface areas of the samples were determined by equilibrium adsorption of N_2 at 77 K with a micromeritics apparatus. The specific surface areas were calculated by applying the BET equation.

FTIR transmission spectra were recorded between 400 and 4000 cm^{-1} at room temperature on a Perkin-Elmer 1600 spectrometer using self-supporting disks of the samples diluted in KBr. This technique is sensitive to the presence of carbonates and pyrophosphates in the samples. It can therefore be used to provide information on the purity.

Diffuse reflectance spectra were recorded at room temperature between 190 and 2500 nm on a Varian Cary 5E spectrometer equipped with a double monochromator and an integrating sphere coated with polytetrafluoroethylene (PTFE). PTFE was also used as a reference.

Temperature programmed reduction (TPR) was chosen to study the Ni and the Ni-Nb loaded CaHAp. This technique can provide valuable information on the location of nickel and its interactions with the carrier. TPR profiles were recorded at atmospheric pressure using a microreactor containing 0.030 g of catalyst and thermal conductivity detector (TCD) to determine the consumption of H_2 from a mixture of H_2/Ar = 2.5/40 introduced in the reactor at total flow rate equal to 42.5 $cm^3.min^{-1}$. The heating rate was chosen equal to 10°C.min^{-1}.

2.3. Catalytic tests

The oxidative dehydrogenation of propane into propylene was carried out in a fixed bed U-shaped reactor operated at atmospheric pressure. The apparatus and the conditions have been

described elsewhere (17-20). In typical experiments, the micro-reactor containing 0.050 g of the catalyst (particle size 120-180 μm) was heated in a flow of synthetic air at 450°C. Then, the catalyst was allowed to stabilize for about 1 h at the reaction temperature before the introduction of the reaction mixture constituted by 3.6/1.8/91 vol.% of propane, oxygen and nitrogen respectively. The total flow rate is 60 mL min^{-1}. The reaction products were analyzed by gas chromatography. The major products formed under these reaction conditions were propylene, CO, CO_2 and cracking products (methane, ethane, ethylene).

3. Results and discussion

3.1. Catalysts characterization

3.1.1 Chemical analysis and specific surface area

Table I reports specific surface areas of the samples calcined for 12 h at 450°C. The hydroxyapatite exhibits a specific surface area similar to that previously measured in other investigations (17-20). Addition of nickel and niobium to CaHAp does not modify significantly the specific surface areas of the catalysts.

Table I

Chemical analysis and specific surface areas of the catalysts.

Catalysts	S (m^2/g)	Ca (wt.%)	P (wt.%)	Ni (wt.%)	Nb (wt.%)	Ca/P
CaHAp	59	39.2	18.3	0	-	1.62
1.5Ni/CaHAp	52	38.9	18.4	1.2	-	1.63
2.5Ni /CaHAp	47	39.3	18.2	2.35	-	1.66
5Ni /CaHAp	48	39.1	18.5	4.94	-	1.63
10Ni /CaHAp	46	39.5	18.1	9.9	-	1.67
12.5Ni /CaHAp	47	39.6	18.3	12.3	-	1.66
15Ni /CaHAp	47	39.4	18.2	15.2	-	1.66
20Ni /CaHAp	50	39.2	18.2	20	-	1.67
10 $Ni_{0.95}Nb_{0.05}$CaHAp	51	39.0	18.2	-	0.05	1.66
10 $Ni_{0.9}Nb_{0.1}$CaHAp	60	39.2	18.3	-	0.1	1.67
10 $Ni_{0.85}Nb_{0.15}$CaHAp	60	39.4	18.1	-	0.15	1.67
10 $Ni_{0.8}Nb_{0.2}$CaHAp	61	39.3	18.2	-	0.2	1.67
10 $Ni_{0.7}Nb_{0.3}$CaHAp	63	39.1	18.1	-	0.3	1.67
10 $Ni_{0.4}Nb_{0.6}$CaHAp	64	39.2	18.1	-	0.6	1.67

The chemical analysis (Table I) of CaHAp and xNi/CaHAp shows that the Ca/P molar ratio is between 1.61 and 1.67, indicating that the samples are slightly calcium deficient or in agreement with the stoichiometric value of calcium-hydroxyapatite (Ca/P = 1.67). These low variations of Ca/P confirm that the Ni^{2+}/Ca^{2+} exchange is probably very small (17-20,38). In fact, at the chosen impregnation pH = 9, the isoelectric point of the calcium hydroxyapatite is

of about 6.5 and therefore, the exchange process cannot, in such conditions, completely take place. On the other hand, the amounts of loaded nickel are quasi equal to those initially introduced into the solutions.

3.1.2 X-ray analysis

The powder diffraction pattern of the support (Fig. 1) showed prominent peaks well resolved which were all attributed to the hexagonal crystal structure of the hydroxyapatite (JCPDS 09-0432) (17-20, 38).

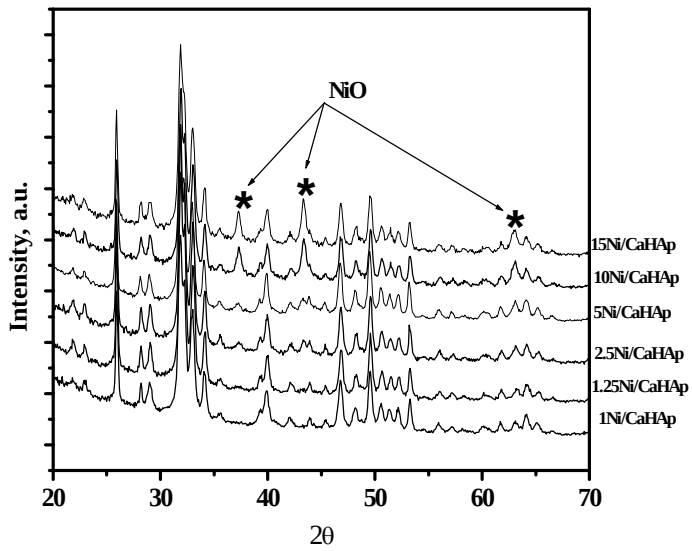


Fig. 1: XRD patterns of xNi/CaHAp.

The xNi/CaHAp samples containing less than 2.5 wt.% Ni showed only diffraction lines belonging to calcium-hydroxyapatite, suggesting that the added nickel is amorphous and well-dispersed on its surface (17-20). As the nickel loading is increased up to 5 wt.%, diffraction lines belonging to bulk NiO start to appear at 2θ = 37.3, 43.3, 62.9° and become more intense for 10 and 20 wt.% as the particles grow larger. Their size calculated using Scherrer's equation, increases with the amount of loaded nickel (39). Indeed, the average size of those nickel oxide particles is equal to 16, 23, and 35 nm for x = 10, 15 and 20 respectively. It can be noticed that in a previous work, particles of comparable size were observed using MET measurements (17).

The diffraction patterns of $10Ni_{1-y}Nb_y$/CaHAp (where $0 \leq y \leq 1$) catalysts are similar to that report on Fig. 1. When the ratio Nb/Ni is below 0.17, only diffraction lines corresponding to NiO and $Ca_{10}(PO_4)_6(OH)_2$ are detected. The increase of the Nb/Ni ratio leads to the decrease of the intensity of the peaks of NiO. No diffraction lines corresponding to niobium

compounds were observed possibly because of the low amount of Nb incorporated in the calcium-hydroxyapatite. It is also important to notice that to our knowledge no work on Nb loaded calcium-hydroxyapatite was published; albeit, several publications have reported the Ca^{2+} exchange with transition metal ions (40).

3.1.3 FTIR spectra

FTIR spectra of the catalysts show the typical bands of calcium-hydroxyapatite structure. The PO_4^{-3} groups are characterised by four vibrational modes: (i) the bands ν_2 (ii) and ν_3 at 1000–1200 cm^{-1} which both are assigned to asymmetric stretching (iii) ν_1 absorption at 950–960 cm^{-1} which is associated with symmetric stretching, and (iiii) the ν_4 at 560–570 cm^{-1} which is due to asymmetric bending vibrations (17-20). The wide absorption band appearing at 3400–3100 cm^{-1} is attributed to water present in the samples. The small and sharp band located at 3670–3570 cm^{-1} is due to the stretching vibration modes of the OH^- groups hosted by the tunnels of the apatite framework. In the domain 1500–1400 cm^{-1} and at 875 cm^{-1}, appear the asymmetric stretching and bending mode out of plane due respectively to the carbonates $(CO_3)^{2-}$ resulting from the atmospheric CO_2 adsorbed by the apatite (17-20).

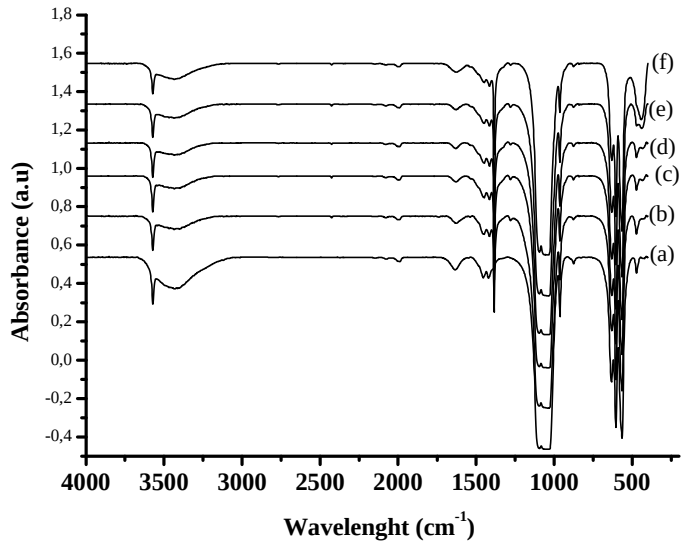


Fig. 2: FT-IR spectra of CaHAp and xNi/CaHAp for (a) x = 0.5; (b) x = 1.25; (c) x = 2.5; (d) x = 5; (e) x = 10; (f) x = 20.

Addition of nickel to CaHAp decreases the intensity of the band assigned to the hydroxyl groups (3580 cm^{-1}) located in the CaHAp channels but does not affect significantly the rest of the spectra (Fig. 2). This decrease is probably due to the incorporation of a small amount of Ni in the apatite network (20). A small band appears at 450 cm^{-1}. Its intensity increases with

the Ni loading. This band is due to the nickel oxide. Addition of different amounts of niobium to 10Ni/CaHAp does not modify significantly the FTIR spectra. The only observed change is the decrease of the intensity of the band at 450 cm^{-1} assigned to Ni oxide.

3.1.4 Diffuse reflectance spectroscopy

Diffuse reflectance spectroscopy was used to study the coordination and the symmetry of nickel species dispersed on CaHAp. The Ni^{2+} ions ($3d_8$) in octahedral surrounding exhibit a spectrum (^3F fundamental term and ^3P excited term) that involves three spin allowed d-d transitions appearing in the near infrared (NIR) and in the visible domain. These transitions are usually located around 930-1660 nm for the ν_1 ($^3A_{2g} \rightarrow {}^3T_{2g}$) transition, 570-1000 nm for ν_2 ($^3A_{2g} \rightarrow {}^3T_{1g}$) transition and 360-520 nm for ν_3 ($^3A_{2g} \rightarrow {}^3T_1g$ (P)) transition. The values of the crystal field parameters are generally small resulting in d-d transition bands displaying low intensities (17).

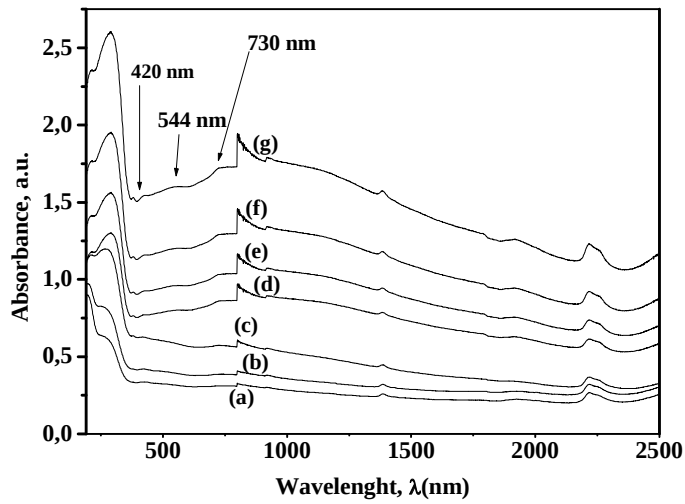


Fig. 3: DRS spectra of xNi/CaHAp for: (a) x = 0; (b) x = 0.5; (c) x = 1.25; (d) x = 2.5; (e) x = 5; (f) x = 10; (g) x = 20.

The spectrum of calcined CaHAp (Fig. 3a) displays: (i) in the NIR region several bands due to free and bonded OH hydroxyls. The bands at 1385 nm and 1425 nm are attributed to $2\nu_{(OH)}$ overtones and the bands at 1930 and 2220 nm to combinations of $\nu_{(OH)}$ and $\delta_{(OH)}$ (17-20); (ii) in the UV-visible domain the band at 210 nm was assigned to $O^{2-} \rightarrow Ca^{2+}$ charge transfer.

Figures 3b-3g show the spectra of Ni(x)/CaHAp. In NIR region they display numerous bands due to the vibration modes of OH groups identical to those observed with CaHAp. In the UV-visible region, the spectra exhibit beside the band at 210 nm due to $O^{2-} \rightarrow Ca^{2+}$ charge transfers a second one centred on 285 nm attributed to $O^{2-} \rightarrow Ni^{2+}$ charge transfers. The

intensity of this band increases with the increase of Ni concentration in the samples. In the visible domain, the bands located at 420 and 730 nm are due to the ν_3 and ν_2 transitions of Ni^{2+} ions in an octahedral or pseudo-octahedral symmetry (17). The band appearing at 544 nm might be associated with the existence of a second type of sites hosting Ni^{2+} ions in symmetry lower to octahedral. The broad band appearing in the NIR (950 nm) is attributed to ν_1 transition. The increase of Ni concentration enhances the broadening of the band indicating the formation of NiO on the surface of the phosphate.

The addition of niobium Ni(x)/CaHAp does not modify notably the spectra.

3.1.5 Temperature Programmed Reduction

The TPR profiles of the Ni(x)/CaHAp catalysts are shown in Fig. 4. They display only one sharp and symmetric peak which was assigned to NiO reduction (Table II). No shift of the maximum of the peak was observed when the Ni loading is increased. However, its intensity increases progressively with the increase of nickel loading and the growth of NiO particles. Previously, Boukha *et al*. have reported that the thermogram of Ni(x)/CaHAp exhibits above 550 °C two peaks attributed to the reduction of the Ni^{2+} ions exchanged with the Ca^{2+} ions in the apatite matrix (17). The difference between the TPR results published by Boukha and those presented here is due to used method for the preparation of the catalyst which minimises the incorporation of Ni^{2+} ions in CaHAp framework.

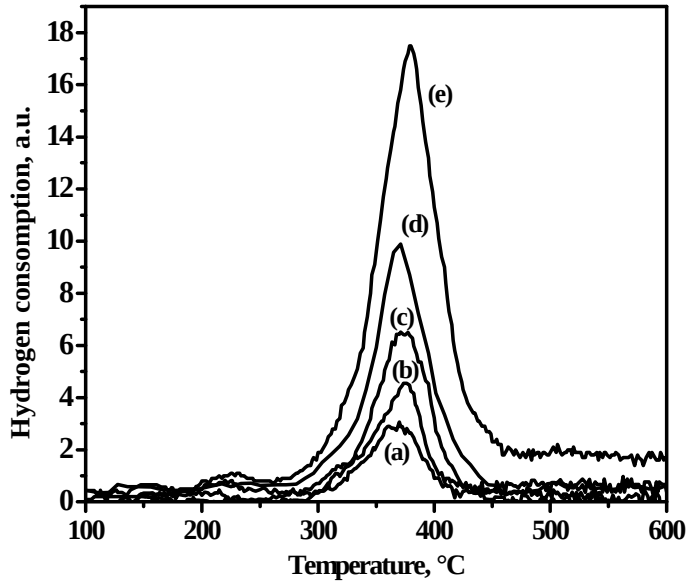


Fig. 4: Temperature Programmed Reduction profiles (H_2-TPR) of xNi/CaHAp where x is: (a) x = 1.25; (b) x = 2.5; (c) x = 5; (d) x = 10; (e) x = 20.

Fig. 5 displays the TPR profiles of $10Ni_{1-y}Nb_y$/CaHAp and shows that the incorporation of Nb into 10Ni/CaHAp induces significant modifications of the thermogram. For $y < 0.2$, it displays only one H_2 uptake (Fig. 5a-d) which shifts slightly towards high temperatures as Nb loading is increased (Table II). Normally, the incorporation of niobium in NiO should lower the reduction temperature of the mixed oxide because Nb^{5+} ions weaken the Ni-O-Ni bonds. The surprising increase the reduction temperature of $10Ni_{1-y}Nb_y$/CaHAp especially at high loadings can be ascribed to the increase of the interactions of the oxide with the support brought by Nb^{5+} ions. For $y \geq 0.2$, the TPR profiles become broad and exhibit two hydrogen consumptions appearing at 403 and 481°C respectively. The second hydrogen uptake can probably be assigned to the reduction of Nb^{5+} ions occupying more energetic sites on the surface of the carrier.

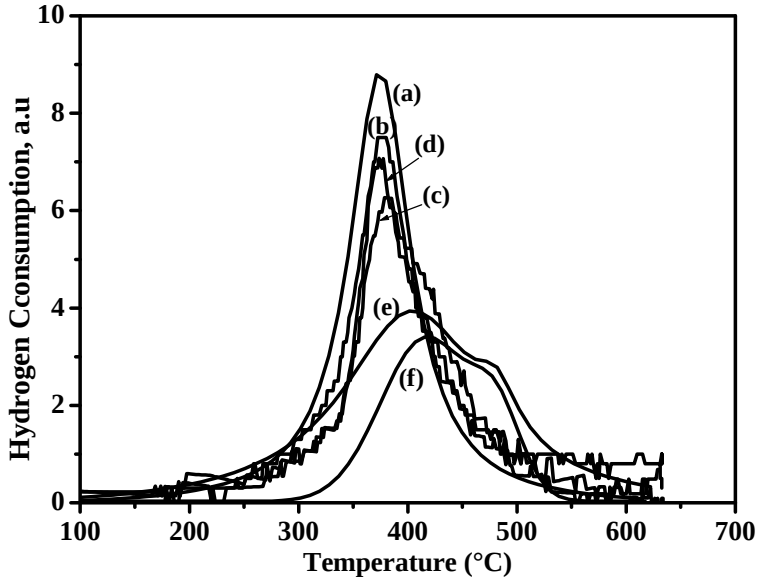


Fig. 5: Temperature Programmed Reduction profiles of $10Ni_{1-y}Nb_y$/CaHAp with: (a) y = 0.05; (b) y = 0.1; (c) y = 0.15; (d) y = 0.2; (e) y = 0.3; (f) y = 0.6.

The comparison of the calculated and the experimental amounts of hydrogen consumed (H_{2Exp}/H_{2Cal}, Table II) suggests that niobium does not uptake any hydrogen until $y \geq 0.2$ indicating that for $y < 0.2$ loadings there is no segregation between the two oxides (34-36). However, the problem needs more precise investigations.

Table II

Hydrogen consumption of the catalysts (H_2-TPR).

Catalysts	Temperature (°C)	Experimental H_2 uptake (Mole g^{-1})	calculated H_2 uptake (Mole g^{-1})	Ratio H_{2Exp}/H_{2Cal}
1.25Ni/CaHAp	362	1.910^{-4}	$2.1.10^{-4}$	0.9
2.5Ni/CaHAp	370	$4.4.10^{-4}$	4.210^{-4}	1.0
5Ni/CaHAp	372	$7.5.10^{-4}$	8.410^{-4}	0.9
10Ni/CaHAp	371	1.710^{-3}	1.610^{-3}	1.0
20Ni/CaHAp	379	3.310^{-3}	$3.4\ 10^{-3}$	1.0
10Ni0.95Nb0.05/CaHAp	375	$2.2.10^{-3}$	$1.9\ 10^{-3}$	1.1
10Ni0.90Nb0.1/CaHAp	381	$2.4.10^{-3}$	$2.2\ 10^{-3}$	1.1
10Ni0.85Nb0.15/CaHAp	383	2.010^{-3}	$2.5\ 10^{-3}$	1.1
10Ni0.80Nb0.20/CaHAp	391	2.310^{-3}	$2.7\ 10^{-3}$	0.9
10Ni0.70Nb0.30/CaHAp	403 481	$2.2.10^{-3}$ $0.1.10^{-4}$	$3.5\ 10^{-3}$	0.6
10Ni0.4Nb0.6/CaHAp	410 468	$2.2.10^{-3}$ $0.4.10^{-4}$	8.010^{-3}	0.3

4. Catalytic activity

Prior to the tests, the catalysts were stabilized by a treatment under synthetic air at 450°C. The reaction temperature was set at 425°C, because above 425°C, it leads exclusively to the production of methane and CO_x accompanied by a strongly reduction of the catalyst. Figure 6 displays the propane conversion versus time on stream (TOS) over the Ni(x)/CaHAp catalysts.

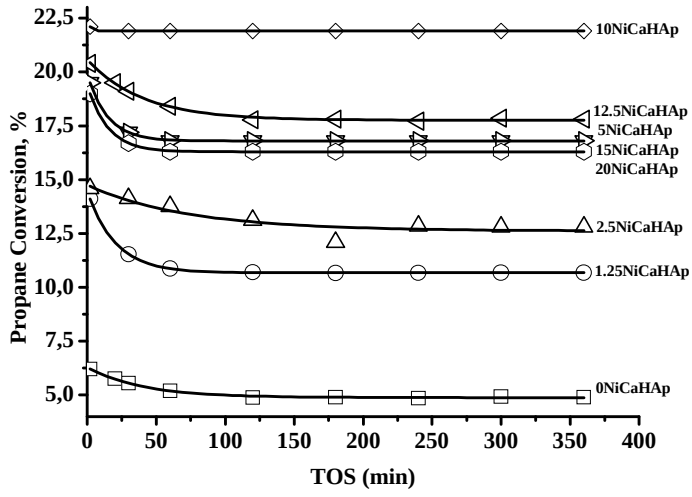


Fig. 6: Propane conversion versus time on stream at 425 °C over xNi/CaHAp catalysts.

On pure hydroxyapatite (free of Ni) the propane conversion riches about 5%. A similar conversion (5.9%) has recently been reported by S. Sugiyama *et al*. (41). Addition of Ni to

CaHAp increases the stationary propane conversion up to 22% for a nickel loading x = 10%. Above this concentration, the conversion decreases to around 16% for x = 20 wt.% Ni.

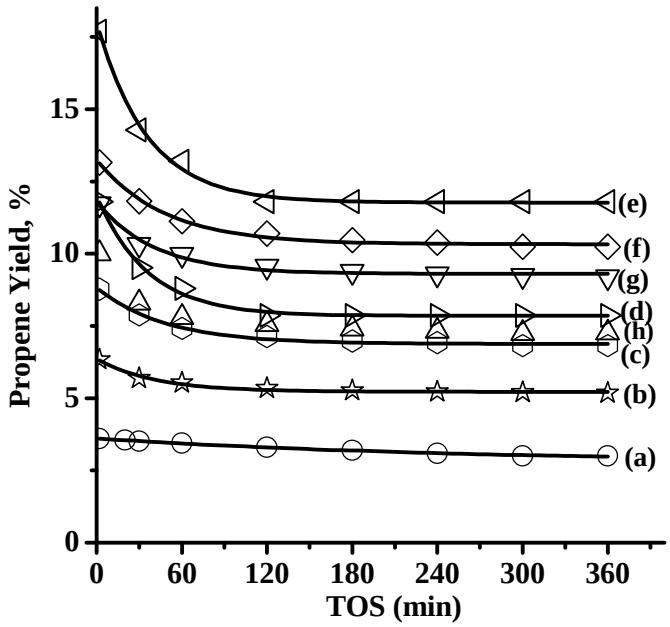


Fig. 7: Propylene yield versus time on stream at 425°C on xNi/CaHAp catalysts where: (a) x = 0; (b) x = 20; (c) x = 15; (d) x = 12; (e) x = 10; (f) x = 5; (g) x = 2.5; (h) x = 1.25.

Fig. 7 depicts the propylene yield versus time on stream over xNi/CaHAp catalysts. All the samples showed before they reach a stationary state a declining propylene production. It is important to signal that the decrease of the propylene yield is accompanied by an increase of methane production. The results presented in Fig. 6 and 7 indicate that the catalytic activity of CaHAp is strongly enhanced by nickel oxide and influenced by NiO particles size. The optimal conversion is achieved with 10Ni/CaHAp.

Fig. 8 summarized the variations of propane conversion, propene yield and propene selectivity, at steady state versus nickel content. The propane conversion and propylene yield increase with the nickel load and reach a maximum for about 10 wt.% Ni before declining (Fig. 8a and 8b). The propylene selectivity (Fig. 8c) displays similar variations but the maximum appears around 2.5 wt.% Ni and is equal to 77%. The observed maxima in the propane conversion and propylene yield are attributed to the modification of the nickel particles size and to the changes of the support basicity. As a matter of fact, the incorporation of nickel in the hydroxyapatite enhances its catalytic activity but concomitantly decreases its basicity, as previously shown when loading it with transition metals (18,20). This competition between the two parameters leads the observed maxima. Moreover, the increase of Ni content above 10 wt% increases the agglomeration the Ni particles and subsequently decreases the

amount of active sites. It is also commonly admitted that desorption of alkenes before their transformation into carbon oxides in ODH reaction is easier when the used catalysts possess basic sites. Corma *et al.* showed that the best conversion and selectivity of supported vanadium oxide in propane ODH are achieved with basic metal oxides as carriers (MgO, Bi_2O_3, La_2O_3 and Sm_2O_3) (42).

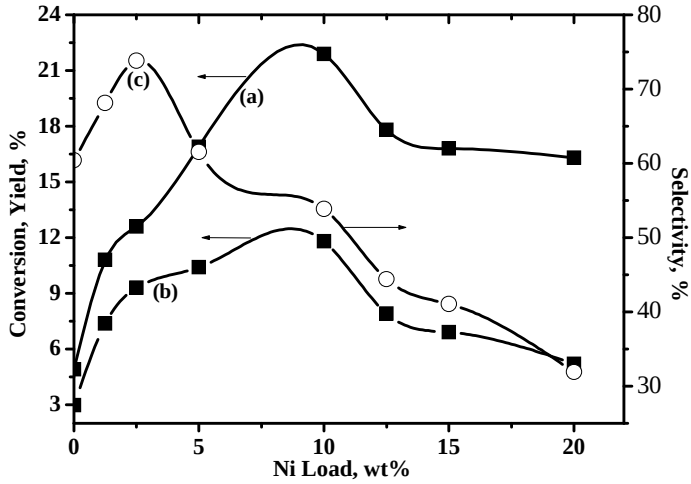


Fig. 8: Propane conversion (a), propylene yield (b) and selectivity (c), versus nickel load over xNi/CaHAp at 425°C.

For Ni loadings below 2.5 wt.% it is not excluded, despite the taken precautions during the preparation of the catalysts, that a small fraction of nickel ions exchanges with calcium in the hydroxyapatite structure. At loading higher than 5 wt.%, the surface of the support is covered by islands of NiO particles in agreement with DRX results. It is important to underline that in comparable conditions ethylene selectivity in ethane ODH is higher than that of propylene. This difference is essentially attributed to the fact that propane is significantly more active at lower temperatures than ethane (43).

A good method to improve the yield of propylene is to use additives. The metal oxide modifiers adjust the active site characteristics and modify the reaction kinetics. Heracleous *et al.* have investigated the effect of some additives on Ni/Al_2O_3 ethane ODH and concluded that the introduction of niobium was the most beneficial for ethane ODH (35). To investigate the influence of the Nb as a promoter of nickel loaded hydroxyapatite, 10Ni/CaHAp was selected because of its high propane conversion.

The propane ODH over the 10Ni/CaHAp was significantly influenced by the presence of niobium oxide. The variations of propane conversion and propylene yield at 425°C versus

time on stream are displayed on Fig. 9 and 10. Nb containing samples are less active than the 10NiCaHAp but the presence of Nb prevents the decrease of the catalysts activity versus time on stream (Fig. 10). The addition of Nb also increases slightly the propylene yield. The best results are obtained with 0.15 wt.% Nb.

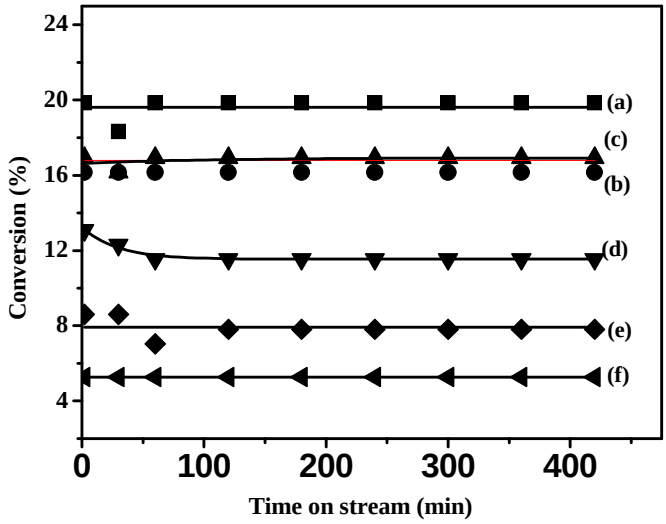


Fig. 9: Propane conversion versus time on stream at 425°C over $10Ni_{1-y}Nb_y$/CaHAp for (a) y = 0.05; (b) y = 0.1; (c) y = 0.15; (d) y = 0.2; (e) y = 0.3; (f) y = 0.4.

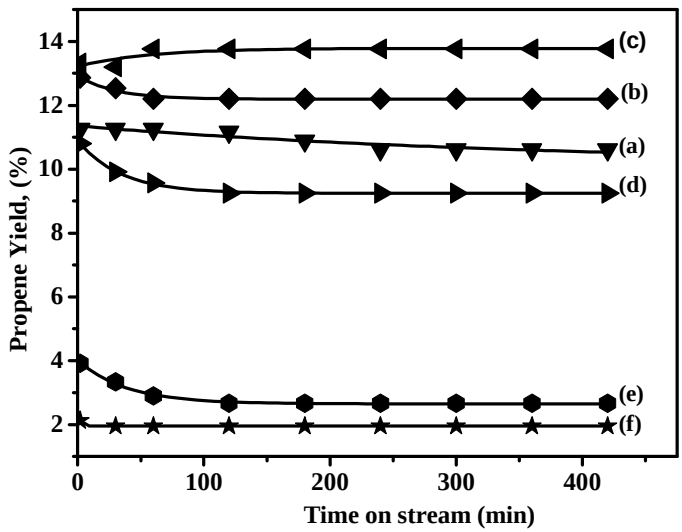


Fig. 10: Propylene yield versus time on stream at 425°C over $10Ni_{1-y}Nb_y$/CaHAp for (a) y = 0.05; (b) y = 0.1; (c) y = 0.15; (d) y = 0.2; (e) y = 0.3; (f) y = 0.4.

Fig. 11 and 12 display the variations of propane conversion and propylene yield at stationary state versus the reaction temperature. The propylene production increases with the

reaction temperature. The best performance was recorded for y = 0.15 with a selectivity above 90% at 425°C.

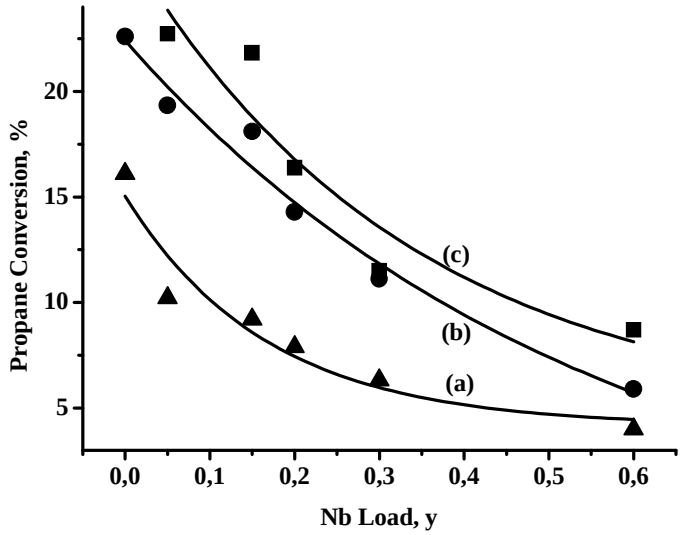


Fig. 11: Conversion versus Nb load at different reaction temperatures over $10Ni_{1-y}Nb_y$/CaHAp: (a) 400°C; (b) 450°C; (c) 500°C.

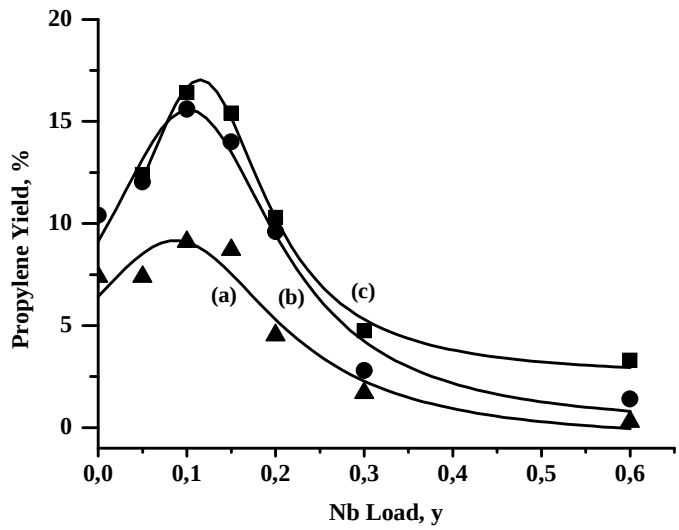


Fig. 12: Propylene yield versus Nb load at different reaction temperatures over $10Ni_{1-y}Nb_y$/CaHAp: (a) 400°C; (b) 450°C; (c) 500°C.

It is well known that, in light alkanes ODH, the selectivity increases when the conversion decreases. In the present case, introduction of niobium improved the selectivity by enhancing the propylene yield as well as the stability of the catalysts. These results suggest that the effect of niobium can be related to the modification of the strength of the oxygen bonded to the

active sites. As a matter of fact, even if the reaction mechanism is not exactly of Mars and van Krevelen type, the catalyst must able to activate oxygen in order to allow propane ODH.

Xinjie Zhang *et al*. in investigations carried out on ethane ODH over NiO/Al_2O_3 showed the presence on the catalyst of two different kinds of active oxygen species (44). The more active species transform ethane and ethylene into carbon dioxide while the less active ones convert ethane only into ethylene. A similar effect might be produced by the addition of niobium to 10Ni/CaHAp. Its incorporation diminishes the strong oxidizing abilities of oxygen species. E. Heracleous and A. A. Lemonidou also concluded that the introduction of Nb in NiO lattice by either substitution of nickel atoms and/or filling of the cationic vacancies in the defective nonstoichiometric NiO surface leads to a reduction of the electrophilic oxygen species (O^-), which are abundant on NiO and are responsible for the total oxidation of ethane to carbon dioxide (36). In the case of propane, the trend is not clear, but the high valence metals may influence the acid-base features of the surface. They probably also modify the oxygen stoichiometry of NiO (36-37). Further evidence of the impact of niobium on the selectivity was observed when several Nb/Ni ratios were synthesised by changing only the Nb content in the catalysts. However, additional work is necessary in order to understand the enhancing effect of Nb on the selectivity.

5. Conclusion

From the presented results the following conclusions might be drawn:

- In the Ni loaded CaHAp prepared by impregnation at pH = 9 and calcined in air, nickel is present as (i) a small quantity of isolated Ni^{2+} ions in octahedral symmetry probably exchanged with Ca^{2+} ions on the surface, (ii) amorphous NiO and large crystallised NiO particles characterized by different techniques.
- Nb promoted $10Ni_{1-y}Nb_y$/CaHAp catalysts were synthesized using a new method. TPR (H_2-TPR) profiles showed surprisingly an increase of the reduction temperature of the promoted 10Ni/CaHAp, especially at high Nb loadings ascribed to the increase of the interactions of the oxide with the support brought by Nb^{5+} ions.
- The two series of catalysts were tested propane ODH. The Propylene selectivity and the catalyst stability were improved by Nb addition.

6. Acknowledgements

The authors wish to acknowledge the financial support provided by CNRST. The help kindly offered in materials analyses and characterization by XRD is greatly appreciated.

7. References

(1) L. F. Albright, B. L. Crynes and W. H. Corcoran, (eds), Pyrolysis: Theory and Industrial Practice (Academic Press, London, 1983).

(2) G. Centi, F. Cavani and F. Trifirò, Selective Oxidation by Heterogeneous Catalysis, Kluwer Academic/Plenum Publishers (2001).

(3) Bo Zheng, W. Hua, Y. Yue , Zi Gao, J. Catal. 232 (2005) 143.

(4) R.K. Grasselli, Catalysis Today 49 (1999) 141.

(5) F. Cavani , N. Ballarini, A. Cericola, Catalysis Today 127 (2007) 113.

(6) I.M. Dahl, K. Grande, K.J. Jens, E. Rytter and A. Slagtern, Appl. Catal. A: 77 (1991)163.

(7) Z. Chao and E. Ruckenstein, Catalysis Letters Vol. 94, Nos. 3–4, May 2004

(8) G. Karamullaoglu, S. Onen, T.r Dogu, Chemical Engineering and Processing 41 (2002) 337.

(9) D. S. Sam, V. Soenen and J. V. Volta, J. Catal. 123 (1990) 417.

(10) J. N. Michaels, D. L. Stern and R. K. Grasseli, Catal. Lett. 42 (1996) 135.

(11) N. Mizuno, M. Tateishi and M. Iwamoto, Appl. Catal. A. 128 (1995) 165.

(12) G. Centi and F. Trifirò, Appl. Catal. A. 143 (1996) 3.

(13) T. Balsco, P. Concepcion, J. M. Lopez-Nieto and J. Perez-Pariente, J. Catal. 152 (1995) 1.

(14) M. Okamoto, L. Luo, J. A. Labinger and M. E. Davis, J. Catal. 192 (2000) 128.

(15) A. A. Teixeira-Neto, L. Marchese, G. Landi, L. Lisi, H. O. Pastore, Catalysis Today 133 (2008) 1.

(16) A. M. Gaffney, S. Chaturvedi, M. B. Clark Jr., S. Hana, D. Lea, S. A. Rykovb, J. G. Chenb, J. Catal. 229 (2005) 12.

(17) Z. Boukha, M. Kacimi, M. F. R. Pereira, J. L. Faria, J. L. Figueiredo and M. Ziyad Appl. Catal. A: General 317 (2007) 299.

(18) N. Cheikhi, M. Kacimi, M. Rouimi, M. Ziyad, L. F. Liotta, G. Pantaleo, G. Deganello J. Catal. 232 (2005) 257.

(19) K. Elkabouss, M. Kacimi, M. Ziyad, S. Ammar and F. Bozon-Verduraz, J. Catal. 226 (2004) 16.

(20) M. Khachani, M. Kacimi, A. Ensuque, J-Y Piquemal, C. Connan, F. Bozon-Verduraz, M. Ziyad, Appl. Catal. A: General 388 (2010) 113.

(21) L. Savary, J. Saussey, G. Costentin, M.M Bettahar, M. Gubelmann- Bonneau, J.C. Lavalley, Catal. Today 32 (1996**)** 57.

(22) X. Zhang, H. Wan, W. Weng, X. Yi, Appl. Surf. Sci. 220 (2003) 117.

(23) A. Jimenez-Lopez, E. Rodriguez-Castellon, J. Santamaria-Gonzalez, P. Braos-Garcia, E. Felici, F. Marmottini, Langmuir 16 (2000**)** 3317.

(24) E. A. Mamedov , V. Cortds Corberfin, Appl. Catal. A: General 127 (1995) 1.

(25) A. Addane, M. Kacimi and M. Ziyad, Catalysis Letters 73 (2001) 47.

(26) G. I. Golodets, Stud. Surf. Sci. Catal. 55 (1990) 693.

(27) A. Pantazidis, S.A. Bucholz. H.W. Zanthoff, Y. Schuurman, C.A. Mirodatos, Catal. Today, 40 (1998) 207.

(28) J. A. Aderson, L. Daza, D. S. Damyanova, J. L. G. Fierro, M. T. Rodrigo, Appl. Catal. A 113 (1994) 75.

(29) K. Schulze, W. Makowski, R. Chyz, R. Dziembaj, G.Geismar, Applied Clay Science 18 (2001) 59.

(30) E. Ruckenstein, Y. H. Hu, J. Catal. 162 (1996) 230.

(31) Y. Schuurman, V. Ducarme, T. Chen, W. Li, C. Mirodatos, G. A. Martin, Appl. Catal. A 163 (1997) 227.

(32) X. Zhang, Y. Gong, G. Yu, Y. Xie, J. Mol. Catal. A 180 (2002) 293.

(33) X. Zhang, J. Liu, Y. Jing, Y. Xie, Appl. Catal. A: 240 (2003) 143.

(34) E. Heracleous, A. F. Lee, K.Wilson, A.A. Lemonidou, J. Catal. 231 (2005)159.

(35) E. Heracleous, A. A. Lemonidou, J. Catal. 237 (2006) 162.

(36) E. Heracleous, A. A. Lemonidou, J. Catal. 237 (2006) 175.

(37) I. Nowak and M. Ziolek, Chem. Rev. 99 (12) (1999) 3603.

(38) K. Mori, T. Hara, T. Mizugaki, K. Ebitani, and K. Kaneda, J. Am. Chem. Soc. 126 (2004) 10657.

(39] C. C. Silva, A. G. Pinheiro, M.A.R. Miranda, J. C. Go´ es, A. S. B. Sombra, Solid State Sciences 5 (2003) 553.

(40) I. Mobasherpour, M. Soulati Heshajin , A. Kazemzadeha, M. Zakeri, Journal of Alloys and Compounds 430 (2007) 330.

(41) S. Sugiyama, T. Osaka , Y. Hirata , K. Sotowa, Appl. Catal. A: General 312 (2006) 52.

(42) A. Corma, J. M. Lopez-Nieto, G. Kremenic, and J. L. Fierro, Appl. Catal. A. General 61 (1990) 235.

(43) D. R. Lide, CRC Handbook of Chemistry and Physics, 83rd ed.: CRC Press LLC. (2002).

(44) X. Zhang, Y. Gong, G. Yu, Y. Xie, J. Mol. Catalysis A: Chemical 180 (2002) 293.

An Integrated Approach for Spare Parts Provisioning

C. Diallo[1,7], **D. Aït-Kadi**[2,3,5,6] **and A. Chelbi**[3,4]

[1] Dalhousie University, Industrial Eng., 5269 Morris Street, Halifax, Nova-Scotia, Canada B3H-4R2
[2] Université Laval, Mechanical Eng., 1065, avenue de la Médecine, Québec, Canada G1V-0A6
[3] Centre Interuniversitaire de Recherche sur les Réseau d'Entreprise, la Logistique et le Transport CIRRELT, Université Laval, Canada G1V-0A6
[4] Centre de Recherche en Productique CEREP, Ecole supérieure des Sciences et Technique de Tunis ESSTT, BP 56 Bab Menara 1008, Tunis, Tunisia
[5] Centre Interdisciplinaire de Recherche en Réadaptation et Intégration Sociale CIRRIS Université Laval, Canada G1V-0A6
[6] Resident Member of the Académie Hassan II des Sciences et Techniques, Royaume du Maroc
[7] *Corresponding Author E-mail: claver.diallo@dal.ca*

Abstract. *This article addresses the problem of spare parts identification and provisioning for multi-component systems. A framework considering available technical, economical and strategic information is presented. Mathematical models are proposed to determine, for each spare part, the required quantity over a given planning horizon. The objective may be to maximize either the reliability or the availability of the system. Analytical models are proposed to determine the management parameters.*

Key words: maintenance, spare parts, identification, inventory, optimization.

1. Introduction

When acquiring a system, one is often faced with the difficult question of identifying which of its components may fail to operate on the horizon considered and for which spare parts stocks are to be supplied. It should be noted that, contrary to standard parts, spare parts are designed for specific use and they can be acquired only from the manufacturer of the system or its authorized representatives. Their provisioning leadtime is generally long, if not often unknown. They cannot be resold easily. Consumption is typically governed by a random process. These parts are, additionally, subject to obsolescence and deterioration. Because they can be used to perform preventive replacement

(planned replacement) or corrective replacement (random replacement), classical mathematical models of inventory management [25] cannot be applied directly. Another dimension to consider stems from the fact that some parts may be remanufactured, or simply repackaged to be used as spares. Others can only be used once. First, we will deal with the latter category of parts and then we will tackle the first category in a subsequent section. Inventory management of spare parts is a strategic and economical issue for all organizations that operate equipment with operating characteristics deteriorating with age and use. Interruption of service can have high economical consequences. The unavailability of a component is certainly very detrimental both in terms of cost and continuity of production or service [30]. On the other hand, maintaining a high stock of spare parts can be expensive.
Section 2 of this article presents several decision tools for the identification of parts to be kept in stock. Section 3 deals with the determination of the spare parts quantity required to achieve a predetermined performance for non-repairable systems. Section 4 proposes a mathematical model to find the quantity of repairable spare parts required to maintain a certain level of service. A conclusion of this work is presented in section 5.

2. Spare Parts Identification

In the absence of data and information to assess the degradation of a component and to estimate the probability of failure over the economic life of the system, we rely generally on the manufacturer's recommendations. Manufacturers have often provided a list of components to keep in stock based on their own feedback, on data from accelerated testing conducted according to standard procedures or from more sophisticated analysis of the failure modes of the main components. To stay in business, equipment manufacturers are expected not only to meet the needs identified by the client, but also to anticipate these needs and to demonstrate that the products and services offered are equivalent if not superior to what is available on the market. Nowadays, customers may claim information that could help them gain maximum advantage from the system being acquired. More and more customers require their suppliers to provide them with lifetime and degradation information, analysis of failure modes, effects and criticality (FMEA), etc. Having access to these data, it becomes possible to identify components for which spare parts are to be kept. Thus, if the lifetime density function $f_i(.)$ or the lifetime cumulative distribution function $F_i(.)$ or the failure rate $h_i(.)$ or the reliability function $R_i(.)$ of component i is known, then spare parts are to be held if

$$F_i(t) > F^*$$

where F^* is the maximum failure risk the buyer (customer) is willing to accept during the mission duration of the system being bought.

It is also possible to base the decision on the criticality index obtained from the FMEA/FMECA of the system or by referring to similar equipment or on recommendations from internal or external experts. Figure 1 presents a generic process for identifying components for which spare parts are expected. For each equipment under consideration, all available data on lifetimes, repair times, suppliers, lead-times, etc., are gathered. If partial or no information is available, then the

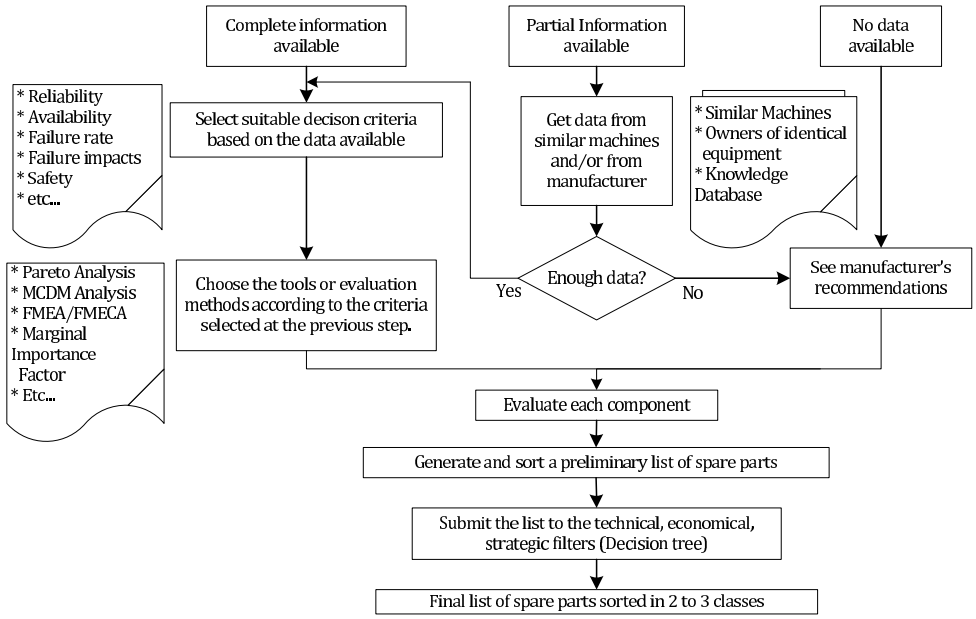


Figure 1: The spare parts identification process.

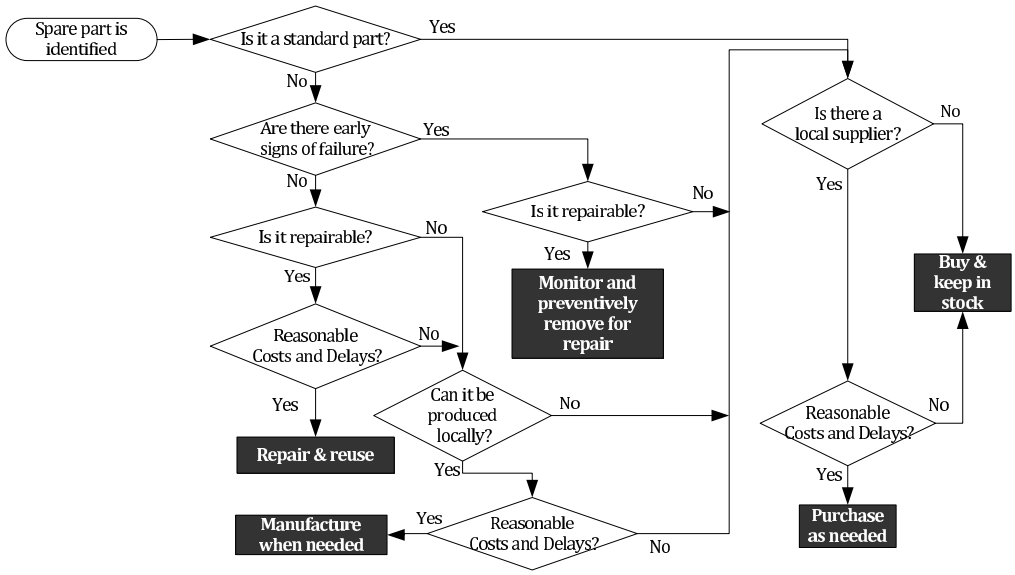


Figure 2: Decision tree with filters.

decisions are based on the manufacturers' recommendations or on information from owners of similar equipment. When complete information is available, then the decision-maker chooses the

criteria according to which the components will be evaluated to determine if they should be on the list of spare parts. A list of potential criteria is given in figure 1. According to the criteria retained, evaluation and final decision-making methods are to be selected next. Each component of the equipment is then evaluated according to the criteria and a total score is obtained. A list of potential methods is also provided in figure 1. A ranking of the components based on the final scores gives an ordered list of potential spare parts. Once the preliminary list of spare parts is established, it is subjected to filters to select the parts to hold in stock and those to be supplied as needed. An example of decision tree with filters is depicted in figure 2. This tree takes into account the cost of acquisition or production, repair costs, delays, whether there are early signs of failure or not, and if the component is a standard part or not. A standard part is a generic mass-produced part readily available at reasonable to low cost (e.g., seals, nuts and bolts, high replacement rate parts). Selected components are then ranked in order of importance. This classification allows to pay more attention to the components considered as being more important, especially if the list of spare parts includes a large number of components and the resources available are limited or scarce. Decision criteria most often used to justify that a spare part must be kept in stock to serve as a replacement are: criticality, reliability, availability, impacts of failure, failure rate and maintenance costs incurred in case of failure. In most cases, for simplicity or ignorance of the analysis tools available, only one or two of these criteria are considered in the analysis. For many organizations, the total cost of maintenance criterion is often the one used in the decision process. For each component i, we calculate the critical ratio RG_i by

$$RG_i = \frac{\text{Indirect Costs}}{\text{Direct Costs}}$$

For more details on the direct and indirect costs, please refer to reference [9]. Any component i with RG_i ratio greater than 1, is then kept in store as a spare.

The analysis of failure modes, their effects and criticality, is increasingly used in industry to ensure sustainable use of assets. Each component is associated with a criticality index (C) obtained by multiplying the severity index (S), the probability of occurrence of the adverse event (O) and the difficulty of detecting it (D) (see [12, 6, 36])

$$C = S \times O \times D$$

For a component subject to several failure modes, a criticality index C_k is defined for each failure mode k (see [16]):

$$C_k = K_A \cdot K_E \cdot \alpha_{kp} \cdot \beta_k \cdot \lambda_p \cdot T$$

where

A system or component whose criticality index exceeds a predetermined threshold will be included in the list of potential spare parts.

In practice, the decision to store a part or not, may involve several criteria (costs, reliability, frequency of failure, response time, etc.) In the event that more than three criteria must be considered, multicriteria decision-making methods (MCDM) or tools have to be used. Several studies

C_k : criticality index for failure mode k
K_A : failure rate adjustment factor to compensate for actual operating conditions
K_E : failure rate adjustment factor to compensate for actual environmental conditions
$\alpha_k p$: proportion of failures of component p due to failure mode k
β_k : conditional probability that failure mode k will cause the failure
λ_p : failure rate of component p
T : mission length

published in the literature have successfully used multicriteria methods for the identification and classification of equipment and spare parts (for example see [18] and [8]). The same tools can be adopted to generate a list of components to include in the priority list of potential replacement parts. Braglia et al. [7] used the AHP multicriteria classification method to classify the parts according to the impact of the failure, utilization, inventory problems and characteristics of parts. Schärlig [32] [31], Roy and Bouyssou [29] propose extensive reviews of multicriteria methods. Many computer programs and web sites offer the possibility of carrying out a multicriteria classification. Hammami [22] presents a very detailed review of computer programs dealing with multicriteria decision support. Eisenhawer et al. [17] propose a method of approximate reasoning based on fuzzy logic to establish a priority list of items to keep in stock for a nuclear facility. Once the components to be considered as spare (replacement) parts are identified, one has to determine the required quantities to be acquired during a given time period in order to achieve the expected performance levels. The following section will discuss models and methods for calculating the quantities of parts required during the economic life cycle of the system.

3. Determination of the required quantity of non-repairable spares

For each component requiring spare parts, it is important to estimate the required amount of spares needed throughout the economic life cycle of the equipment. To achieve this, one must estimate the average number of replacements at failure and, where applicable, the average number of preventive replacements. In this study, the data and information available will primarily be used to determine the density function $f(.)$, the distribution function $F(.)$, the survivor function or reliability $R(.)$, the failure rate $h(.)$ associated with the lifetime of the component under consideration. Figure 3 shows how field data is processed to obtain one of the four reliability characteristics. Knowledge of any one of these four characteristics is sufficient to obtain the other three. Table 1 recalls the relationships between these four functions.

For a component with lifetime density function $f(t)$ and negligible replacement duration, the average number of replacements at failure $M(t)$, with replacements carried-out with new spare

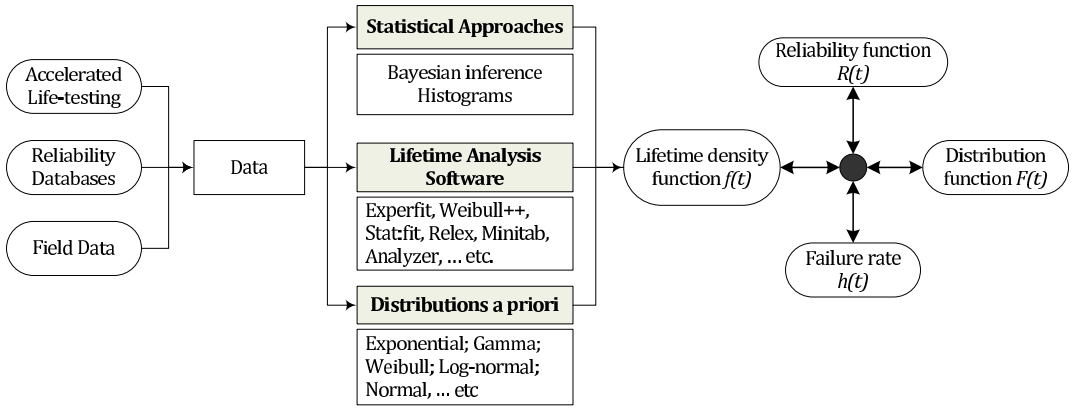


Figure 3: Field data processing diagram.

	$f(t)$	$F(t)$	$R(t)$	$h(t)$
$f(t)$	–	$\int_0^t f(x)dx$	$\int_t^\infty f(x)dx$	$\frac{f(t)}{\int_t^\infty f(x)dx}$
$F(t)$	$\frac{dF(t)}{dt}$	–	$1-F(t)$	$\frac{dF(t)}{[1-F(t)]dt}$
$R(t)$	$\frac{-dR(t)}{dt}$	$1-R(t)$	–	$\frac{-dR(t)}{R(t)dt}$
$h(t)$	$h(t).e^{-\int_0^t h(x)dx}$	$1-e^{-\int_0^t h(x)dx}$	$e^{-\int_0^t h(x)dx}$	–

Table 1: Basic reliability relationships.

parts during a mission of length t, satisfies the following fundamental renewal equation:

$$M(t) = F(t) + \int_0^t M(t-x)f(x)dx.$$

If $F^{(i)}(t)$ denotes the i–fold convolution of $F(t)$ with itself, then

$$M(t) = \sum_{i=1}^{\infty} F^{(i)}(t).$$

If at failure the component is minimally repaired without affecting its failure rate $h(t)$, then the average number of failure during the time interval [0,t] is given by

$$M(t) = \int_0^t h(x)dx.$$

When the repair or replacement durations are random, the average number of failures during the time interval [0,t] is given by

$$M(t) = \sum_{i=1}^{\infty} G^{(i)}(t)$$

where $G^{(i)}(t)$ denotes the i–fold convolution of $G(t)$ with itself, $g(t) = \frac{dG(t)}{dt}$ being the convolution of the lifetime density function $f(t)$ with the repair or replacement duration density function $v(t)$ such that

$$g(t) = \int_0^t f(t-x)v(x)dx.$$

Closed-form expressions for the renewal function $M(t)$ are only known to a relatively short list of distributions used in reliability and maintenance modeling, such as the Uniform, Exponential and Erlang distributions. However, several numerical methods have been proposed to compute $M(t)$ (see [2, 24, 10]). Diallo and Aït-Kadi [15] have proposed an approximation based on the Dirac function to compute $g(t)$ and $M(t)$ when repair or replacement durations are not negligible. Once $M(t)$ is known, the average number n of spare parts required for the time interval [0,t[is obtained by rounding its value to the next integer:

$$n = \lceil M(t) \rceil.$$

For a component replaced at failure and after T units of time, according to the age replacement policy (ARP), the upper bound of the expected number of spare parts n_a for a mission of length t is given by [5]:

$$n_a = \left\lceil \frac{t \cdot [1 - R(T)]}{\int_0^T R(x)dx} \right\rceil.$$

If the component is replaced at failure or at predetermined instants kT ($k = 1, 2, 3, \cdots$) regardless of its age and state, according to the block replacement policy (BRP), then the expected number of spare parts n_b for a mission of length t is given by [5]:

$$n_b = \lceil k\,[M(T) + 1] + M(t - kT) \rceil$$

with $kT \leq t < (k+1)T$.

For some applications, a spare part can be considered as a stand-by component in the reliability point of view. The determination of the system reliability $R_S(t, n)$ for a stand-by structure with n components allows to calculate the number of spares n to keep in stock to achieve a desired predetermined reliability level R^* for a given mission duration t . This is equivalent to finding the smallest integer n^* such that:

$$\begin{aligned} R_S\,[t, n^*(t)] &\geq R^* \\ \int_t^{\infty} f^{[n^*(t)+1]}(x)dx &\geq R^* \end{aligned}$$

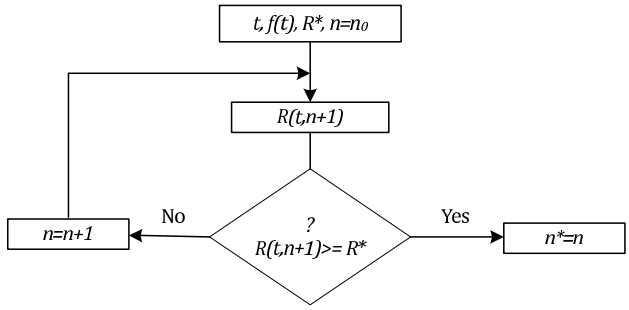


Figure 4: Computation procedure for determining the number of spare parts.

where $f^{(i)}(t)$ denotes the i-fold convolution of $f(t)$ with itself. Once $R_S(t,n)$ is known the number of spare parts to keep can be computed using a simple iterative algorithm from Aït-Kadi et al. [1] and depicted by figure 4.

Table 2 in Diallo et al. [13] gives the expressions of $R_S\left[t, n^*(t)\right]$ for different configurations of the problem. In the general case, when k components are in operation and $(n-k)$ are kept in stock, the expression of $R_S\left[t, n^*(t)\right]$ becomes:

$$R_S\left[t, n^*(t)\right] = R_S\left[t, n^*(t), k\right]$$

where

$$R_S\left[t, n^*(t), k\right] = e^{-k\lambda t} \sum_{j=0}^{n-k} \frac{(k\lambda t)^j}{j!}.$$

Note that the number of spare components can be calculated using a predetermined availability value A^*. The approach is to find $n^*(t)$ such that $A[t, n*(t)] \geq A^*$, where $A[t, n*(t)]$ is the availability of a stand-by structure consisting of $n^*(t)$ reserve components and one component in operation. Diallo et al. [14] proposed a mathematical model for the maximization of the system's availability under joint preventive maintenance and spare parts provisioning strategy.

4. Determination of the required quantity of repairable spares

A failed component, according to its degradation state, is repaired and put back in a "as new condition" or put in a state where it can resume operation. The acquisition cost of these repairable components is generally high. If the repair is found to be not feasible for technical, economical or other reasons, the failed component is sent to a recycling or disposal facility. The method based on the stand-by structure presented in the previous section can also be used here to determine the number of repairable components to use in a single system.

For a repairable component (as good as new) with failure rate $h(t) = \lambda$ and repair rate $\mu(t) = \mu$, the expression of $R_S\left[t, n^*(t)\right]$ is given by

$$R_S\left[t, n^*(t)\right] = e^{-zt}$$

where

$$z = \frac{(1-\gamma)^2 \gamma^{n^*(t)} \lambda}{1 - \gamma^{n^*(t)+1}\{1 + [n^*(t)+1](1-\gamma)\}}$$

and $\gamma = \lambda/\mu$.

Inventory management of a fleet of repairable systems is much more complex than that of standard nonrepairable components. This problem occurs mainly in mining facilities, fleet maintenance, oil industry, civil aviation, military nuclear industry, etc. The complexity of the problem is usually due to the randomness of the failures, the varying restoration times, the uncertainties about the state of degradation of the component. Several analytical models, dealing with some variants of the problem of inventory management for repairable components, were published in the literature. Figure 5 presents, schematically, a configuration with two bases (operations centers) and a central depot for the maintenance of components and systems. The objective is to determine systems and components stock levels, at the bases and at the depot, in order to guarantee a given level of service at minimum cost. Since the repairs take time and given that any non-availability of the systems results in high costs, an inventory of systems in operational state is maintained at each base. Once a failure occurs at a base, a replacement system is taken from the stock and is immediately put in place to ensure continuity of service. The failed system is either repaired on site (at the base) with a given probability or sent to the repair center at the depot with a certain shipment leadtime. According to the level of degradation of the components, the repair can be undertaken or not. If the repair is not possible, a new one is ordered. When the repair is possible, spare components are used to carry-out the replacement. If replacement components are not available, an order is immediately placed. After repair, the system is sent back to the originating base (decentralized management) or kept in stock at the depot (centralized management). The repair time depends on the availability of spare parts, the repair capacities and the workload in the repair shops.

For the problem illustrated above, consider that the fleet has N machines (systems) in operation. These systems are independent and identically distributed each with a failure rate $\lambda(t)$ and a repair rate $\mu(t)$. A stock of y spare systems is held. A repair shop, consisting of c parallel channels, is used to repair systems that fail during operation. The production of goods or services is stopped whenever a total of $(N + y)$ systems have failed. Taylor and Jackson [35] were the first to apply queueing theory to solve this spare parts provisioning problem. Several studies have subsequently been devoted to the subject. These include, among others, the work done by Sherbrooke [33, 34] and others such as [4, 23, 27, 28, 3, 11, 26]. Two cases are distinguished. In the first case, the capacity of repair stations are assumed sufficient and therefore no waiting line is formed (see Sherbrooke [33, 34]). The modeling assumptions adopted lead to lower stocks than would be required to achieve the specified level of service [3]. In the second case, finite repair capacity is assumed. The analytical treatment becomes more complex as shown in [27, 3, 19] . A comprehensive review of articles on inventory management of repairable systems is carried out by Guide and Srivastava [21]. We will now present the model proposed by Gross et al. [19] to address the problem illustrated in figure 5. This is a Markovian model for a single echelon repairman problem whose transition diagram is shown in figure 6.

The number y of spare parts to keep in stock should allow to reach the service level NS defined

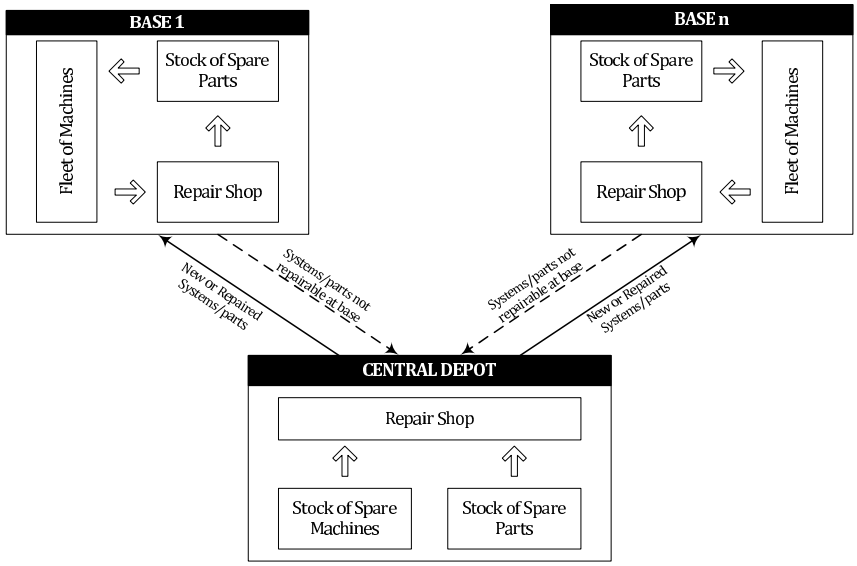


Figure 5: Flow of parts in the two-echelon repairman problem.

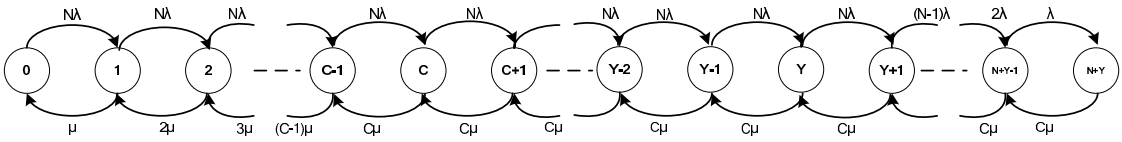


Figure 6: Transition diagram of the repairman model.

as the probability of having at least one spare machine in stock.

P_i is the steady-state probability of having i failed machines awaiting repair or being repaired. When the failure and repair rates are constant such that $\lambda(t) = \lambda$ and $\mu(t) = \mu$, then the P_i are given by Gross et al. [19]:

$$\text{for } c > y \qquad P_i = \begin{cases} \dfrac{N^i}{i!}\left(\dfrac{\lambda}{\mu}\right)^i P_0 & \text{for } i = 1, \cdots, y \\ \dfrac{N^y N!}{(N-i+y)!i!}\left(\dfrac{\lambda}{\mu}\right)^i P_0 & \text{for } i = y+1, \cdots, c-1 \\ \dfrac{N^y N!}{(N-i+y)!c^{i-c}c!}\left(\dfrac{\lambda}{\mu}\right)^i P_0 & \text{for } i = c, \cdots, y+N \end{cases}$$

$$\text{for } c \leq y \qquad P_i = \begin{cases} \dfrac{N^i}{i!}\left(\dfrac{\lambda}{\mu}\right)^i P_0 & \text{for } i = 1, \cdots, c-1 \\ \dfrac{N^i}{c^{i-c}c!}\left(\dfrac{\lambda}{\mu}\right)^i P_0 & \text{for } i = c, \cdots, y-1 \\ \dfrac{N^y N!}{(N-i+y)!c^{i-c}c!}\left(\dfrac{\lambda}{\mu}\right)^i P_0 & \text{for } i = y, \cdots, y+N \end{cases}$$

The expression of P_0 is given by

$$P_0 = \begin{cases} \left[1 + \sum_{i=1}^{c-1} \frac{N^i}{i!}\left(\frac{\lambda}{\mu}\right)^i + \sum_{i=c}^{y} \frac{N^i}{c^{i-c}c!}\left(\frac{\lambda}{\mu}\right)^i + \sum_{i=y+1}^{y+N} \frac{N^y N!}{(N-i+y)!c^{i-c}c!}\left(\frac{\lambda}{\mu}\right)^i\right]^{-1} & \text{for } c \leq y \\ \left[1 + \sum_{i=1}^{y-1} \frac{N^i}{i!}\left(\frac{\lambda}{\mu}\right)^i + \sum_{i=y}^{c} \frac{N^i N!}{(N-i+y)!c!}\left(\frac{\lambda}{\mu}\right)^i + \sum_{i=y+1}^{y+N} \frac{N^y N!}{(N-i+y)!c^{i-c}c!}\left(\frac{\lambda}{\mu}\right)^i\right]^{-1} & \text{for } c > y \end{cases}$$

The source population (the fleet of machines+spares) is usually finite ($N + y$), therefore the probability that a failure is about to occur when there are i broken machines in the repair shop is given by Q_i (see Gross et al. [19]):

$$Q_i = \begin{cases} \dfrac{N \cdot P_i}{N - \sum_{i=y}^{y+N}(i-y)P_i} & \text{for } i = 0, \cdots, y-1 \\ \dfrac{(N-i+y) \cdot P_i}{N - \sum_{i=y}^{y+N}(i-y)P_i} & \text{for } i = y, \cdots, y+N \end{cases}$$

Once the Q_i are known, it suffices to find the smallest integer y such that

$$\sum_{i=0}^{y-1} Q_i \geq NS$$

Logistical delays are not taken into account by this model. Moreover, the assumption that failure and repair rates are constant, sometimes causes decision-makers to question the validity of the results obtained by this model. Gross [20] studied the sensitivity of the model to the exponentiality assumption and derived easy rules of thumb to estimate the error induced by such a hypothesis. Kim et al. [27] present an algorithm to determine the level of the stock of spare parts required in the different bases in a two-tier system with a central depot. Note that in their model, the depot does not have a stock of spare parts. It only carries out repairs. The authors propose a mathematical model to minimize the average total cost of inventory management while satisfying a minimum service level. They also use the results of queueing theory to determine the probability of shortages and the likelihood of excess inventory. Logistical delays are included in their model.

5. Conclusion

An integrated approach for the identification and management of spare parts has been proposed. We have described a methodology for the identification of components for which replacement parts must be kept. For each spare part, analytical models are presented for determining the quantities required over a given operating horizon. Models of inventory management were then proposed for repairable systems. Several factors that affect system performance such that the replenishment leadtime and random demand are taken into account in the mathematical models presented.
Because the machines and their operating environment tend to change over time, it is wise to frequently update the management parameters and decision variables to take into account any technological, economical and strategic change. It should also be worthwhile to implement proper maintenance procedures and monitoring of spare parts when they are stored for long periods of time by an appropriate control of their environment (moisture control, greasing, repositioning).
Future research work should focus on the determination of inventory levels at bases and depot levels when risk-pooling initiatives are possible and lateral transshipments allowed. Setting of inventory levels of repairable spare parts also called "rotables" in some industries will be studied when scheduled inspections create random and deterministic demand on top of the demand generated by random failures.

Acknowledgements

The authors would like to thank the anonymous referees and the editors for their useful suggestions.

References

[1] Aït-Kadi D, Diallo C, Chelbi A. *Spare parts identification and provisioning models*, Journal of Decision Systems; 12 (2003), No. 1, 47-65.

[2] Aït-Kadi D, Chelbi A. Procédure numérique de calcul de la fonction de renouvellement, Technical Report, Université Laval, 1998.

[3] Albright SC. *Approximation to the stationary distribution of a multi-echelon repairable-item inventory system with finite sources and repair channels*, Naval Research Logistics; 36 (1989), No. 2, 179-195.

[4] Balania AR, Gross D, Soland RM. *Optimal provisioning for single-echelon repairable item inventory control in a time-varying environment*, IIE Transactions; 21 (1989), No. 3, 202-212.

[5] Barlow RE, Proschan F. Mathematical theory of reliability. New-York NY: John Wiley and Sons, 1965.

[6] Bouti A, Aït-Kadi D. *A state-of-the-art review of FMEA/FMECA*, International Journal of Reliability, Quality and Safety Engineering; 1 (1994), No. 4, 515-543.

[7] Braglia M, Grassi A, Montanari R. *Multi-attribute classification method for spare parts inventory management*, Journal of Quality in Maintenance Engineering; 10 (2004), No. 1, 55-65.

[8] Chelbi A, Aït-Kadi D. Classement des équipements par ordre de priorité pour la maintenance, Journal of Decision Systems; 11 (2002), No. 1, 91-108.

[9] Chelbi A. Stratégies optimales d'inspection et de remplacements d'équipements sujets á des défaillances aléatoires [PhD Thesis]. Québec, Canada: Université Laval; 1996.

[10] Cléroux R, McConalogue DJ. *A numerical algorithm for recursively defined convolution integrals involving distribution functions*, Management Science; 22 (1976), 1138-1146.

[11] Dada M. *A two-echelon inventory system with priority shipments*, Management Science; 38 (1992), No. 8, 1140-1153.

[12] Dekker R, Kleijn MJ, Rooij PJ. *A spare parts stocking policy based on equipment criticality*, International Journal of Production Economics; 56 (2005), No. 1, 69-77.

[13] Diallo C, Aït-Kadi D, Chelbi A. *Integrated Spare Parts Management*. In Ben Daya M, Duffuaa SO, Raouf A, Knezevic J, Ait-Kadi D, editors. Handbook of Maintenance Management and Engineering. London. Springer; 2009.

[14] Diallo C, Aït-Kadi D, Chelbi A.(s, Q) *spare parts provisioning strategy for periodically replaced system*, IEEE Transactions on Reliability; 57 (2008), No. 1, 134-139.

[15] Diallo C , Aït-Kadi D. *A simple approximation to compute the mean number of failures for a repairable system and application to the availability determination* in Proceedings of the Congres International de Génie industriel Trois-Rivieres, Canada, 2007.

[16] Ebeling CE. An Introduction to reliability and maintainability engineering. Long Grove IL: Waveland Press, Inc.; 2010.

[17] Eisenhawer SW, Bott TF, Jackson JW. *Prioritizing the purchase of spare parts using an approximate reasoning model*, Proceedings of the Annual Reliability and Maintenability Symposium; 2002, 20-25.

[18] Gajpal PP, Ganesh LS, Rajnedran C. *Criticality analysis of spare parts using the analytic hierarchy process*, International Journal of Production Economics; 35 (1994), No. 1-3, 293-297.

[19] Gross D, Kahn HD, Marsh JD. *Queueing models for spares provisioning*, Naval Research Logistics Quarterly; 24 (1977), No. 4, 521-536.

[20] Gross D. *Sensitivity of queueing models to the assumption of exponentiality*, Naval Research Logistics Quarterly; 22 (1975), No. 2, 271-287.

[21] Guide VDR, Srivastava R. *Repairable inventory theory: Models and applications*, European Journal of Operational Research; 102 (1997), 1-20.

[22] Hammami A. Modélisation technico-économique d'une chaine logistique dans une entreprise réseau [PhD Thesis]. Québec, Canada: Université Laval - Ecole Nationale Supérieure des Mines de St-Etienne; 2003.

[23] Hausman WH, Erkip N. *Multi-echelon vs. single-echelon inventory control policies for low-demand items*, Management Science; 40 (1994), No. 5, 597-602.

[24] Huiskonen J. *Maintenance spare parts logistics: special characteristics and strategic choices*,International Journal of Production Economics; 71 (2001), 125-133.

[25] Kennedy WJ, Patterson JW, Fredendall LD. *An overview of recent literature on spare parts inventories*, International Journal of Production Economics; 76 (2002), 201-215.

[26] Kukreja A, Schmidt CP. *A model for lumpy demand parts in a multi-location inventory system with transshipments*. Computers and Operations Research; 32 (2005), 20591775.

[27] Kim JS, Shin kc, Yu hk. *Optimal algorithm to determine the spare inventory level for a repairable-item inventory system*, Computers and Operations Research; 23 (1996), No. 3, 289-297.

[28] Miller BL. *A queueing reward system with several customer classes*, Management Science; 16 (2005), No. 3, 234-245.

[29] Roy B, Bouyssou D. Aide multicritere á la décision: méthodes et cas, Paris: Economica; 1993.

[30] Sarker R, Haque A. *Optimization of maintenance and spare provisioning policy using simulation*, Applied Mathematical Modelling; 24 (2000),751170.

[31] Schärlig A. Pratiquer Electre et Prométhée. Lausanne: Presses polytechniques romandes; 1996.

[32] Schärlig A. Décider sur plusieurs criteres: panorama de l'aide á la décision multicritere. Lausanne: Presses polytechniques romandes, 1985.

[33] Sherbrooke CC. Optimal Inventory Modeling of systems. New York, NY: John Wiley; 1992.

[34] Sherbrooke CC. *Multiechelon inventory systems with lateral supply*, Naval Research Logistics; 39 (1992), 29-40.

[35] Taylor J, Jackson RRP. *An application of the birth and death process to the provision of spare machines*, Operations Research; 5 (1954), No. 4, 95-108.

[36] Wei BC. *A unified approach to failure mode effects and criticality analysis*, Proceedings of the Annual Reliability and Maintenability Symposium; (1991), 260-271.

A Non Homogeneous Riemann Solver for Two-phase Shallow Water Flows

Fayssal Benkhaldoun[1] **and Saida Sari**[1] **and Mohammed Seaid**[2]

[1] LAGA, Université Paris 13, 99 Av J.B. Clement, 93430 Villetaneuse, France
[2] School of Engineering and Computing Sciences, University of Durham, DH1 3LE, UK
[3] *Corresponding Author E-mail: fayssal@math.univ-paris13.fr*

Abstract. *The purpose of this research is to develop a simulation method for two-phase flows using shallow water equations. The hydraulics is modeled by the two-dimensional shallow water flows with variable horizontal density. The variation of density in the water flows can be attributed to the variation of thermal and salinity properties of the water. As an example of two-phase shallow water flows is the inclusion of the salty water from the sea into the fresh water of a river. Driving force of the phase separation and the mixing is the gradient of the density. For the numerical solution procedure we propose a non-homogeneous Riemann solver in the finite volume framework. The proposed method consists of a predictor stage for the discretization of gradient terms and a corrector stage for the treatment of source terms. The gradient fluxes are discretized using a modified Roe's scheme using the sign of the Jacobian matrix in the coupled system. A well-balanced discretization is used for the treatment of source terms. The efficiency of the solver is evaluated by several test problems for two-phase shallow water flows. The numerical results demonstrate high resolution of the proposed non-homogeneous Riemann solver and confirm its capability to provide accurate simulations for two-phase shallow water equations under flow regimes with strong shocks.*

Key words: Shallow water equations; variable density; finite volume method; unstructured grids; two-phase flows.

1. Introduction

Mathematical modelling of two-phase shallow water flows is based on the formulation and solution of the appropriate equations of continuity and motion of water height and density. In general, hydrodynamical flows represent a three-dimensional turbulent Newtonian flow in complicated geometrical domains. The costs of incorporating three-dimensional data in natural water courses

is often excessively high. Computational efforts needed to simulate three-dimensional turbulent flows can also be significant. In view of such considerations, many researchers have tended to use rational approximations in order to develop two-dimensional hydrodynamical models for water flows. Indeed, under the influence of gravity, many free-surface water flows can be modelled by the shallow water equations with the assumption that the vertical scale is much smaller than any typical horizontal scale. These equations can be derived from the depth-averaged incompressible Navier-Stokes equations using appropriate free-surface and boundary conditions along with a hydrostatic pressure assumption. The shallow water equations in depth-averaged form have been successfully applied to many engineering problems and their application fields include a wide spectrum of phenomena other than water waves. For instance, the shallow water equations have applications in environmental and hydraulic engineering such as tidal flows in an estuary or coastal regions, rivers, reservoir and open channel flows, see [16, 14, 13, 12, 10] among others. In general, two-phase shallow water flows are determined by the characteristics of the hydraulic flow and the properties of the suspended sediments. Thus, dynamics of the water and dynamics of the sediments must be studied using a mathematical model made of three different but dependent model variables: (i) a hydraulic variable defining the dynamics of the water flow, (ii) a sediment variable defining the transport and dispersion of the sediments and (iii) a state equation for the density to close the system. The model presented in this study includes the shallow water equations for the flow, conservation of sediment concentration, and state equation for updating the water density. Similar coupled sediment transport with surface water flow models have been developed in [8, 9] and the references therein. The governing equations form a two-dimensional nonlinear system of hyperbolic conservation laws with source terms. Such practical coupled hydrodynamical and sediment problems are not trivial to simulate since the geometry can be complex and the topography irregular.

The main concern of the two-phase shallow water flows is to determine the evolution of water free-surface and density for hydrodynamics systems such as rivers, estuaries, bays and other nearshore regions where water flows interact at an interface separating two bodies of water with different densities. Example of applications include among others, water interaction at the river-sea junctions, beach profile changes due to severe wave climates, water free-surface response to temperature or salinity variations, and harbour siltation. The ability to design numerical methods able to predict the hydraulics evolution of the coastal flows has a clear mathematical and engineering relevances. In practice, two-phase shallow water flows involve coupling between a hydrodynamics model, which provides a description of the flow field leading to a specification of local suspendent sediment rates, and an equation for the water density which expresses the conservative balance of suspendent sediment volume and its continual redistribution with time. Here, the hydrodynamic model is described by the shallow water equations and the suspended sediment transport is modelled by an advection equation accounting for density effects. The coupled models form a hyperbolic system of conservation laws with a source term. Nowadays, much effort has been devoted to develop numerical schemes for hydraulics models able to resolve all hydrodynamics and sediment scales. In the current study, a class of finite volume methods is proposed for numerical simulation of transient flows involving density variations. The method consists of a predictor stage where the numerical fluxes are constructed and a corrector stage to recover the conservation equations. The

sign matrix of the Jacobian matrix is used in the reconstruction of the numerical fluxes. Most of these techniques have been recently investigated in [1, 2, 3, 4, 6] for solving single-phase shallow water models without accounting for density variation. The current study presents an extension of this method to transient flows involving density variation in the water flows. A detailed formulation of the sign matrix and the numerical fluxes is presented. The proposed method also satisfies the property of well-balancing flux-gradient and source-term in the system. Numerical results will be shown for several two-phase shallow water flow problems.

The organization of this paper is as follows. In section 2. we present the governing equations for two-dimensional shallow water equations with horizontal variable density. The finite volume method is formulated in section 3.. This section includes both the discretization of gradient fluxes and treatment of source terms. Section 4. contains numerical results and applications. Conclusions are summarized in section 5.

2. Equations for Two-phase Shallow Water Flows

In the current study, the governing equations represent depth-averaged mass and momentum conservation of a water-species mixture, and mass conservation of the species in horizontal or/and vertical directions. These equations consist of the conservation of mass and momentum balance

$$\begin{aligned}
\frac{\partial(\rho_1 h)}{\partial t}+\frac{\partial}{\partial x}(\rho_1 hu)+\frac{\partial}{\partial y}(\rho_1 hv) &= R,\\
\frac{\partial}{\partial t}(\rho_1 hu)+\frac{\partial}{\partial x}\left(\rho_1 hu^2+\frac{1}{2}g\rho_1 h^2\right)+\frac{\partial}{\partial y}\left(\rho_1 huv\right) &= -g\rho_1 h\frac{\partial Z}{\partial x}-\Omega\rho_1 hv-\tau_{bx}+\tau_{wx}-\\
&\quad \nu\left(\frac{\partial^2}{\partial x^2}\left(\rho_1 hu\right)+\frac{\partial^2}{\partial y^2}\left(\rho_1 hu\right)\right), \quad (2.1)\\
\frac{\partial}{\partial t}(\rho_1 hv)+\frac{\partial}{\partial x}\left(\rho_1 huv\right)+\frac{\partial}{\partial y}\left(\rho_1 hv^2+\frac{1}{2}g\rho_1 h^2\right) &= -g\rho_1 h\frac{\partial Z}{\partial y}+\Omega\rho_1 hu-\tau_{by}+\tau_{wy}-\\
&\quad \nu\left(\frac{\partial^2}{\partial x^2}\left(\rho_1 hv\right)+\frac{\partial^2}{\partial y^2}\left(\rho_1 hv\right)\right),
\end{aligned}$$

for the first water body and

$$\begin{aligned}
\frac{\partial(\rho_2 h)}{\partial t}+\frac{\partial}{\partial x}(\rho_2 hu)+\frac{\partial}{\partial y}(\rho_2 hv) &= R,\\
\frac{\partial}{\partial t}(\rho_2 hu)+\frac{\partial}{\partial x}\left(\rho_2 hu^2+\frac{1}{2}g\rho_2 h^2\right)+\frac{\partial}{\partial y}\left(\rho_2 huv\right) &= -g\rho_2 h\frac{\partial Z}{\partial x}-\Omega\rho_2 hv-\tau_{bx}+\tau_{wx}-\\
&\quad \nu\left(\frac{\partial^2}{\partial x^2}\left(\rho_2 hu\right)+\frac{\partial^2}{\partial y^2}\left(\rho_2 hu\right)\right), \quad (2.2)\\
\frac{\partial}{\partial t}(\rho_2 hv)+\frac{\partial}{\partial x}\left(\rho_2 huv\right)+\frac{\partial}{\partial y}\left(\rho_2 hv^2+\frac{1}{2}g\rho_2 h^2\right) &= -g\rho_2 h\frac{\partial Z}{\partial y}+\Omega\rho_2 hu-\tau_{by}+\tau_{wy}-\\
&\quad \nu\left(\frac{\partial^2}{\partial x^2}\left(\rho_2 hv\right)+\frac{\partial^2}{\partial y^2}\left(\rho_2 hv\right)\right),
\end{aligned}$$

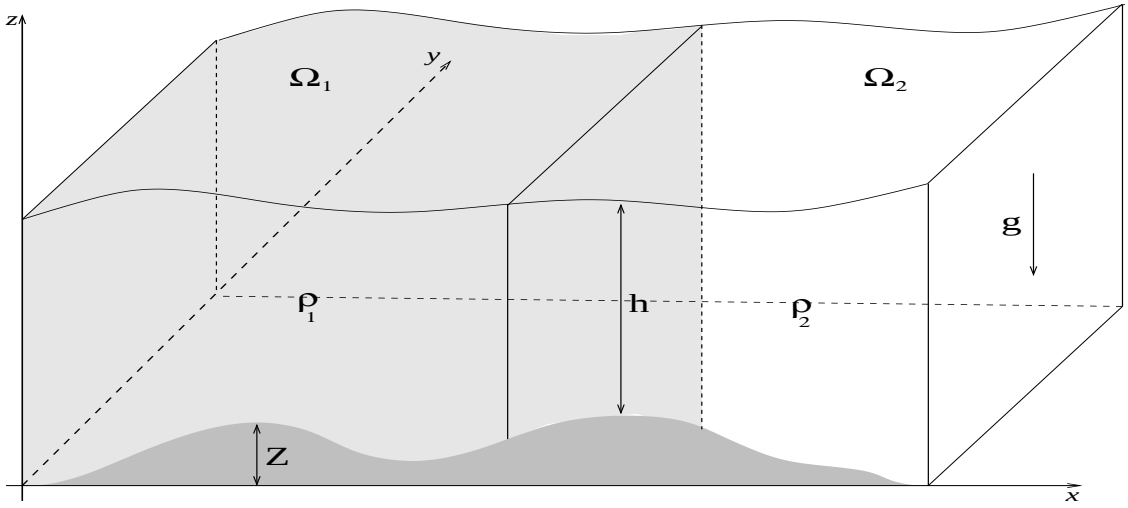


Figure 1: Sketch of a domain for two-dimensional shallow water flows.

for the second water body. Here, t is the time variable, $\mathbf{x} = (x, y)^T$ the space coordinates, $\mathbf{u} = (u, v)^T$ the depth-averaged water velocity, h the water depth, Z the bottom topography, g the gravitational acceleration, ρ_1 and ρ_2 the water density of the two water flows, ν the water viscosity, R the source term due to rainfall, Ω the Coriolis parameter defined by $\Omega = 2\omega \sin\phi$, with $\omega = 0.000073\ rad\ s^{-1}$ is the angular velocity of the earth and ϕ is the geographic latitude, see Figure 1 for an illustration. For a constant density, the above equations reduce to the standard shallow water equations widely used in the literature, see [**?**, **?**] among others. In the current work, we assume that the density depends on space and time variables, *i.e.*, $\rho = \rho(x, y, t)$. For a two-phase system as shown in Figure 1, the density is given as

$$\rho(x,y,t) = \begin{cases} \rho_1, & \text{if} \quad (x,y) \in \Omega_1, \\ \\ \rho_2, & \text{if} \quad (x,y) \in \Omega_2. \end{cases}$$

$$\begin{aligned} \frac{\partial(\rho h)}{\partial t} + \frac{\partial}{\partial x}(\rho h u) + \frac{\partial}{\partial y}(\rho h v) &= R, \\ \frac{\partial}{\partial t}(\rho h u) + \frac{\partial}{\partial x}\left(\rho h u^2 + \frac{1}{2} g \rho h^2\right) + \frac{\partial}{\partial y}\left(\rho h u v\right) &= -g\rho h \frac{\partial Z}{\partial x} - \Omega \rho h v - \tau_{bx} + \tau_{wx} - \\ &\quad \nu\left(\frac{\partial^2}{\partial x^2}\left(\rho h u\right) + \frac{\partial^2}{\partial y^2}\left(\rho h u\right)\right), \\ \frac{\partial}{\partial t}(\rho h v) + \frac{\partial}{\partial x}\left(\rho h u v\right) + \frac{\partial}{\partial y}\left(\rho h v^2 + \frac{1}{2} g \rho h^2\right) &= -g\rho h \frac{\partial Z}{\partial y} + \Omega \rho h u - \tau_{by} + \tau_{wy} - \\ &\quad \nu\left(\frac{\partial^2}{\partial x^2}\left(\rho h v\right) + \frac{\partial^2}{\partial y^2}\left(\rho h v\right)\right), \end{aligned} \tag{2.3}$$

For the time evolution of the density, the equation used to close the system is given by

$$\rho = \rho_w(1-c) + \rho_s c, \tag{2.4}$$

where ρ_s is the sediment density and c is the depth-averaged concentration of the suspended sediment. The equation for mass conservation of species is modeled by

$$\frac{\partial}{\partial t}(\rho_s hc) + \frac{\partial}{\partial x}(\rho_s huc) + \frac{\partial}{\partial y}(\rho_s hvc) = S - \kappa\left(\frac{\partial^2}{\partial x^2}(\rho_s hc) + \frac{\partial^2}{\partial y^2}(\rho_s hc)\right), \tag{2.5}$$

where κ is the sediment viscosity and S is a source/sink term. In (2.3), τ_{bx} and τ_{by} are the bed shear stress in the x- and y-direction, respectively, defined by the depth-averaged velocities as

$$\tau_{bx} = \rho C_b u\sqrt{u^2+v^2}, \qquad \tau_{by} = \rho C_b v\sqrt{u^2+v^2}, \tag{2.6}$$

where C_b is the bed friction coefficient. The surface stress τ_w is usually originated by the shear of the blowing wind and is expressed as a quadratic function of the wind velocity,

$$\tau_{wx} = \rho C_w w_x\sqrt{w_x^2+w_y^2}, \qquad \tau_{by} = \rho C_w w_y\sqrt{w_x^2+w_y^2}, \tag{2.7}$$

with C_w is the coefficient of wind friction and $\mathbf{w} = (w_x, w_y)^T$ is the velocity of wind at $10\ m$ above water surface.

For simplicity in the presentation, let us rewrite the equations (2.3) and (2.5) in the following vector form

$$\frac{\partial \mathbf{W}}{\partial t} + \frac{\partial}{\partial x}\left(\mathbf{F}(\mathbf{W}) - \tilde{\mathbf{F}}(\mathbf{W})\right) + \frac{\partial}{\partial y}\left(\mathbf{G}(\mathbf{W}) - \tilde{\mathbf{G}}(\mathbf{W})\right) = \mathbf{S}(\mathbf{W}) + \mathbf{Q}(\mathbf{W}), \tag{2.8}$$

where $\mathbf{W}$ is the vector of conserved variables, $\mathbf{F}$ and $\mathbf{G}$ are the physical fluxes in x- and y-direction, $\tilde{\mathbf{F}}$ and $\tilde{\mathbf{G}}$ are the diffusion fluxes, $\mathbf{S}$ and $\mathbf{Q}$ are the source terms. These variables are defined as

$$\mathbf{W} = \begin{pmatrix} \rho h \\ \rho hu \\ \rho hv \\ \rho_s hc \end{pmatrix}, \qquad \mathbf{F}(\mathbf{W}) = \begin{pmatrix} \rho hu \\ \rho hu^2 + \frac{1}{2}g\rho h^2 \\ \rho huv \\ \rho_s huc \end{pmatrix}, \qquad \mathbf{G}(\mathbf{W}) = \begin{pmatrix} \rho hv \\ \rho huv \\ \rho hv^2 + \frac{1}{2}g\rho h^2 \\ \rho_s hvc \end{pmatrix},$$

$$\tilde{\mathbf{F}}(\mathbf{W}) = \begin{pmatrix} 0 \\ \nu\frac{\partial}{\partial x}(\rho hu) \\ \nu\frac{\partial}{\partial x}(\rho hv) \\ \kappa\frac{\partial}{\partial x}(\rho_s hc) \end{pmatrix}, \qquad \tilde{\mathbf{G}}(\mathbf{W}) = \begin{pmatrix} 0 \\ \nu\frac{\partial}{\partial y}(\rho hu) \\ \nu\frac{\partial}{\partial y}(\rho hv) \\ \kappa\frac{\partial}{\partial y}(\rho_s hc) \end{pmatrix},$$

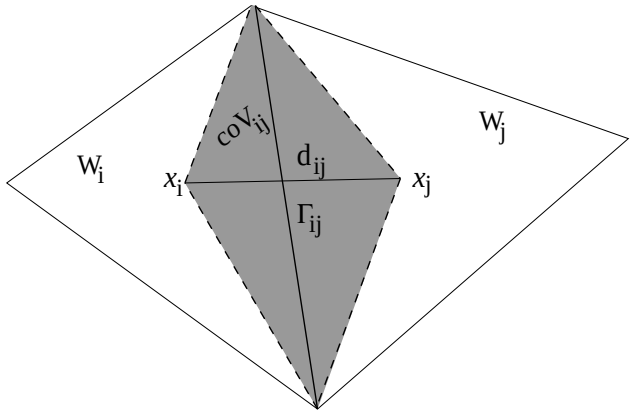


Figure 2: Generic control volumes and notations.

$$\mathbf{S}(\mathbf{W}) = \begin{pmatrix} 0 \\ -g\rho h\dfrac{\partial Z}{\partial x} \\ -g\rho h\dfrac{\partial Z}{\partial y} \\ 0 \end{pmatrix}, \qquad \mathbf{Q}(\mathbf{W}) = \begin{pmatrix} R \\ -\Omega\rho hv - \tau_{bx} + \tau_{wx} \\ \Omega\rho hu - \tau_{by} + \tau_{wy} \\ S \end{pmatrix}.$$

Note that the equations (2.8) has to be solved in a bounded spatial domain $\mathcal{D}$ with smooth boundary $\partial\mathcal{D}$, equipped with given boundary and initial conditions. In practice, these conditions are problem dependent and their discussion is postponed for section 4. where numerical examples are discussed.

3. Unstructured Non Homogeneous Riemann Solver

The governing equations (2.8) are formulated in Cartesian coordinates and will be discretized into the unstructured grids by the finite volume method. The unstructured grids are polygons and the number of edges of the grids is not limited in theory, but only triangular grids are considered in the current study. Hence, we divide the time interval into sub-intervals $[t_n, t_{n+1}]$ with stepsize Δt and discretize the spatial domain in conforming triangular elements $\mathcal{T}_i$. Each triangle represents a control volume and the variables are located at the geometric centres of the cells. Hence, using the control volume depicted in the left plot of Figure 2, a finite volume discretization of (2.8) yields

$$\begin{aligned} \mathbf{W}_i^{n+1} &= \mathbf{W}_i^n - \frac{\Delta t}{|\mathcal{T}_i|}\sum_{j\in N(i)}\int_{\Gamma_{ij}} \mathcal{F}(\mathbf{W}^n;\mathbf{n})\, d\sigma + \frac{\Delta t}{|\mathcal{T}_i|}\sum_{j\in N(i)}\int_{\Gamma_{ij}} \tilde{\mathcal{F}}(\mathbf{W}^n;\mathbf{n})\, d\sigma \qquad (3.1) \\ &+ \frac{\Delta t}{|\mathcal{T}_i|}\int_{\mathcal{T}_i} \mathbf{S}(\mathbf{W}^n)\, dV + \frac{\Delta t}{|\mathcal{T}_i|}\int_{\mathcal{T}_i} \mathbf{Q}(\mathbf{W}^n)\, dV, \end{aligned}$$

where $N(i)$ is the set of neighboring triangles of the cell $\mathcal{T}_i$, $\mathbf{W}_i^n$ is an averaged value of the solution $\mathbf{W}$ in the cell $\mathcal{T}_i$ at time t_n,

$$\mathbf{W}_i = \frac{1}{|\mathcal{T}_i|}\int_{\mathcal{T}_i} \mathbf{W}\ dV,$$

where $|\mathcal{T}_i|$ denotes the area of $\mathcal{T}_i$. Here, Γ_{ij} is the interface between the two control volumes $\mathcal{T}_i$ and $\mathcal{T}_j$, $\mathbf{n} = (n_x, n_y)^T$ denotes the unit outward normal to Γ_{ij}, and

$$\mathcal{F}(\mathbf{W};\mathbf{n}) = \mathbf{F}(\mathbf{W})n_x + \mathbf{G}(\mathbf{W})n_y, \qquad \tilde{\mathcal{F}}(\mathbf{W};\mathbf{n}) = \tilde{\mathbf{F}}(\mathbf{W})n_x + \tilde{\mathbf{G}}(\mathbf{W})n_y.$$

To deal with the source terms $\mathbf{Q}$, a standard splitting procedure (see for instance [11]) is employed for the discrete system (3.2) as

$$\begin{aligned}
\mathbf{W}_i^* &= \mathbf{W}_i^n - \frac{\Delta t}{|\mathcal{T}_i|}\sum_{j\in N(i)}\int_{\Gamma_{ij}} \mathcal{F}(\mathbf{W}^n;\mathbf{n})\ d\sigma + \frac{\Delta t}{|\mathcal{T}_i|}\int_{\mathcal{T}_i}\mathbf{S}(\mathbf{W}^n)\ dV,\\
\mathbf{W}_i^{**} &= \mathbf{W}_i^* + \frac{\Delta t}{|\mathcal{T}_i|}\sum_{j\in N(i)}\int_{\Gamma_{ij}} \tilde{\mathcal{F}}(\mathbf{W}^*;\mathbf{n})\ d\sigma,\\
\mathbf{W}_i^{n+1} &= \mathbf{W}_i^{**} + \frac{\Delta t}{|\mathcal{T}_i|}\int_{\mathcal{T}_i}\mathbf{Q}(\mathbf{W}^{**})\ dV.
\end{aligned} \tag{3.2}$$

Note that the time splitting (3.2) is only first-order accurate in time. A second-order splitting for the system (3.2) can be derived analogously using the Strang method [15]. The finite volume discretization (3.2) is complete once the gradient fluxes $\mathcal{F}(\mathbf{W};\mathbf{n})$, diffusion fluxes $\tilde{\mathcal{F}}(\mathbf{W};\eta)$ and a discretization of source terms $\mathbf{S}(\mathbf{W})$ are well defined.

3.1. Discretization of the gradient fluxes

Let us consider the hyperbolic parts in the system (2.8) given by

$$\frac{\partial \mathbf{W}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{W})}{\partial x} + \frac{\partial \mathbf{G}(\mathbf{W})}{\partial y} = \mathbf{S}(\mathbf{W}). \tag{3.3}$$

Applied to the system (3.3), the finite volume discretization over the control volume $\mathcal{T}_i$ yields

$$\begin{aligned}
\frac{\partial}{\partial t}\int_{\mathcal{T}_i}\rho h\ dV + \oint_{\mathcal{S}_i}(\rho h u n_x + \rho h v n_y)\ d\sigma &= 0,\\
\frac{\partial}{\partial t}\int_{\mathcal{T}_i}\rho h u\ dV + \oint_{\mathcal{S}_i}\left(\left(\rho h u^2 + \frac{1}{2}g\rho h^2\right)n_x + \rho h u v n_y\right)d\sigma &= -g\rho h\oint_{\mathcal{S}_i} Z n_x\ d\sigma,\\
\frac{\partial}{\partial t}\int_{\mathcal{T}_i}\rho h v\ dV + \oint_{\mathcal{S}_i}\left(\rho h u v n_x + \left(\rho h v^2 + \frac{1}{2}g\rho h^2\right)n_y\right)d\sigma &= -g\rho h\oint_{S_i} Z n_y\ d\sigma,\\
\frac{\partial}{\partial t}\int_{\mathcal{T}_i}\rho_s h c\ dV + \oint_{\mathcal{S}_i}(\rho_s h u c n_x + \rho_s h v c n_y)\ d\sigma &= 0,
\end{aligned}$$

where $\mathcal{S}_i$ is the surface surrounding the control volume $\mathcal{T}_i$. Using the expressions of the normal velocity $u_\eta = un_x + vn_y$ and tangential velocity $u_\tau = -un_y + vn_x$, the above equations can be reformulated as

$$\begin{aligned}
\frac{\partial}{\partial t}\int_{\mathcal{T}_i} \rho h \, dV + \oint_{\mathcal{S}_i} \rho h u_\eta \, d\sigma &= 0, \\
\frac{\partial}{\partial t}\int_{\mathcal{T}_i} \rho h u \, dV + \oint_{\mathcal{S}_i} \left(\rho h u u_\eta + \frac{1}{2} g \rho h^2 n_x\right) d\sigma &= -g\rho h \oint_{S_i} Z n_x \, d\sigma, \\
\frac{\partial}{\partial t}\int_{\mathcal{T}_i} \rho h v \, dV + \oint_{\mathcal{S}_i} \left(\rho h v u_\eta + \frac{1}{2} g \rho h^2 n_y\right) d\sigma &= -g\rho h \oint_{S_i} Z n_y \, d\sigma, \\
\frac{\partial}{\partial t}\int_{\mathcal{T}_i} \rho_s h c \, dV + \oint_{\mathcal{S}_i} \rho_s h u_\eta c d\sigma &= 0,
\end{aligned} \tag{3.4}$$

In order to simplify the system (3.4), we first sum the second equation multiplied by n_x to the third equation multiplied by n_y, then we subtract the third equation multiplied by n_x from the second equation multiplied by n_y. The result of these operations is

$$\begin{aligned}
\frac{\partial}{\partial t}\int_{\mathcal{T}_i} \rho h \, dV + \oint_{\mathcal{S}_i} \rho h u_\eta \, d\sigma &= 0, \\
\frac{\partial}{\partial t}\int_{\mathcal{T}_i} \rho h u_\eta \, dV + \oint_{\mathcal{S}_i} \left(\rho h u_\eta u_\eta + \frac{1}{2} g \rho h^2\right) d\sigma &= -g\rho h \oint_{S_i} Z \, d\sigma, \\
\frac{\partial}{\partial t}\int_{\mathcal{T}_i} \rho h u_\tau \, dV + \oint_{\mathcal{S}_i} \rho h u_\tau u_\eta \, d\sigma &= 0, \\
\frac{\partial}{\partial t}\int_{\mathcal{T}_i} \rho_s h c \, dV + \oint_{\mathcal{S}_i} \rho_s h u_\eta c d\sigma &= 0,
\end{aligned} \tag{3.5}$$

which can be reformulated in a differential form as

$$\begin{aligned}
\frac{\partial \rho h}{\partial t} + \frac{\partial (\rho h u_\eta)}{\partial \eta} &= 0, \\
\frac{\partial (\rho h u_\eta)}{\partial t} + \frac{\partial}{\partial \eta}\left(\rho h u_\eta^2 + \frac{1}{2} g \rho h^2\right) &= -g\rho h \frac{\partial Z}{\partial \eta}, \\
\frac{\partial (\rho h u_\tau)}{\partial t} + \frac{\partial}{\partial \eta}\left(\rho h u_\eta u_\tau\right) &= 0, \\
\frac{\partial (\rho_s h c)}{\partial t} + \frac{\partial}{\partial \eta}\left(\rho_s h u_\eta c\right) &= 0,
\end{aligned} \tag{3.6}$$

An equivalent system of (3.6) can also be rewritten in a vector form as

$$\frac{\partial \mathbf{U}}{\partial t} + \mathbf{A}_\eta(\mathbf{U}) \frac{\partial \mathbf{U}}{\partial \eta} = \mathbf{0}. \tag{3.7}$$

where

$$\mathbf{U} = \begin{pmatrix} \rho h \\ \rho h u_\eta \\ \rho h u_\tau \\ \rho_s h c \end{pmatrix},$$

$$\mathbf{A}_\eta(\mathbf{U}) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \dfrac{g(\rho+\rho_w)h}{2\rho_w} - u_\eta^2 & 2u_\eta & 0 & -\dfrac{g(\rho_s-\rho_w)\rho h}{2\rho_s\rho_w} \\ -u_\eta u_\tau & u_\tau & u_\eta & 0 \\ -\dfrac{\rho_s c u_\eta}{\rho} & \dfrac{\rho_s c}{\rho} & 0 & u_\eta \end{pmatrix}.$$

One of the advantages in considering the projected system (3.7) is that no discretization of source terms is required. Thus, in the predictor stage, we use the projected system (3.7) to compute the averaged states as

$$\mathbf{U}_{ij}^n = \frac{1}{2}\left(\mathbf{U}_i^n + \mathbf{U}_j^n\right) - \frac{1}{2}\operatorname{sgn}\left[\mathbf{A}_\eta\left(\overline{\mathbf{U}}\right)\right]\left(\mathbf{U}_j^n - \mathbf{U}_i^n\right), \tag{3.8}$$

where the sign matrix of the Jacobian is defined as

$$\operatorname{sgn}\left[\nabla \mathbf{F}_\eta\left(\overline{\mathbf{U}}\right)\right] = \mathcal{R}(\overline{\mathbf{U}})\operatorname{sgn}\left[\Lambda(\overline{\mathbf{U}})\right]\mathcal{R}^{-1}(\overline{\mathbf{U}}),$$

with $\Lambda(\overline{\mathbf{U}})$ is the diagonal matrix of eigenvalues, and $\mathcal{R}(\overline{\mathbf{U}})$ is the right eigenvector matrix. These matrices can be explicitly expressed using the associated eigenvalues of $\mathbf{A}_\eta(\mathbf{U})$ in (3.7). Here, $\overline{\mathbf{U}}$ is the Roe's averaged state given by

$$\overline{\mathbf{U}} = \begin{pmatrix} \dfrac{h_i+h_j}{2}\left(\dfrac{\rho_i\sqrt{h_i}+\rho_j\sqrt{h_j}}{\sqrt{h_i}+\sqrt{h_j}}\right) \\ \dfrac{h_i+h_j}{2}\left(\dfrac{u_i\sqrt{h_i}+u_j\sqrt{h_j}}{\sqrt{h_i}+\sqrt{h_j}}\eta_x + \dfrac{v_i\sqrt{h_i}+v_j\sqrt{h_j}}{\sqrt{h_i}+\sqrt{h_j}}\eta_y\right)\left(\dfrac{\rho_i\sqrt{h_i}+\rho_j\sqrt{h_j}}{\sqrt{h_i}+\sqrt{h_j}}\right) \\ \dfrac{h_i+h_j}{2}\left(-\dfrac{u_i\sqrt{h_i}+u_j\sqrt{h_j}}{\sqrt{h_i}+\sqrt{h_j}}\eta_y + \dfrac{v_i\sqrt{h_i}+v_j\sqrt{h_j}}{\sqrt{h_i}+\sqrt{h_j}}\eta_x\right)\left(\dfrac{\rho_i\sqrt{h_i}+\rho_j\sqrt{h_j}}{\sqrt{h_i}+\sqrt{h_j}}\right) \\ \rho_s\dfrac{h_i+h_j}{2}\left(\dfrac{c_i\sqrt{h_i}+c_j\sqrt{h_j}}{\sqrt{h_i}+\sqrt{h_j}}\right) \end{pmatrix}. \tag{3.9}$$

Once the states $\mathbf{U}_{ij}^n$ are calculated in the predictor stage (3.8), the states $\mathbf{W}_{ij}^n$ are recovered by using the transformations $v = (u_\tau, u_\eta)\cdot\eta$ and $u = (u_\tau, u_\eta)\cdot\tau$. Thus, applied to the system (3.3), the proposed finite volume method consists of a predictor stage and a corrector stage and can be

formulated as:

$$\begin{aligned} \mathbf{U}_{ij}^{n} &= \frac{1}{2}\left(\mathbf{U}_{i}^{n}+\mathbf{U}_{j}^{n}\right)-\frac{1}{2}\operatorname{sgn}\left[\mathbf{A}_{\eta}\left(\overline{\mathbf{U}}\right)\right]\left(\mathbf{U}_{j}^{n}-\mathbf{U}_{i}^{n}\right), \\ \mathbf{W}_{i}^{n+1} &= \mathbf{W}_{i}^{n}-\frac{\Delta t}{|\mathcal{T}_{i}|}\sum_{j\in N(i)}\mathcal{F}\left(\mathbf{W}_{ij}^{n};\eta_{ij}\right)|\Gamma_{ij}|+\Delta t\mathbf{S}_{i}^{n}, \end{aligned} \tag{3.10}$$

Next we discuss the formulation of matrices $\mathcal{R}(\overline{\mathbf{U}})$ and $\mathcal{R}^{-1}(\overline{\mathbf{U}})$, the treatment of source terms $\mathbf{S}_i^n$ in the proposed finite volume scheme and also the extension of the scheme to a second-order accuracy.

3.2. Determination of the sign matrix

The four eigenvalues corresponding to the projected system (3.7) are

$$\begin{aligned} \lambda_1 &= u_\eta-\sqrt{gh}, \\ \lambda_2 &= 0, \\ \lambda_3 &= u_\eta, \\ \lambda_4 &= u_\eta, \\ \lambda_5 &= u_\eta+\sqrt{gh}. \end{aligned}$$

Hence, the sign matrix in (3.8) is defined as

$$\operatorname{sgn}\left[\mathbf{A}_{\eta}\left(\overline{\mathbf{U}}\right)\right]=\mathcal{R}(\overline{\mathbf{U}})\operatorname{sgn}\left[\Lambda(\overline{\mathbf{U}})\right]\mathcal{R}^{-1}(\overline{\mathbf{U}}),$$

where $\overline{\mathbf{U}}$ is the Roe's average state given by (3.9), $\mathcal{R}(\overline{\mathbf{U}})$ and $\Lambda(\overline{\mathbf{U}})$ are respectively, the right eigenvector and the diagonal matrices reconstructed as

$$\mathcal{R}(\overline{\mathbf{U}})=\begin{pmatrix} 1 & 1 & 1 & 0 & 1 \\ \bar{\lambda}_1 & 0 & \bar{u}_\eta & 0 & \bar{\lambda}_5 \\ \bar{u}_\tau & \bar{u}_\tau & 0 & 1 & \bar{u}_\tau \\ \dfrac{\rho_s\bar{c}}{\bar{\rho}} & \dfrac{\rho_s\bar{c}}{\bar{\rho}} & \dfrac{\rho_s(\bar{\rho}+\rho_w)}{\bar{\rho}(\rho_s-\rho_w)} & 0 & \dfrac{\rho_s\bar{c}}{\bar{\rho}} \\ 0 & \dfrac{-4\bar{\lambda}_1\bar{\lambda}_5}{(\bar{\lambda}_1-\bar{\lambda}_5)^2} & 0 & 0 & 0 \end{pmatrix},$$

$$\Lambda(\overline{\mathbf{U}})=\begin{pmatrix} \bar{\lambda}_1 & 0 & 0 & 0 & 0 \\ 0 & \bar{\lambda}_2 & 0 & 0 & 0 \\ 0 & 0 & \bar{\lambda}_3 & 0 & 0 \\ 0 & 0 & 0 & \bar{\lambda}_4 & 0 \\ 0 & 0 & 0 & 0 & \bar{\lambda}_5 \end{pmatrix},$$

$$\mathcal{R}^{-1}(\overline{\mathbf{U}}) = \begin{pmatrix} \dfrac{\bar{u}_\eta(\bar{\rho}-\rho_w)-\bar{\lambda}_5(\bar{\rho}+\rho_w)}{2\rho_w(\bar{\lambda}_1-\bar{\lambda}_5)} & \dfrac{1}{\bar{\lambda}_1-\bar{\lambda}_5} & 0 & \dfrac{\bar{\rho}(\rho_s-\rho_w)}{2\rho_s\rho_w}\dfrac{\bar{\lambda}_5-\bar{u}_\eta}{\bar{\lambda}_1-\bar{\lambda}_5} & -\dfrac{\bar{\lambda}_1-\bar{\lambda}_5}{4\bar{\lambda}_1} \\ 0 & 0 & 0 & 0 & -\dfrac{(\bar{\lambda}_1-\bar{\lambda}_5)^2}{4\bar{\lambda}_1\bar{\lambda}_5} \\ -\dfrac{\bar{\rho}-\rho_w}{2\rho_w} & 0 & 0 & \dfrac{\bar{\rho}(\rho_s-\rho_w)}{2\rho_s\rho_w} & 0 \\ -\dfrac{\bar{u}_\tau(\bar{\rho}+\rho_w)}{2\rho_w} & 0 & 1 & \dfrac{\bar{u}_\tau\bar{\rho}(\rho_s-\rho_w)}{2\rho_s\rho_w} & 0 \\ -\dfrac{\bar{u}_\eta(\bar{\rho}-\rho_w)-\bar{\lambda}_1(\bar{\rho}+\rho_w)}{2\rho_w(\bar{\lambda}_1-\bar{\lambda}_5)} & -\dfrac{1}{\bar{\lambda}_1-\bar{\lambda}_5} & 0 & -\dfrac{\bar{\rho}(\rho_s-\rho_w)}{2\rho_s\rho_w}\dfrac{\bar{\lambda}_1-\bar{u}_\eta}{\bar{\lambda}_1-\bar{\lambda}_5} & \dfrac{\bar{\lambda}_1-\bar{\lambda}_5}{4\bar{\lambda}_5} \end{pmatrix}$$

It is easy to verify that the sign matrix in (3.8) is given by

$$\operatorname{sgn}\Big[\mathbf{A}_\eta\left(\overline{\mathbf{U}}\right)\Big] = \begin{pmatrix} s_{11} & s_{12} & 0 & s_{14} & s_{15} \\ s_{21} & s_{22} & 0 & s_{24} & s_{25} \\ s_{31} & s_{32} & \operatorname{sgn}(\bar{\lambda}_4) & s_{34} & s_{35} \\ s_{41} & s_{42} & 0 & s_{44} & s_{45} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\begin{aligned}
s_{11} &= l_{11}\operatorname{sgn}(\bar{\lambda}_1) + l_{31}\operatorname{sgn}(\bar{\lambda}_3) + l_{51}\operatorname{sgn}(\bar{\lambda}_5),\\
s_{12} &= l_{12}\operatorname{sgn}(\bar{\lambda}_1) + l_{52}\operatorname{sgn}(\bar{\lambda}_5),\\
s_{14} &= l_{14}\operatorname{sgn}(\bar{\lambda}_1) + l_{34}\operatorname{sgn}(\bar{\lambda}_3) + l_{54}\operatorname{sgn}(\bar{\lambda}_5),\\
s_{15} &= l_{15}\operatorname{sgn}(\bar{\lambda}_1) + l_{55}\operatorname{sgn}(\bar{\lambda}_5),\\
s_{21} &= l_{11}r_{21}\operatorname{sgn}(\bar{\lambda}_1) + l_{31}r_{23}\operatorname{sgn}(\bar{\lambda}_3) + l_{51}r_{25}\operatorname{sgn}(\bar{\lambda}_5),\\
s_{22} &= l_{12}r_{21}\operatorname{sgn}(\bar{\lambda}_1) + l_{52}r_{25}\operatorname{sgn}(\bar{\lambda}_5),\\
s_{24} &= l_{14}r_{21}\operatorname{sgn}(\bar{\lambda}_1) + l_{34}r_{23}\operatorname{sgn}(\bar{\lambda}_3) + l_{54}r_{25}\operatorname{sgn}(\bar{\lambda}_5),\\
s_{25} &= l_{15}r_{21}\operatorname{sgn}(\bar{\lambda}_1) + l_{55}r_{25}\operatorname{sgn}(\bar{\lambda}_5),\\
s_{31} &= l_{11}r_{31}\operatorname{sgn}(\bar{\lambda}_1) + l_{41}\operatorname{sgn}(\bar{\lambda}_4) + l_{51}r_{35}\operatorname{sgn}(\bar{\lambda}_5),\\
s_{32} &= l_{12}r_{31}\operatorname{sgn}(\bar{\lambda}_1) + l_{52}r_{35}\operatorname{sgn}(\bar{\lambda}_5),\\
s_{34} &= l_{14}r_{31}\operatorname{sgn}(\bar{\lambda}_1) + l_{44}\operatorname{sgn}(\bar{\lambda}_4) + l_{54}r_{35}\operatorname{sgn}(\bar{\lambda}_5),\\
s_{35} &= l_{15}r_{31}\operatorname{sgn}(\bar{\lambda}_1) + l_{55}r_{35}\operatorname{sgn}(\bar{\lambda}_5),\\
s_{41} &= l_{11}r_{41}\operatorname{sgn}(\bar{\lambda}_1) + l_{31}r_{43}\operatorname{sgn}(\bar{\lambda}_3) + l_{51}r_{45}\operatorname{sgn}(\bar{\lambda}_5),\\
s_{42} &= l_{12}r_{41}\operatorname{sgn}(\bar{\lambda}_1) + l_{52}r_{45}\operatorname{sgn}(\bar{\lambda}_5),\\
s_{44} &= l_{14}r_{41}\operatorname{sgn}(\bar{\lambda}_1) + l_{34}r_{43}\operatorname{sgn}(\bar{\lambda}_3) + l_{54}r_{45}\operatorname{sgn}(\bar{\lambda}_5),\\
s_{45} &= l_{15}r_{41}\operatorname{sgn}(\bar{\lambda}_1) + l_{55}r_{45}\operatorname{sgn}(\bar{\lambda}_5),
\end{aligned}$$

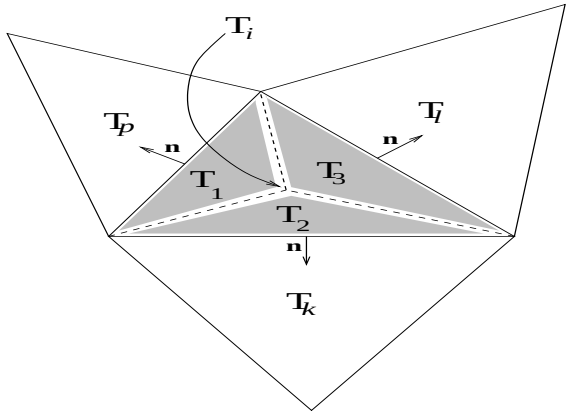


Figure 3: Sub-triangles used in the discretization of source terms.

where r_{ij} and l_{ij} are the entries of the matrix $\mathcal{R}(\overline{\mathbf{U}})$ and $\mathcal{R}^{-1}(\overline{\mathbf{U}})$, respectively.

3.3. Discretization of the source term

The treatment of the source terms in the shallow water equations presents a challenge in many numerical methods, compare [7, **?**, 1] among others. In our scheme, the source term approximation $\mathbf{S}_i^n$ in the corrector stage is reconstructed such that the still-water equilibrium (C-property) [**?**] is satisfied. Here, a numerical scheme is said to satisfy the C-property for the equations (2.8) if the condition

$$u = v = 0, \qquad h + Z = H, \qquad \rho = C, \tag{3.11}$$

holds for stationary flows at rest. In (3.11), H and C are nonnegative constants. Therefore, the treatment of source terms in (3.10) is reconstructed such that the condition (3.11) is preserved at the discretized level. Remark that the last condition in (3.11) means that at the equilibrium the sediment medium is assumed to be saturated. Furthermore, from the density equation (2.5), a constant density is equivalent to a constant concentration c. Hence, $\mathbf{S}_i^n$ should be a consistent discretization of the source term in (3.3) defined as

$$\mathbf{S}_i^n = \begin{pmatrix} 0 \\ -g\bar{h}_{xi}^n \bar{\rho}_{xi}^n \displaystyle\sum_{j\in N(i)} Z_{ij} n_{xij} \left|\Gamma_{ij}\right| \\ -g\bar{h}_{yi}^n \bar{\rho}_{yi}^n \displaystyle\sum_{j\in N(i)} Z_{ij} n_{yij} \left|\Gamma_{ij}\right| \\ 0 \end{pmatrix}. \tag{3.12}$$

The approximations $\bar{h}_{xi}^n$ and $\bar{h}_{yi}^n$ are reconstructed using a technique recently developed in [1] for the proposed finite volume method to satisfy the well-known C-property in standard shallow water flow over fixed beds. In this section we briefly describe the formulation of this procedure and more details can be found in [1]. Hence, at the stationary state, the numerical flux in the corrector stage

yields

$$\sum_{j\in N(i)} \mathcal{F}\left(\mathbf{W}_{ij}^n;\mathbf{n}_{ij}\right) = \begin{pmatrix} 0 \\ -g\displaystyle\int_{\mathcal{T}_i} \rho h \frac{\partial Z}{\partial x}\, dV \\ -g\displaystyle\int_{\mathcal{T}_i} \rho h \frac{\partial Z}{\partial y}\, dV \\ 0 \end{pmatrix},$$

which is equivalent to

$$\begin{pmatrix} 0 \\ \displaystyle\sum_{j\in N(i)} \frac{1}{2} g \rho_{ij}^n \left(h_{ij}^n\right)^2 N_{xij} \\ \displaystyle\sum_{j\in N(i)} \frac{1}{2} g \rho_{ij}^n \left(h_{ij}^n\right)^2 N_{yij} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -g\displaystyle\int_{\mathcal{T}_i} \rho h \frac{\partial Z}{\partial x}\, dV \\ -g\displaystyle\int_{\mathcal{T}_i} \rho h \frac{\partial Z}{\partial y}\, dV \\ 0 \end{pmatrix}. \tag{3.13}$$

where $N_{xij} = n_{xij}\ |\Gamma_{ij}|$ and $N_{yij} = n_{yij}\ |\Gamma_{ij}|$. Next, to approximate the source terms we proceed as follows. First we decompose the triangle $\mathcal{T}_i$ into three sub-triangles as depicted in Figure 3. Then, the source term is approximated as

$$\int_{\mathcal{T}_i} \rho h \frac{\partial Z}{\partial x}\, dV = \int_{\mathcal{T}_1} \rho h \frac{\partial Z}{\partial x}\, dV + \int_{\mathcal{T}_2} \rho h \frac{\partial Z}{\partial x}\, dV + \int_{\mathcal{T}_3} \rho h \frac{\partial Z}{\partial x}\, dV, \tag{3.14}$$

where

$$\int_{\mathcal{T}_1} \rho h \frac{\partial Z}{\partial x}\, dV = \rho_1 h_1 \int_{\mathcal{T}_1} \frac{\partial Z}{\partial x}\, dV,$$

with h_1 is an average value of h on the sub-triangle $\mathcal{T}_1$. Hence,

$$\begin{aligned} \int_{\mathcal{T}_1} \rho h \frac{\partial Z}{\partial x}\, dV &= \rho_1 h_1 \sum_{j\in N(1)} \int_{\Gamma_{1j}} Z n_x\, d\sigma, \\ &= \rho_1 h_1 \sum_{j\in N(1)} Z_{1j}\ N_{x1j}, \\ &= \rho_1 h_1 \sum_{j\in N(1)} \frac{Z_1 + Z_j}{2}\ N_{x1j}. \end{aligned} \tag{3.15}$$

Again, using the stationary flow condition $h_1 + Z_1 = H$, $h_j + Z_j = H$ and $\rho_1 = C$, one gets

$$h_1 + Z_1 + h_j + Z_j = 2H \qquad \text{and} \qquad \frac{Z_1 + Z_j}{2} = H - \frac{h_1 + h_j}{2}.$$

Thus, (3.15) gives

$$\int_{\mathcal{T}_1} \rho h \frac{\partial Z}{\partial x}\, dV = \rho_1 h_1 \sum_{j\in N(1)} \left(H - \frac{h_1+h_j}{2} \right) N_{x1j}.$$

Using the fact that $\sum_{j\in N(1)} N_{x1j} = 0$,

$$\begin{aligned} \int_{\mathcal{T}_1} \rho h \frac{\partial Z}{\partial x}\, dV &= -\frac{\rho_1 h_1}{2} \sum_{j\in N(1)} h_j\, N_{x1j}, \\ &= -\frac{\rho_1 h_1}{2} \left(h_p N_{x1p} + h_2 N_{x12} + h_3 N_{x13} \right). \end{aligned}$$

A similar procedure leads to the following approximations of the other terms in (3.14)

$$\begin{aligned} \int_{\mathcal{T}_2} \rho h \frac{\partial Z}{\partial x}\, dV &= -\frac{\rho_2 h_2}{2} \left(h_k N_{x2k} + h_1 N_{x21} + h_3 N_{x23} \right), \\ \int_{\mathcal{T}_3} \rho h \frac{\partial Z}{\partial x}\, dV &= -\frac{\rho_3 h_3}{2} \left(h_l N_{x3l} + h_1 N_{x31} + h_2 N_{x32} \right). \end{aligned}$$

Notice that h_p, h_k and h_l are the average values of h respectively, on the triangle $\mathcal{T}_p$, $\mathcal{T}_k$ and $\mathcal{T}_l$, see Figure 3. Summing up, the discretization (3.14) gives

$$\int_{\mathcal{T}_i} \rho h \frac{\partial Z}{\partial x}\, dV = -\frac{\rho_1 h_1}{2} h_p N_{x1p} - \frac{\rho_2 h_2}{2} h_k N_{x2k} - \frac{\rho_3 h_3}{2} h_l N_{x3l}.$$

For this reconstruction, the source terms in (3.13) result in

$$\begin{aligned} \sum_{j\in N(i)} \rho_{ij}^n \left(h_{ij}^n \right)^2 N_{xij} &= \rho_1 h_1 \left(h_p N_{x1p} \right) + \rho_2 h_2 \left(h_k N_{x2k} \right) + \rho_3 h_3 \left(h_l N_{x3l} \right), \\ \sum_{j\in N(i)} \rho_{ij}^n \left(h_{ij}^n \right)^2 N_{yij} &= \rho_1 h_1 \left(h_p N_{y1p} \right) + \rho_2 h_2 \left(h_k N_{y2k} \right) + \rho_3 h_3 \left(h_l N_{y3l} \right). \end{aligned} \tag{3.16}$$

Here, (3.16) forms a linear system of two equations for the three unknowns h_1, h_2 and h_3. To complete the system we add the natural conservation equation

$$h_1 + h_2 + h_3 = 3h_i.$$

Analogously, the bottom values Z_j, $j = 1, 2, 3$ are reconstructed in each sub-triangle of $\mathcal{T}_i$ as

$$Z_j + h_j^n = Z_i + h_i^n, \qquad j = 1, 2, 3.$$

Finally, the source terms in (3.14) are approximated as

$$\begin{aligned} \rho_1 h_1 \int_{\mathcal{T}_1} \frac{\partial Z}{\partial x}\, dV &= \rho_1 h_1 \left(\frac{Z_1+Z_p}{2} N_{x1p} + \frac{Z_1+Z_2}{2} N_{x12} + \frac{Z_1+Z_3}{2} N_{x13} \right), \\ \rho_2 h_2 \int_{\mathcal{T}_1} \frac{\partial Z}{\partial x}\, dV &= \rho_2 h_2 \left(\frac{Z_2+Z_k}{2} N_{x2k} + \frac{Z_2+Z_1}{2} N_{x21} + \frac{Z_2+Z_3}{2} N_{x23} \right), \\ \rho_3 h_3 \int_{\mathcal{T}_1} \frac{\partial Z}{\partial x}\, dV &= \rho_3 h_3 \left(\frac{Z_3+Z_l}{2} N_{x3l} + \frac{Z_3+Z_1}{2} N_{x31} + \frac{Z_3+Z_2}{2} N_{x32} \right), \end{aligned} \tag{3.17}$$

with a similar equation for the other source terms in the y-direction. Numerical results reported in [1] have shown that the above reconstruction exactly preserves the C-property to the machine precision.

The treatment of boundary conditions in the SRNH scheme is performed using similar techniques as those described in [4]. For the computational examples considered in this paper, boundary conditions are enforced on the corrector solution by computing fluxes at boundaries. On the predictor solution and the slopes of dependent variables, boundary conditions are enforced in boundary cells by setting the required variables to the corresponding values of the adjacent inner cells. When slopes are based on vertex values, the solution at boundary vertices is computed by interpolation from two neighboring centroids. When slopes are based on centroid values, the three points used to estimate the slopes are the centroid and the two neighboring centroids inside the computational domain. For further details on the implementation of boundary conditions for the SRNH scheme we refer to [4, 6].

4. Numerical Results

We present numerical results for a class of test problems for dam-break with single and double density discontinuities. In all the computations reported herein, the Courant number Cr is set to 0.8 and the time stepsize Δt is adjusted at each step according to the stability condition

$$\Delta t = Cr \min_{\Gamma_{ij}} \left(\frac{|T_i| + |T_j|}{2\,|\Gamma_{ij}| \max_p |(\lambda_p)_{ij}|} \right),$$

where Γ_{ij} is the edge between two triangles T_i and T_j.

4.1. Dam-break with single density discontinuity

We consider a density dam-break problem with a single initial discontinuity. The problem consists of solving the equations (2.8) in a flat and frictionless channel of length $1000\ m$ and width $500\ m$ filled with two waters with density $\rho_1 = 10\ kg/m^3$ in the left section and $\rho_2 = 1\ kg/m^3$ in the right section *i.e.*,

$$\rho(x, y, 0) = \begin{cases} 10\ kg/m^3, & \text{if} \quad x < 500\ m, \\ 1\ kg/m^3, & \text{otherwise.} \end{cases}$$

Initially, the system is at rest with constant water height $h = 1\ m$. The gravity constant $g = 1\ m/s^2$, the water viscosity $\nu = 1.2 \times 10^{-6}\ m^2/s$ and the sediment viscosity $\kappa = 3.5 \times 10^{-4}\ m^2/s$.

In Figure 4 we display the time evolution of the density, water height, velocity and concentration variables using a mesh with 500 gridpoints. It is clear from these results that at the initial time, the hydrostatic pressure difference at the interface of the two liquids drives a flow of higher density liquid towards the right, pushing the lower density liquid ahead. To conserve mass, the free surface of the lower density liquid rises and a rightward propagating shock-like bore forms. This flow features have been accurately captured by our generalized Rusanov scheme. It should be stressed that

the mechanisms of the density dam-break problems are similar to that of the standard dam-break induced by change in free-surface depth, in that a leftward rarefaction, a rightward shock and a contact wave are formed. Similar wave structures also occur in shock tube gas dynamics.

4.2. Dam-break with double density discontinuity

In this example we solve a density dam-break problem with two initial discontinuities. Here, a flat channel of length $100\ m$ is filled at the left-hand side and right-hand side of the channel with a liquid with density $\rho = 1\ kg/m^3$. At the centre of the channel there is a liquid column of density $\rho = 10\ kg/m^3$ and width of $1\ m$. Initially, the system is at rest with constant water height $h = 1\ m$ and $g = 1\ m/s^2$. The computed results are illustrated in Figure 5 for the t-x phase space. It is evident that the sudden collapse of the denser liquid in the central column causes primary shock waves to be created and propagate as bores in the direction from high to low density. Two outward propagating bores are generated, traveling in opposite directions. Each primary bore decreases in strength with time, which can be seen from the curved shock path. On the other hand, a pair of rarefaction waves travels inward from the interfaces. The rarefaction waves are almost immediately reflected at the center, and then move outward, weakening rapidly. The accuracy of the proposed finite volume is highly achieved in reproducing these physical features.

4.3. Shallow water flow in an idealized inlet system

5. Concluding Remarks

In this paper we have presented a class of finite volume methods for solving two-phase shallow water flows on unstructured triangular grids. The method consists of two stages which can be interpreted as a predictor-corrector procedure. In the first stage, the scheme uses the projected system of the coupled equations and introduces the sign matrix of the flux Jacobian which results in an upwind discretization of the characteristic variables. In the second stage, the solution is updated using the conservative form of the equations and a special treatment of the bed bottom in order to obtain a well-balanced discretization of the flux gradients and the source terms. To increase the accuracy of the scheme we have incorporated slope limiters. Verification of the proposed method has been carried out using test problems of dam-break with single and double density discontinuities. The method exhibited good shape, high accuracy and stability behavior for all hydraulic regimes considered. The presented results demonstrate the capability of the unstructured finite volume method that can provide insight to complex two-phase shallow water flows.

An extension of the proposed finite volume to coupled models of two-phase shallow water flows and bed-load transport in viscous shallow water flows will be the topic of future research. The diffusion effects and tidal waves can be important in many coastal scenarios. Another planned activity will consist in a thorough comparison of different physical models for sediment discharge in the Exner equation.

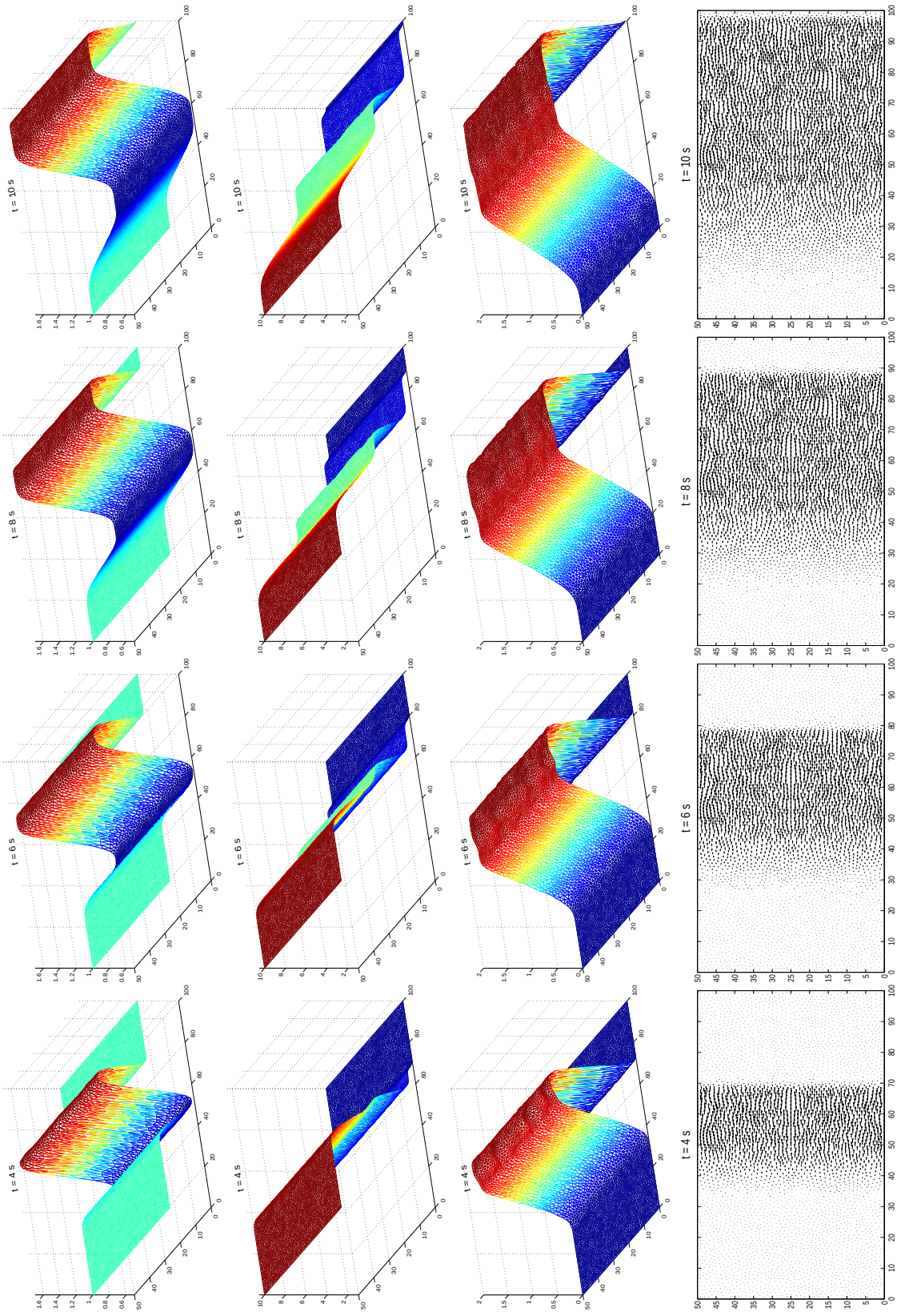


Figure 4: Water height h (first row), water density ρh (second row), x-velocity u (third row) and velocity vectors (fourth row) for dam-break with single density discontinuity at four different simulation times. From left to right $t = 4, 6, 8$ and $10\ s$.

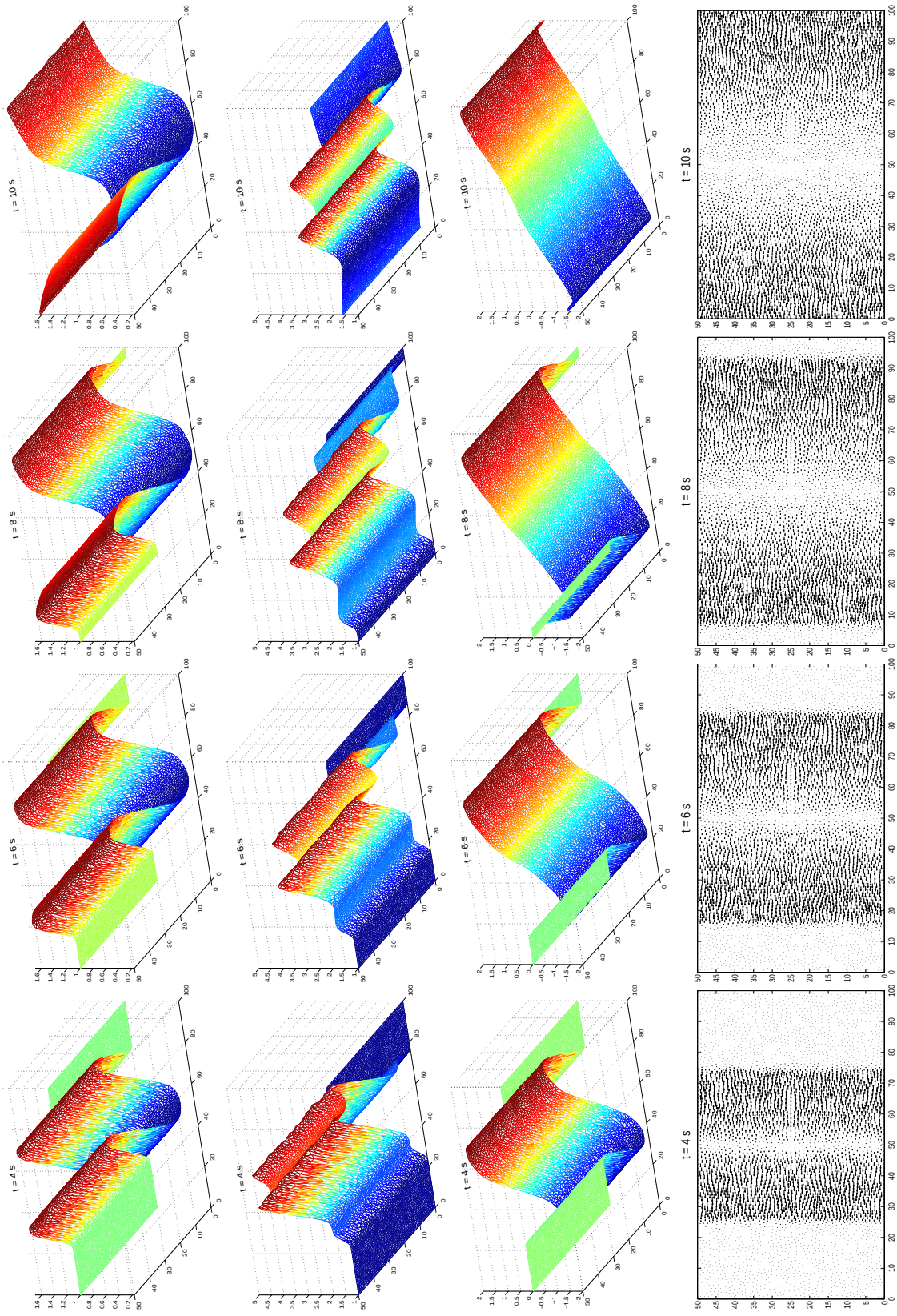


Figure 5: Water height h (first row), water density ρh (second row), x-velocity u (third row) and velocity vectors (fourth row) for dam-break with double density discontinuity at four different simulation times. From left to right $t = 4, 6, 8$ and $10\ s$.

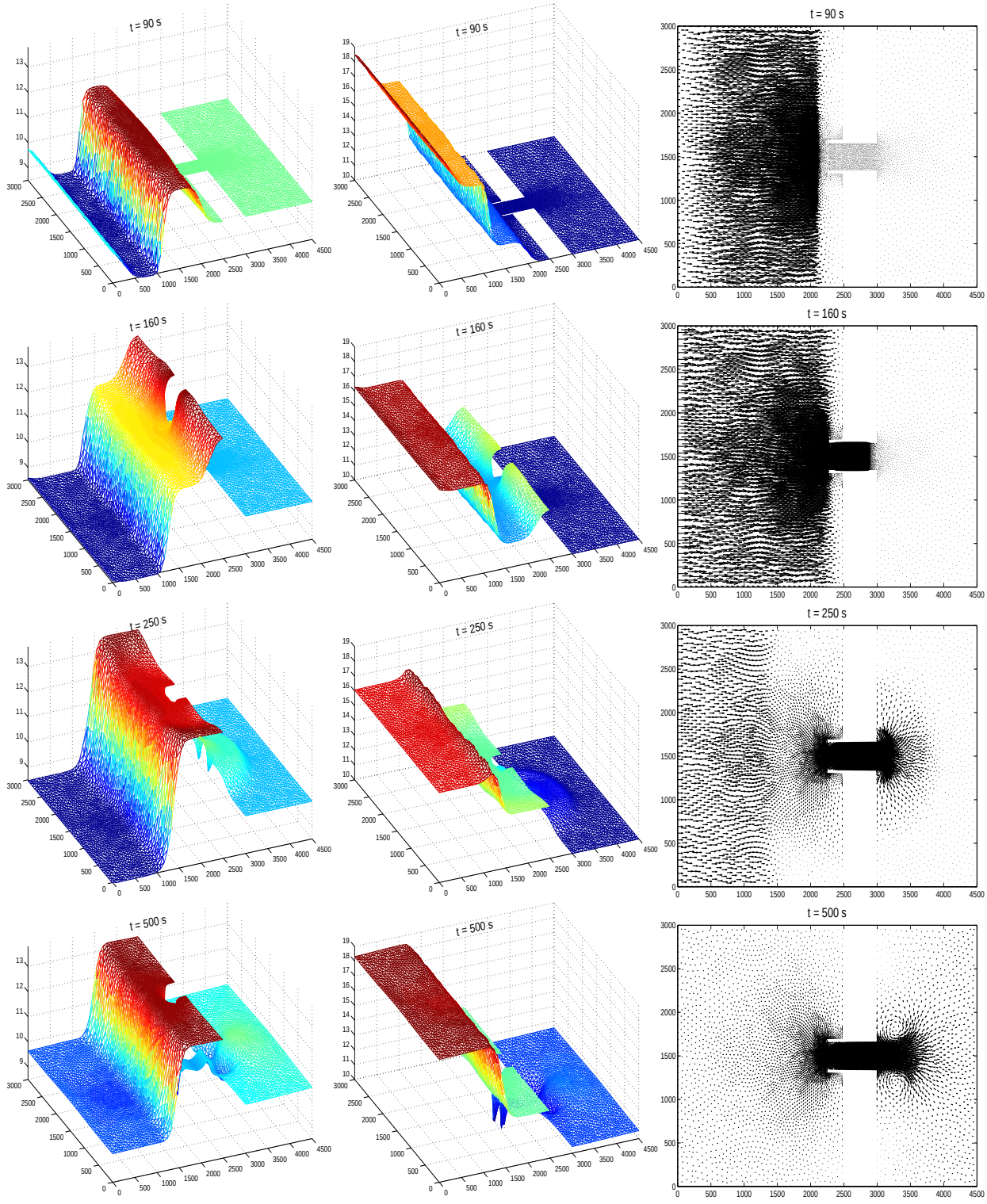


Figure 6: Water height h (first column), water density ρh (second column) and velocity vectors (third column) for the idealized inlet system at four different simulation times. From top to bottom $t = 90$, 160, 250 and 500 s.

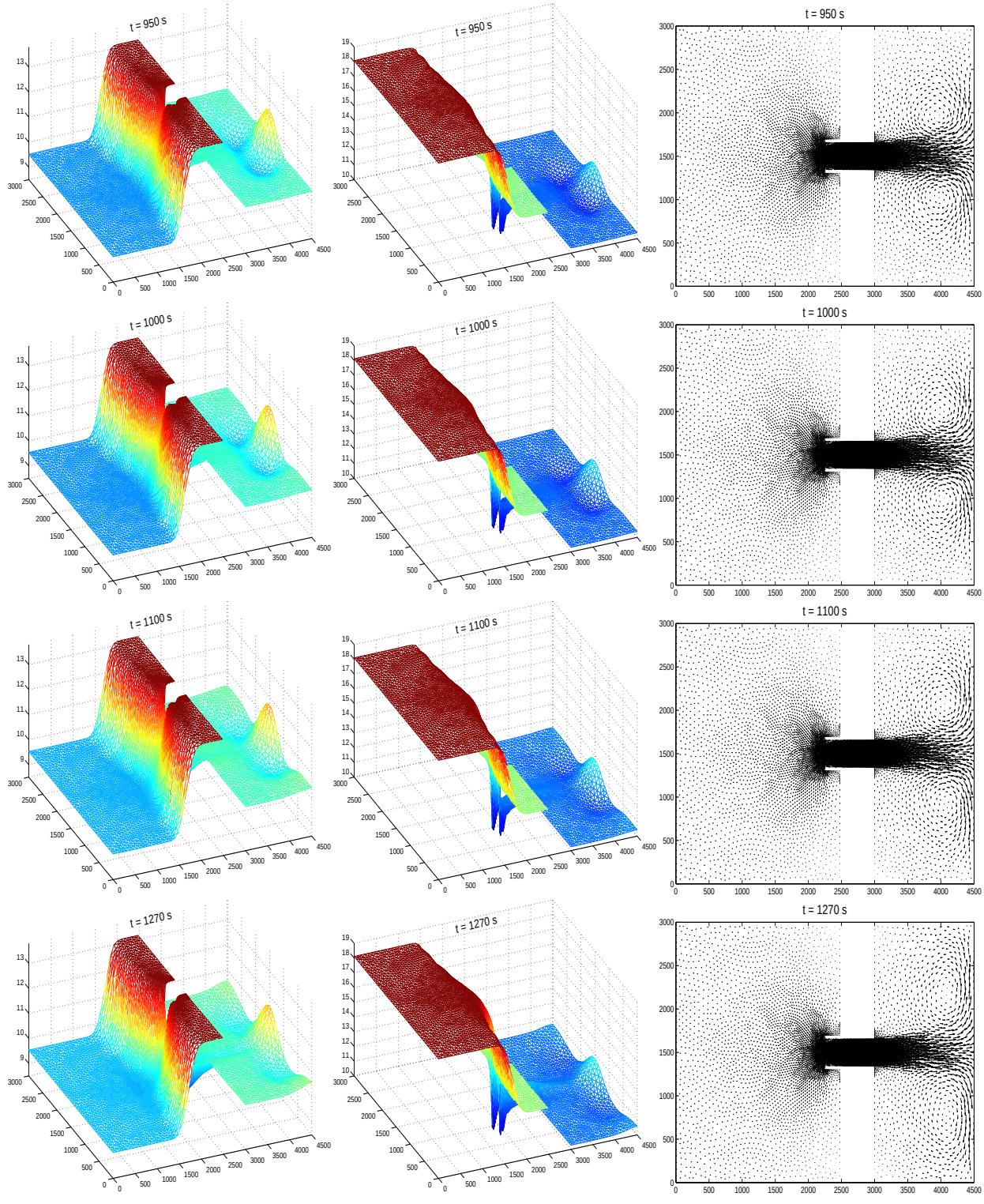


Figure 7: Water height h (first column), water density ρh (second column) and velocity vectors (third column) for the idealized inlet system at four different simulation times. From top to bottom $t = 950$, 1000, 1100 and 1270 s.

Acknowledgements

The work of M. Seaïd was supported in part by Centre National de la Recherche Scientifique (CNRS) under the contract # 209821. Financial support provided by the project Mhycof, Université Paris 13 is also gratefully acknowledged.

References

[1] F. Benkhaldoun, I. Elmahi, M. Seaïd, A new finite volume method for flux-gradient and source-term balancing in shallow water equations, Computer Methods in Applied Mechanics and Engineering, **199**, 49-52 (2010)

[2] F. Benkhaldoun, S. Sahmim, M. Seaïd, A two-dimensional finite volume morphodynamic model on unstructured triangular grids. Int. J. Num. Meth. Fluids. 63 (2010) 1296–1327.

[3] F. Benkhaldoun, S. Sahmim, M. Seaïd, Solution of the sediment transport equations using a finite volume method based on sign matrix. SIAM J. Sci. Comp. 31 (2009) 2866–2889.

[4] F. Benkhaldoun, I. Elmahi, M. Seaïd, "Well-balanced finite volume schemes for pollutant transport by shallow water equations on unstructured meshes", J. Comp. Physics. **226** pp:180-203 (2007).

[5] P.L. Roe, Approximate Riemann Solvers, Parameter Vectors and Difference Schemes, J. Comp. Physics. **43**, 357-372 (1981).

[6] S. Sahmim, F. Benkhaldoun, F. Alcrudo, A Sign Matrix Based Scheme for Quasi-Hyperbolic Non-Homogeneous PDEs with an Analysis of the Convergence Stagnation Problem, J. Comput. Phys. **226**, 1753-1783 (2007).

[7] M.E. Vázquez, Improved Treatment of Source Terms in Upwind Schemes for the Shallow Water Equations in Channels with Irregular Geometry. J. Comp. Physics. **148**, 497–526 (1999).

[8] Leighton, F.Z. Borthwick, A.G.L. Taylor, P.H.: 1-D numerical modelling of shallow flows with variable horizontal density. International Journal for Numerical Methods in Fluids. **62**, 1209–1231 (2010)

[9] Zhou, J.G. Borthwick, A.G.L.: Lattice Boltzmann method for variable density shallow water equations. Computers & Fluids. **49**, 146–149 (2011)

[10] M.B. Abbott, Computational hydraulics: Elements of the theory of free surface flows, Fearon-Pitman Publishers, London, 1979.

[11] S.J. Billett, E.F. Toro, On WAF-Type Schemes for Multidimensional Hyperbolic Conservation Laws, J. Comp. Physics. **130**, 1–24 (1997).

[12] Z. Cao and P. Carling. Mathematical modelling of alluvial rivers: reality and myth. Part I: General overview. Water Maritime Engineering. **154**, 207-220 (2002)

[13] J. Fredsøe, R. Deigaard, Mechanics of Coastal Sediment Transport. *Advanced Series on Ocean Engineering - Vol. 3*, 1992.

[14] G. Simpson and S. Castelltort. Coupled model of surface water flow, sediment transport and morphological evolution. Computers & Geosciences. **32**, 1600-1614 (2006)

[15] G. Strang. On the Construction and the Comparison of Difference Schemes. SIAM J. Numer. Anal. **5**, 506517 (1968)

[16] R.L. Soulsby: Dynamics of marine sands, a manual for practical applications. HR Wallingford, Report SR 466 (1997)

Strictly Real Locally Convex Algebras

Rodia I. Hadjigeorgiou[1] **and M. Oudadess**[2]

[1] Institution Name 1, Department if any, Full address including the country
[2] Ecole Normale Supérieure, Department of Mathematics, B.P. 5118 Takaddoum 10105 Rabat
[3] *Corresponding Author E-mail: oudadessm@yahoo.fr*

Abstract. *Different kinds of strictly real locally convex algebras are considered. We extend, to this more general context, the basic result of I. Kaplansky, in the Banach case, that is the commutativity modulo the Jacobson radical.*
Key words: Strictly real, algebras, m-convex, A-convex, uniformly convex.

1. Introduction

In 1949, I. Kaplansky has introduced [10] the class of strictly real Banach algebras. Since then, not much work has been done on them**.** We intend, in this paper, to examine locally *convex* algebras which are not Banach, but still have the same or similar properties. These are often encountered in different branches of functional analysis, especially in spectral theory.

In Section 3, we show that the (real) multiplier algebra $M(E)$, of a semisimple strictly real Banach algebra, is also strictly real (Proposition 10). This algebra $M(E)$ will be endowed with another topology in Section 6, and serve as an example of a different kind.

Section 4 is devoted to m-*convex* algebras. We obtain Kaplansky's result (as in the Banach case), that is the commutativity modulo the (Jacobson) radical of any complete strictly real $l.m.c.a.$ (Proposition 13). There is also an observation on Michael's problem (Proposition 16). The same results are obtained for locally A-*convex* algebras, in Section 5 (Proposition 22, Proposition 24).

Locally uniformly A-*convex* algebras are examined in Section 6. In that context, Mackey completeness is sufficient to get Kaplansky's result (Proposition 27).

In Section 7, we consider locally uniformly *convex* algebras introduced in [19] (see also [21]). The same results are obtained (Proposition 40), as for locally uniformly A-*convex* algebras. We also provide an example of a such sequentially complete algebra, which has not been done in [19] and [21].

2. Preliminaries

Let (E, τ) be a locally *convex* algebra ($l.c.a.$), with separately continuous multiplication, whose topology τ is given by a family $(p_\lambda)_{\lambda\in\Lambda}$ of seminorms. For simplicity and provided there is nowhere

a risk of confusion, we will write only $(p_\lambda)_\lambda$ instead of $(p_\lambda)_{\lambda\in\Lambda}$. The algebra (E,τ) is said to be locally A-*convex* (l-A-$c.a.$; [4], [5]) if, for every x and every λ, there is $M(x,\lambda)>0$ such that

$$\max\left[p_\lambda(xy),p_\lambda(yx)\right]\leq M(x,\lambda)p_\lambda(y);\forall y\in E.$$

In the case of a single vector space norm, $(E,\|.\|)$ is called an A-normed algebra. If $M(x,\lambda)=M(x)$ depends only on x, we say that (E,τ) is a locally uniformly A-*convex* algebra ($l.u$-A-$c.a.$; [5]). If it happens that, for every λ,

$$p_\lambda(xy)\leq p_\lambda(x)p_\lambda(y);\forall x,y\in E.$$

then (E,τ) is named a locally m-*convex* algebra ($l.m.c.a.$; cf. [12], [11]). Recall also that a $l.c.a.$ has continuous multiplication if, for every λ, there is $\lambda^{'}$such that

$$p_\lambda(xy)\leq p_{\lambda'}(x)p_{\lambda'}(y);\forall x,y\in E.$$

A $l.c.a.$ is said to be locally uniformly *convex* ($l.u.c.a.$; [19], [21]) if

$(\forall x)(\exists M(x)>0)(\forall\lambda)(\exists\lambda^{'}):p_\lambda(xy)\leq M(x)p_{\lambda'}(y);\forall y\in E$, with $\lambda^{'}$depending only on λ but not on x.

An element x of E is said to be bounded [1] (i-bounded in the sense of S. Warner [24]) if there is $\alpha>0$ such that $\{(\alpha x)^n:n=1,2,...\}$ is bounded.

The bounded structure (bornology) of a locally *convex* space ($l.c.s.$) (E,τ) is the collection $\mathbb{B}\tau$ of all the subsets B of E which are bounded in the sense of Kolmogorov-von Neumann, that is B is absorbed by every neighborhood of the origin. If $\tau_{\|.\|}$ is the topology induced by a norm $\|.\|$, we write $\mathbb{B}\tau_{\|.\|}$. We say that a $l.c.s.$ (E,τ) is Mackey complete (M-complete) if its bounded structure $\mathbb{B}\tau$ admits a fundamental system $\mathcal{B}$ of Banach discs that is, for every B in $\mathcal{B}$, the vector space generated by B is a Banach space when endowed with the gauge $\|.\|_B$ of B. For all notions concerning bornology, see [8].

As in the Banach case, a $l.c.a.$ E is said to be *strictly real* if $Sp(x)\subset\mathbb{R}$, for every $x\in E$. Actually, the notion of strict reality can be defined without any reference to a given topology. It is essentially algebraic. A real algebra is said to be strictly real if $Sp(x)\subset\mathbb{R}$, for every $x\in E$, where $Sp(x)$ is the spectrum of $x\in E$ in the complexification $E_{\mathbb{C}}$ of E. The fundamental property is the non voidness of $Sp_{\mathbb{R}}(x)=\{\alpha\in\mathbb{R}:x-\alpha$ is not invertible$\}$. Here are some general facts, already noticed in the Banach case.

Lemma 1. *Let E be a real algebra such that $Sp_{\mathbb{R}}(x)\neq\varnothing$, for every $x\in E$. Then the following are equivalent*

(i) *E is strictly real.*

(ii) *$Sp(x^2)\subset\mathbb{R}_+$, for every $x\in E$.*

(iii) *$e+x^2$ is invertible, for every $x\in E$.*

We give here some examples. Others will be given in different contexts.

Example 2. *Let $E = C_b(\mathbb{R}, \mathbb{R})$ be the algebra of continuous bounded real functions on the real line $\mathbb{R}$. Endow it with the topology given by the family of seminorms $(p_\Phi)_\Phi$, $\Phi \in C_b^0(\mathbb{R}, \mathbb{R})$, where $C_b^0(\mathbb{R}, \mathbb{R})$ is the algebra of functions $\Phi \in E$ vanishing at infinity and*

$$p_\Phi(f) = \sup\{|\Phi(x)|\,|f(x)| : x \in \mathbb{R}\}.$$

It is a $l.u$-A-$c.a.$ which is not a $l.m.c.a.$[4].

Example 3. *Let $E = C(\mathbb{R}, \mathbb{R})$ be the algebra of continuous real valued functions on the real line $\mathbb{R}$. Endow it with the topology of uniform convergence on compacta. This is a $l.m.c.a.$ which is not a $l.u$-A-$c.a.$.*

Example 4. *Considering the standard cartesian product $C_b(\mathbb{R}, \mathbb{R}) \times C(\mathbb{R}, \mathbb{R})$, one obtains a l-A-$c.a$ which is not m-convex nor uniformly A-convex.*

Example 5. *The following example, due to S. Warner [21] in the context of $l.m.c.a.$'s, is a strictly real $l.c.a.$ which is not locally A-convex. Let $E = C[0, 1]$ be the Banach algebra of continuous real functions on the interval $[0, 1]$, endowed with the uniform topology. If E' is the topological dual of E and $\sigma = \sigma(E, E')$ is the associated weak topoloy, then (E, σ) is not a $l.m.c.a.$ ([23], p. 1028). Now, by a result of S. Warner ([23], Theorem 1) and another one of A. C. Cochran ([6], Theorem 3.4), it can not be a l-A-$c.a.$.*

Remark 6. *To enlarge the class of strictly real $l.c.a.$'s, one notices that the property is kept by the following operations: Adjunction of unity, cartesian product, projective limit, the image by an algebra morphism, ...*

3. Banach algebras

This section is devoted to the (real) multiplier algebra $M(E)$, of a commutative strictly real Banach algebra E without order. In the complex case, a notable use is made of the carrier space $\mathcal{M}(E)$. In the real case, the latter can be void. But if we depart with a real algebra such that $\mathcal{M}(E)$ is separating, then there is no obstacle to reproduce the arguments in [22]. Indeed, they are of an algebraic or a general topological character. No appeal is made to complex analysis. So, one can show here that the (real) multiplier algebra of $M(E)$, of E, is also strictly real. But first a concrete example. Recall that an algebra E is said to be without order if $x = 0$ whenever $xy = 0$ for every y. A linear map S from E to E is named a multiplier if $S(xy) = xS(y)$, for every x and y in E. All multipliers are automatically continuous, and the algebra $M(E)$ is commutative [22].

Example 7. *Let $E = C_0(\mathbb{R}, \mathbb{R})$ be the algebra of continuous real valued functions on the real line $\mathbb{R}$, vanishing at infinity. Endowed with the topology of the supremum norm $\|.\|_\infty$, it becomes a Banach algebra. The evaluations χ_x, on E, given by $\chi_x(f) = f(x)$, are characters of E. Hence $\mathcal{M}(E)$ is separating. Actually, E is a semisimple strictly real Banach algebra E without order.*

Remark 8. *Though not used in this paper, we give the following simple but interesting facts.*

(1) The kernel $\ker T$ *, of a multiplier* T *of* E*, is a bilateral ideal. This is due to* $T(xy) = T(x)y$*, for every* x *and every* y*.*

(2) If $(E, \|.\|)$ *is a strictly real Banach algebra and* T *is a multiplier* E*, then* $T(Rad(E)) \subset Rad(E)$*. Indeed, take* $x \in Rad(E)$*. One has to show that* $\varrho(T(x)y) = 0$*, for every* y*. But* $\varrho(T(x)y) = \varrho(xT(y))$*. Now* E *being strictly real (hence commutative modulo its radical) , one has* $\varrho(xT(y)) \leq \varrho(x)\varrho(T(y))$*, with* $\varrho(x) = 0$*.*

Remark 9. *We now give a simple property, we will use in the sequel. It is general and has an interest in its own. If* E *is a real Banach algebra, then* $E/Rad(E)$ *is without order. Indeed if* $\overline{x} \in E/Rad(E)$ *is such that* $\overline{xy} = \overline{xy} = 0$ *for every* $\overline{y}$*, then* $xy \in Rad(E)$*, for every* y*. Whence* $x \in Rad(E)$*. So* $\overline{x} = \overline{0}$*.*

Proposition 10. *Let* $(E, \|.\|)$ *be a strictly real Banach algebra. Then*

(i) The quotient algebra $M(E)/Rad(E)$ *is also strictly real.*

(ii) If moreover E *is semisimple, then* $M(E)$ *is strictly real.*

Proof. **(i)** The algebra $F = E/Rad(E)$ is commutative [10]. So it is a semisimple commutative Banach algebra. It is also without order (Proposition 9). Now, exactly as in [18], every multiplier $T \in M(F)$ can be identified with a continuous bounded real valued function on the carrier space $\mathcal{M}(F)$ of F. So $SpT = Spg \subset \mathbb{R}$.

(ii) Immediat from (i). □

Remark 11. *By a general and simple construction, we can obtain non commutative strictly real Banach algebras from commutative ones. Let* E *be a strictly real commutative Banach algebra that admits a unital isometric morphism* T *(i.e., a morphism such that* $\|T(x)\| = \|x\|$*, for every* x*) which is not a multiplier. Endow* E *with the multiplication* $\times_T$ *given by* $x \times_T y = xT(y)$*. It is not commutative. One easily checks that* E *remains strictly real.*

We do not know , in general, how strict reality behaves concerning completion. We can say the following.

Proposition 12. *Let* $(E, \|.\|)$ *be a strictly real normed algebra. If* E *is commutative, then the completion* $\widehat{E}$ *of* E *is also strictly real.*

Proof. Let $x \in \widehat{E}$, with $x = \lim x_n$, $x_n \in E$. One has $Sp(x) = \{\chi(x) : \chi \in \mathcal{M}(\widehat{E})\}$ and $\chi(x) = \lim \chi(x_n)$. Since, by hypothesis, every $\chi(x_n) \in \mathbb{R}$, one has $\chi(x)$ for every $\chi \in \mathcal{M}(\widehat{E})$. □

4. Locally m-convex algebras

First, let us recall the Arens-Michael decomposition [12] of a complete $l.m.c.a.(E,(|.|_\lambda)_\lambda)$. Put $N_\lambda = \{x \in E : |x|_\lambda = 0\}$ and $E_\lambda = E/N_\lambda$ endowed with the algebra norm $\|.\|_\lambda$ given by $\|x + N_\lambda\|_\lambda = |x|_\lambda$. The norm in the completion $E_\lambda^\wedge$ of E_λ is still denoted $\|.\|_\lambda$. Then

$$(E,(|.|_\lambda)_\lambda) = \varprojlim(E_\lambda^\wedge, \|.\|_\lambda).$$

Denote by π_λ the canonical surjection $x \longrightarrow \overline{x} = x + N_\lambda$, from E to $E_\lambda^\wedge$.

Proposition 13. *Let $(E,(|.|_\lambda)_\lambda)$ be a real complete $l.m.c.a.$. Then*

(i) E is strictly real if, and only if, every factor $E_\lambda^\wedge$ is a strictly real Banach algebra.

(ii) if E is strictly real, then it is commutative modulo its (Jacobson) radical $Rad(E)$.

Proof. (i) If $x = (x_\lambda)_\lambda \in E$, then $Sp(x) = \cup_\lambda Sp(x_\lambda)$. Whence the claim.

(ii) If $x = (x_\lambda)_\lambda \in E$ and $y = (y_\lambda)_\lambda \in E$, then $xy - yx = (x_\lambda y_\lambda - y_\lambda x_\lambda)_\lambda$. Now, since $E_\lambda^\wedge$ is a strictly real Banach algebra, one has $x_\lambda y_\lambda - y_\lambda x_\lambda \in Rad(E_\lambda^\wedge)$, by a result of I. Kaplansky [6]. But ([8], Proposition 7.3, p. 29),

$$Rad(E) = \cap_\lambda \pi_\lambda^{-1}(Rad(E_\lambda^\wedge).$$

Whence $xy - yx \in Rad(E)$. □

We have an application to topological inductive limits of Banach algebras. But first, a general result. Recall that $(E,\tau) = \varinjlim (E_i, \|.\|_i)$ is a topological strict inductive limit of Banach algebras if the restriction of the topology τ to each E_i is exactly $\tau_{\|.\|_i}$ given by the norm $\|.\|_i$.

Proposition 14. *Let $(E,\tau) = \varinjlim (E_i, \|.\|_i)$ be a topological strict inductive limit of Banach algebras. If the set I of indices is well ordered, then E is strictly real.*

Proposition 15. *Let $(E,\tau) = \varinjlim (E_n, \|.\|_n)$ be a topological strict inductive limit of a sequence of strictly real Banach algebras. Then E is commutative modulo its (Jacobson) radical $Rad(E)$.*

Proof. By the previous proposition, it is strictly real. By a result of Arosio **[3]** it is a $l.m.c.a.$. Then use (ii) of Proposition 13. □

Another consequence is in relation with Michael's problem [12]. Using Do Sin-Sya's technique, T. Husain has shown [9] that every (real) character of a sequentially complete $l.m.c.a.$ is bounded, when such character exists. Now Proposition 13 ensures that existence in our context. So, it is worthwhile to state the following. Observe also that the complex case is included.

Proposition 16. *Let (E,τ) be a complete strictly real $l.m.c.a.$. Then*

(i) Every (real) character of E is bounded.

(ii) Every (real) character of the complexification $E_{\mathbb{C}}$, of E, is bounded.

Proof. (i) cf. [9] or [13].
(ii) Actually, there is a bijection between $\mathcal{M}(E)$ and $\mathcal{M}(E_{\mathbb{C}})$. □

Example 17. *$C([0,1],\mathbb{R})$ endowed with the topology of uniform convergence on denumerable compact subsets of $[0,1]$ is a complete $l.m.c.a.$. Actually it is a locally C^*-algebra. So it can not be a Q-algebra for it should then be normable.*

Example 18. *Let Ω be the first uncountable ordinal and endow $[0,\Omega]$ with the order topology. Then the topological algebra $(C[0,\Omega],\mathbb{R})$, endowed with the compact-open topology, is a unital commutative and complete $l.m.c.a.$*

To let Proposition 13 be meaningful, one has to exhibit non commutative strictly real $l.m.c.a.$'s.

Example 19. *Take F the real Banach algebra of $n \times n$ upper triangular matrices $M = (a_{ij})$ such that $a_{ij} \in \mathbb{C}$ for $i \neq j$ and $a_{ii} \in \mathbb{R}$, endowed with the norm $\|.\|$ given by $\|M\| = \max_i \Sigma_j |a_{ij}|$. Consider the standard cartesian product $E = \Pi_n F_n$, where $n \in \mathbb{N}^*$ and $F_n = F$ for every n. Then (E,τ), τ being the product topology, is a non commutative strictly real Fréchet $l.m.c.a.$*

Example 20. *Take the algebra F in the previous example. Put $E_n = F \times ... \times F$ (n times) and $E = \varprojlim(E_n,(f_{nm})_{n,m})$, where $f_{nm} : E_n \longrightarrow E_m$ are the standard projections, for $m \leq n$. Then (E,τ), τ being the projective topology, is a non commutative strictly real Fréchet $l.m.c.a.$*

Remark 21. *To obtain more examples, one can use Remark 3.5, stability properties of the m-convex structure and Proposition 13.*

5. Locally A-convex algebras

These algebras have been introduced by A.C. Cochran, R. Keown and C.R. Williams [4]. Their study is actually reduced to *m-convex* ones [10] (see also [18]). As in the complex case, there is an *m-convex* topology $M(\tau)$ stronger than τ. If the latter is given by a family $(p_\lambda)_\lambda$ of seminorms (cf. Preliminaries), one may suppose that $p_\lambda(e) = 1$ for every λ, where e is the unit element. Then one has $p_\lambda(x) \leq M(\lambda,x)$, for every λ. Put $q_\lambda(x) = \sup\{\, p_\lambda(xu) : p_\lambda(u) \leq 1\}$. The topology $M(\tau)$ is given by the family $(q_\lambda)_\lambda$ of submultiplicative seminorms. Moreover, it is complete if τ is. So, using (ii) of Proposition 13, one obtains the following.

Proposition 22. *A strictly real complete l-A-$c.a.$ $(E,(|.|_\lambda)_\lambda)$ is commutative modulo its (Jacobson) radical $Rad(E)$.*

Remark 23. *Recall that an A-normed algebra is complete if and only if it is a (Banach) normed algebra. It is known that a non m-convex l-A-$c.a.$ $(E, (|.|_\lambda)_\lambda)$ is isomorphic to a subalgebra of the cartesian product of the A-normed algebra's $E_\lambda = E/N_\lambda$, where $N_\lambda = \{x \in E : |x|_\lambda = 0\}$. But the E_λ's are not complete, even if $(E, (|.|_\lambda)_\lambda)$ is, for otherwise the latter should be m-convex. So one can not apply Kaplansky's result. It is the existence of the stronger m-convex toplogy $M(\tau)$ that allows it.*

Concerning Michael's problem, one has also a positive answer.

Proposition 24. *Let (E, τ) be a complete strictly real l-A-$c.a$. Then*

(i) Every (real) character of E is bounded.

(ii) Every (real) character of the complexification $E_\mathbb{C}$, of E, is bounded.

Proof. (i) Since (E, τ) is complete, the topologies τ and $M(\tau)$ have the same bounded sets [20]. Then use Proposition 16.

(ii) By (ii) of Proposition 16. □

Example 25. *Let E be a strictly real $l.u$-A-$c.a$. that is not a $l.m.c.a$ and F a strictly real $l.m.c.a$ which is not uniformly A-convex. Then the standard cartesian product algebra $E \times F$ is a strictly real l-A-$c.a$. which is not m-convex, nor uniformly A-convex. Take e.g., $E = C_b(\mathbb{R}, \mathbb{R})$ of Example 2 and $F = C(\mathbb{R}, \mathbb{R})$ of Example 3.*

Example 26. *To have non commutative strictly real l-A-$c.a$.'s, take a non commutative strictly real $l.u$-A-$c.a$. E that is not m-convex and F a non commutative (or even commutative) strictly real m-convex algebra which is not uniformly A-convex. Then $E \times F$ is a non commutative strictly real l-A-$c.a$., which is not m-convex nor uniformly A-convex.*

6. Locally uniformly A-convex algebras

The approach is structural. Let (E, τ) be a unital $l.u$-A-$c.a$. As in the complex case, there is an algebra norm $\|.\|_0$ stronger than τ (cf. [14], [15], [16]). If the latter is given by a family $(p_\lambda)_\lambda$ of seminorms (cf. Preliminaries), one may suppose that $p_\lambda(e) = 1$ for every λ, where e is the unit element. Then one has $p_\lambda(x) \leq M(x)$, for every λ. Just put $\|x\|_0 = \sup\{ p_\lambda(x) : \lambda\}$. Moreover, $(E, \|.\|_0)$ is a Banach algebra if (E, τ) is M-complete.

Proposition 27. *Let $(E, (|.|_\lambda)_\lambda)$ be a strictly real $l.u$-A-$c.a$.*

(i) If $(E, (|.|_\lambda)_\lambda)$ is unital and M-complete, then E is commutative modulo its (Jacobson) radical $Rad(E)$.

(ii) If $(E, (|.|_\lambda)_\lambda)$ is non unital and complete, then E is commutative modulo its (Jacobson) radical $Rad(E)$.

Proof. **(i)** If E is unital, then $(E, \|.\|_0)$ is a Banach algebra. So, here one uses directly the result of I. Kaplansky.

(ii) If E is not unital, the unitization E_1 of E a l-A-$c.a.$, one can use Proposition 22. □

Remark 28. *The unitization of a $l.u$-A-$c.a.$is not always a $l.u$-A-$c.a.$ (cf. [18]).*

Concrete examples to which the previous proposition applies are the following.

Example 29. *Let $C_b(R,R)$ be the algebra of real continuous bounded functions on the real field R with the usual pointwise operations. Denote by $C_0^+(R,R)$ the strictly positive elements of $C_b(R,R)$. Consider the family $\{p_\varphi : \varphi \in C_0^+(R)\}$ of seminorms given by*

$$p_\varphi(f) = \sup\{|f(x)\varphi(x)| : x \in \mathbb{R}\}; f \in C_b(\mathbb{R}).$$

They determine a locally convex topology β. The space $(C_b(R,R), \beta)$ is a complete locally convex algebra. It is not a $l.m.c.a.$ [4]. But it is a $l.u$-A-$c.a.$ with continuous multiplication.

Example 30. *Consider the multiplier algebra $M(E)$ of a semisimple strictly real Banach algebra (see Section 3). Endow it with the strong operator topology, given by the family $(|.|_x)_x$ of seminorms defined by $|T|_x = \|T(x)\|$. As in the complex case [18], $\big(M(E), (|.|_x)_x\big)$ is a complete $l.u$-A-$c.a.$, the asssociated norm $\|x\|_0$ of which is the operator norm.*

To have non unital strictly real $l.u$-A-$c.a.$'s, one can use the following general construction.

Example 31. *Take a non unital strictly real Banach algebra E. Put $E_1 = E$, $E_n = E \times ... \times E$ (n times) and $F = \cup E_n$. Then clearly $F = \varinjlim(E_n, \|.\|_n)$, the transmission maps being the canonical injections $f_{nm} : E_m \longrightarrow E_n$, for $m \leq n$. Since E_m is closed in E_n, one has $Sp(x) = Sp_{E_n}(x)$, for every n such that $x \in E_n$. Thus E is strictly real, as in Proposition 15.*

We can use the idea of Remark 11, to obtain non commutative strictly $l.u$-A-$c.a.$'s.

Example 32. *Consider a multiplier algebra $M(E)$ as in Section 3. Endow it with a multiplication $\times_T$ as in Remark 11, that is $P \times_T Q = PT(Q)$ and the strong topology β given by the family $(|.|_x)_x$, $x \in E$, wher $|S|_x = \|S(x)\|$. One has*

$$|P \times_T Q|_x = \|PT(Q)(x)\| = \|T(Q)P(x)\| \leq \|T(Q)\| \, |P|_x, \forall P.$$

And also

$$|P \times_T Q|_x = \|PT(Q)(x)\| \leq \|P\| \, \|T(Q)(x)\| = \|P\| \, \|Q(x)\| = \|P\| \, |Q|_x, \forall Q.$$

So $M(E)$ becomes a non commutative strictly real $l.u$-A-$c.a.$'s.

7. Locally uniformly convex algebras

Such complex algebras have been introduced in [21] (see also [19]). The definition remains the same (see Preliminaries). We will exhibit examples of $l.u.c.a.$'s which are not uniformly A-*convex*. But examine first the subnormability of such algebras. A locally *convex* algebra (E,τ) is said to be subnormable if it can be endowed with a vector space norm which induces a topology stronger than τ. Exactly as in the complex case [21] one shows some basic facts.

Proposition 33. *(i) A unital $l.u.c.a.$ is subnormable. If moreover it is barrelled, then it is normable.*

(ii) A non unital $l.u.c.a.$ is subnormable if, and only if, its unitization E_1 is also a $l.u.c.a.$

Remark 34. *The norm in (i) of the previous proposition is a vector space but not necessarily an algebra norm (Example 36 below). This example shows also that, unlike the case of $l.u$-A-$c.a.$'s, not every element here is necessarily bounded.*

Under an additional completeness condition, one can say more.

Proposition 35. *([21], [19]). Let (E,τ) be a unital and Mackey complete $l.u.c.a.$ Then there is a Banach algebra norm stronger than τ with the same bound structure.*

The first example, adapted here to the real case, has been given in [21].

Example 36. *Let $\mathbb{R}[X]$ be the algebra of real polynomials and $(x_m)_m$ a sequence of real numbers such that $|x_m| \longrightarrow +\infty$. Endow it with the topology τ given by the seminorms $P \mapsto |P|_m = |P(x_m)|$. It becomes a unital commutative and metrizable $l.m.c.a.$ Moreover it has a denumerable algebraic basis, so it is subnormable (cf. a Lemma in [7], p. 1039). Let $\|.\|$ be a vector space norm stronger than τ. Then, for every m, there is a $k_m > 1$ such that*

$$|PQ|_m = |P|_m \, |Q|_m \leq \|P\| \, k_m \, |Q|_m \, ; \forall Q.$$

To see the local uniform convexity of $\mathbb{R}[X]$, consider the family $(|.|_m)_m \cup (\alpha \, |.|_m)_{m,\alpha \geq 1}$ of seminorms which also define the topology. It can not be uniformly A-convex, for otherwise we should have a stronger algebra norm $\|.\|_0$ than τ. But then the characters $P \mapsto P(x_m)$ should be continuous for that norm, which contradicts $|P(x_m)| \longrightarrow +\infty$.

In the previous example, the algebra is not sequentially complete. Here we give an example where it is.

Example 37. *Let $E = C^1[0,1]$ be the algebra of $\mathbb{K}$-valued ($\mathbb{K} = \mathbb{R}, \mathbb{C}$) continuous functions, on $[0,1]$, with continuous derivative also at extreme points. endow it with the topology τ given by the seminorms $\|f\|_\infty = \sup\{|f(t)| : 0 \leq t \leq 1\}$ and $|f|_{K_d} = \sup\{\left|f^{'}(t)\right| : t \in K_d\}$, where K_d runs over denumerable compact subsets of $[0,1]$. Then (E,τ) is a unital commutative sequentially complete $l.m.c.a.$. It is a $l.u.c.a.$ Indeed,*

$$\begin{aligned} |fg|_{K_d} &= \sup\{\left|f^{'}(t)g(t) + f(t)g^{'}(t)\right| : t \in K_d\} \\ &\leq M(f)\left[|g|_{K_d} + |g^{'}|_{K_d}\right], \text{ with } M(f) = \max\left(\|f\|_\infty, \|f^{'}\|_\infty\right). \end{aligned}$$

It can not be a $l.u$-A-$c.a.$ for otherwise, there should be an algebra norm $\|.\|_0$ stronger than τ, which is moreover the coarsest norm with such properties (see [15] or [16]). But then

$$\tau_{\|.\|_\infty} \preceq \tau \preceq \tau_{\|.\|_0} \text{ and } \tau_{\|.\|_0} \preceq \tau_{\|.\|_\infty}.$$

So τ should be equivalent to $\tau_{\|.\|_0}$, which is not the case. Here the norm of Proposition 33 is $\|.\|_\infty + \|.\|'_\infty$, where $\|f\|'_\infty = \|f'\|_\infty$.

Example 38. *Consider $(L^\omega[0,1], (\|.\|_n)_n)$ the real Arens algebra [2], where $\|f\|_n = \left(\int_0^1 |f(t)|^n\, dt\right)^{\frac{1}{n}}$, $n \in \mathbb{N}^*$. It is known that $C[0,1]$ endowed with the induced topology is dense in $L^\omega[0,1]$. So here $C[0,1]$ is a $l.u$-A-$c.a.$(hence a $l.u.c.a.$) the completion of which is an algebra. But, it is not of the same type, since as metrizable and complete (a B_0-algebra) it should be normable (Proposition 7.1). So $(L^\omega[0,1], (\|.\|_n)_n)$ is a strictly real Fréchet $l.c.a$. It can not be A-convex, nor uniformly convex.*

Example 39. *$C^1[0,1] \times C_b(\mathbb{R},\mathbb{R})$ is a commutative strictly real $l.u.c.a.$, which is not uniformly A-convex nor m-convex. It is a l-A-$c.a.$.*

Now using Proposition 35 and Kaplansky's result, one obtains the following.

Proposition 40. *A unital strictly real M-complete $l.u.c.a.$ $(E, (|.|_\lambda)_\lambda)$ is commutative modulo its (Jacobson) radical $Rad(E)$.*

We can use the idea of Remark 11, to obtain non commutative strictly real $l.u.c.a.$'s. Actually one gets real or complex algebras, which is a noticeable fact.

Example 41. *Consider a multiplier $M_T(E)$ as in Example 32, endowed with the multiplication $\times_T$ and the strong topology β given by the family $(|.|_x)_x$, $x \in E$, wher $|S|_x = \|S(x)\|$. It is a non commutative strictly $l.u$-A-$c.a.$'s. Take the algebra $C^1[0,1]$ of Example 37. Then $M_T(E) \times C^1[0,1]$ is non commutative strictly $l.u.c.a.$. It is not a $l.u$-A-$c.a.$, but it is locally A-convex.*

Example 42. *One may also consider $M_T(E) \times C_b(R,R)$ or even $M_T(E) \times C^1[0,1] \times C_b(R,R)$.*

8. Conclusion

I. Kaplansky introduced strictly real Banach algebras, and showed that they are commutative modulo their Jacobson radical. In this paper, we extend this result to classes of locally convex (non normable) algebras. First, it is worthwhile to notice that the multiplier algebra $M(E)$ of a strictly real Banach algebra is also strictly real. Kaplansky's result is obtained for m-convex algebras, and a remark is made about Michaël problem. The same for A-convex algebras. Finally, we treat the case of uniformly convex algebras.

Acknowledgements

The authors offer their thanks to the referee for a careful checking of the manuscript and valuable remarks.

References

[1] G. R. Allan**,** *A spectral theory for locally convex algebras*, Proc. London Math. Soc. (3), 15 (1965), 399–421.

[2] R. Arens**,** *The space L^{ω} and convex topological rings*, Bull. Amer. Math.Soc. 52 (1946), 913-935.

[3] A. Arosio**,** *Locally convex inductive limits of normed algebras*, Rend. Sem.Mat. Univ. Padova 51 (1974), 331–359.

[4] A. C. Cochran, R. Keown, C. R. Williams**,** *On a class of topological algebras*, Pacific J. Math. 34 (1970), 17-25.

[5] A. C. Cochran**,** *Representation of A-convex* algebras, Proc. Amer. Math. Soc. 41 (1973), 473-479.

[6] A. C. Cochran**,** *Weak A-convex algebras*, Proc. Amer. Math. Soc. 26 (1970), 73–77.

[7] J. Esterle**,** *Sur la non normabilité de certaines algèbres d'opérateurs,* C. R. Acad. Sc. Paris, t. 278, Série A (1974), 1037-1040.

[8] H. Hogbé–Nlend, *Théorie des bornologies et applications*, Springer Lectures Notes, 213, (1971).

[9] T. Husain**,** *Multiplicative functionals on topological algebras*, Pitman research notes Math. Series 85, Longman, 1983.

[10] I. Kaplansky**,** *Normed Algebras*, Duke Math. J. 16 (1949), 399-418.

[11] A. Mallios**,** *Topological algebras. Selected topics*, North-Holland, Amsterdam, 1986.

[12] E. A. Michael**,** *Locally multiplicatively convex topological algebras,* Mem. Amer. Math. Soc. 11 (1952).

[13] M. Oudadess**,** *Michael's problem on M-convexity and Do Sin-Sya's technique*, Southeast Asian Bull. Math. (2005) 29: 1113-1116.

[14] M. Oudadess**,** *Théorèmes de structures et propriétés fondamentales des algèbres localement uniformémnt A-convex*es, C. R. Acad. Sc. Paris, t. 296, Série J (1983), 851-853.

[15] M. Oudadess**,** *Une norme d'algèbre de Banach dans les algèbres uniformément A-convex*es, Africa Math. 9 (1987), 15-22.

[16] M. Oudadess**,** *Théorème du type Gelfand-Naimark dans les algèbres uniformément A-convexes*, Ann. Sc. Math. Québec 9 (1985), 73-82.

[17] M. Oudadess**,** *Discontinuity of the multiplication in multiplier algebras*, Publications Mathématiques 34 (1990), 397-401.

[18] M. Oudadess**,** *Remarks on locally A-convex algebras*, Bull. Greek Math. Soc. 56 (2009), 47-55.

[19] M. Oudadess**,** *Locally uniformly convex algebras*, Bull. Greec Math. Soc. (to appear).

[20] M. Oudadess**,** *Functional boundeness of some M-complete m-convex* algebras, Bull. Greec Math. Soc. 39 (1997), 17-20.

[21] M. Oudadess**,** *Bounded structures in locally A-convex algebras*, Math. Studies 4, Estonian Math. Soc., (2008), 80-88.

[22] Ju-Kwei Wang**,** *Multipliers of commutative Banach algebras*, Pacific J. Math. 11 (1961), 1131–1149.

[23] S. Warner**,** *Weakly topologized algebras*, Proc. Amer. Math. Soc. 8 (1957), 314-16.

[24] S.Warner**,** *Inductive limits of normed algebras*,Trans. Amer. Math. Soc. 82 (1956), 190–216.

Positive Feedback Trading and Stock Return Autocorrelation: The Case of Morocco

M. Achkir and A. Charafi

School of Business Administration, Al Akhawayn University, PO Box 104, Hassan II Avenue, Ifrane 53000 Morocco

Corresponding Author E-mail: a.charafi@aui.ma

Abstract. *This paper investigates the presence of positive feedback trading in the Casablanca stock exchange and measures the profitability and the effectiveness of selected herding strategies. The MADEX returns from 2004 to 2010 are analyzed, modeled, and forecasted for that purpose using linear autoregressive models, GARCH processes, and E-GARCH processes. Relying on the Sentana and Wadhwani's positive feedback model, this paper explores the link between feedback trading, serial autocorrelations, and volatility. It presents supporting evidence on the persistence of serial autocorrelations in the index returns suggesting the prevailing influence of feedback trading activity on return dynamics. The signaling-based simulation results reveal that herd trading dominates the simple buy and hold strategy and the smart money investors' strategy both on daily and weekly bases. The results also unveil the impact of the day of trade on weekly trading outcomes, volatilities, and Sharpe ratios.*

Key words: positive feedback trading, autocorrelation, GARCH.

1. Introduction

Throughout the last decade, there has been growing interest in the emerging and developing stock exchanges and the investment opportunities present in the countries that host them. These markets are referred to as "frontier markets"; and are characterized by small market capitalizations, low liquidity levels, and imminent privatization trends. In addition, these economies are distinguished by the constrained impact of international events on their prosperity and progress. The reason for this is that the stock exchanges in such countries list local companies that have limited ties at the international level. They are also hosted in countries where the restrictive regulation confines the impact of external shocks and crises on the overall economy. Thus, the stock returns in these markets often

exhibit negative correlations with the more developed ones. These various factors highlight risk reduction prospects that attract long term investors that are interested in the diversification opportunities offered by such economic environments.

The attractiveness of the frontier markets sheds light on their well-functioning and raises concerns about their efficiency. The question whether the prices incorporate information instantly and correctly or whether the returns exhibit a random walk or not is the prime concern of agents involved in these markets. Answers to these questions would give an indication about the predictability of the markets and about the manner news releases disseminate into stock prices. Unveiling the nature of the trading conducts in those markets would also allow a better grasp of how the prices are influenced and how they are impacted by the various trading strategies.

In a broad sense, one can consider the trading conduct as being governed by two schools; the fundamental analysis and the technical analysis. The fundamental analysis' adepts engage in trading with a solid knowledge of the companies' standing and in-depth studies of financial information and industry settings. They are referred to as "rational investors" or "smart money investors" (SMI) since they rely on forecasting techniques and valuation models drawn from historical data incorporating various economic factors. The second method of trading; the technical analysis, is based on examining prices and volumes of strictly historical data. The practitioners of this type of trading are called trend chasers, or rational speculators. They trade with the ultimate conviction that the trends' history repeats itself and are believed to influence the market movements.

One of the possible ways in which the technical analysts' behavior could affect share prices is through feedback trading. Feedback traders are trend followers who base their strategy on price movements. Positive (negative) feedback traders buy (sell) when prices rise and sell (buy) when prices fall (rise). Hence, if their presence is significant in a market they can induce serial autocorrelations of the returns. Such behavior was first documented by Cutler, Poterba, & Summers (1990) as they revealed the significant presence of serial autocorrelation in US stock returns. The empirical conclusions that were reached by the aforementioned authors were a breakthrough in behavioral finance and triggered the curiosity of other scholars that were more interested in developing a theoretical framework for such market interactions.

The theoretical aspect of the herding behavior was first dealt with in Shiller's (1984) work when he documented evidence of overreaction following dividend announcements in the US and attributed this phenomenon to social trends. Shiller (1984) builds a model to capture such behavior which was improved upon by Sentana and Wadhwani (1992) to become the positive feedback trading model. Empirical studies based on that model succeed in finding evidence of positive feedback trading in developed markets as well as in emerging markets. Koutmos,1997; and Koutmos and Saidi, 2001 among others show how positive feedback traders induce negative autocorrelations in returns in the US and emerging Asian markets respectively.

This paper complements the positive feedback trading studies by investigating the presence and effectiveness of positive feedback trading in the Casablanca stock exchange (CSE). It also compares trading gains of feedback trading to other chosen strategies.

Daily data of the Casablanca stock exchange over the 2001-2010 period is used for that purpose. The theoretical facet of this paper is based on the model developed by Shiller (1984) and Sentana and Wadhwani (1992) and the adopted methodology is grounded on a GARCH mean model.

2. Motivation and Purpose

There has been no formal empirical study of the positive feedback traders' activity in the Moroccan stock market although several analysts report reversion to the mean phenomena and describe herding effects. Squalli (2006) explains how Colorado's IPO was subject to a trend effect in the first weeks of trading. The stock price was expected to increase, so investors rushed to buy the IPO in masses. After weeks of trading activity and unfounded price increases, the stocks suffered a series of declines dragging it back to its fundamental value. Other analysts studying the general market trends relate the CSE boom to rational speculation. Drissi El Bouzidi (2006) declares that the overall market is overvalued and that behavioral aspects keep the prices artificially higher than they should be. After periods where the market could gain 30% in a matter of weeks, the CSE indices entered a period a repetitive declines in 2009. The market suffered a "psychological crisis" driven by the small investors' panic as pointed out by Nhaili (2009). The aforementioned market dynamics are significant indicators of the prevailing presence of positive feedback trading. Lack of literature documenting market occurrences leaves analysts, investors, and scholars with a poor understanding of the effects taking place during different trading phases.

Our main objective is to investigate evidence of positive feedback trading activity in the CSE and measure its intensity during market ups and downs. We establish a relationship between positive feedback trading and the presence of serial autocorrelation in the returns of the CSE drawing a link between the level of volatility and the nature of trading. We also examine the effectiveness of various trading strategies in the CSE for several scenarios, comparing the annualized returns over a seven years period (2004-2010) for daily trading versus weekly trading and for short selling possibility and no short selling conduct. We study the profitability prospects for selected trading strategies and techniques; passive strategy, smart money investor's strategy, feedback trading, simple hybrid strategy, and complex hybrid strategy and how they are affected by the day of execution.

3. The Positive Feedback Trading Model

3.1. Positive Feedback Traders

The main assumption of the model developed by Shiller (1984) and Sentana and Wadhwani (1992) is that the market is mainly composed of two types of agents. The first group is named “smart money investors” or “expected utility maximizers” (S). This type of investors relies mostly on the fundamentals corresponding to shares such as profitability, leverage, or cash flows, and its behavior is mainly characterized by risk aversion. The second group of investors, the positive feedback traders also referred to as trend chasers (F) primary trade on price movement and evolution. The demand for the first group is established by the Dynamic Capital Asset Pricing Model developed by Merton (1973) and is given by the following Equation:

$$S_t = \frac{E_{t-1}(R_t)-\alpha}{\mu_t} \qquad (1)$$

Where S_t is the fraction of shares demanded by the smart money investors at period t, E_{t-1} is the expectation operator calculated from all available information at period t-1 as the average yearly return assuming the investor buys at the beginning of the year and sells at the end of the same year; R_t the rate of return at period t using closing prices; α is the rate at which the demand for shares by the smart money investors is null. Setting α equal to the risk free rate, Equation (1) becomes equivalent to the Dynamic Capital Asset Pricing Model developed by Merton (1973). μ_t is the volatility measure as a function of the conditional variance, $\mu_t = \mu(\sigma_t^2)$ is the conditional variance measuring risk at time t. To account for the risk aversion of rational investors $\mu'(\sigma_t^2) \geq 0$ so the higher the volatility, the lower the proportion of shares demanded by the smart money investors.

If all shares are held by Smart Money investors so that $\mathrm{S}_t = 1$ Equation (1) simply becomes the standard Capital Asset Pricing Model (CAPM):

$$\mathrm{E}_{t-1}(\mathrm{R}_t) - \alpha = \mu(\sigma_t^2) \qquad (2)$$

The Demand for the second group; the feedback traders, is set by the following equation:

$$\mathrm{F}_t = \begin{cases} \rho^+ \mathrm{R}_{t-1} & \text{if } \mathrm{R}_{t-1} \geq 0 \\ \rho^- \mathrm{R}_{t-1} & \text{if } \mathrm{R}_{t-1} < 0 \end{cases} \qquad (3)$$

Where ρ^+ and ρ^- indicate the nature of the feedback trading with ρ^+, $\rho^- > 0$ to capture positive feedback trading. In the opposite case where ρ^+, $\rho^- < 0$, there is negative

feedback trading; i.e. selling (buying) when prices increase (decrease). This demand equation is more general than the one suggested by Sentana and Wadhwani (1992) where $F_t = \rho\, R_{t-1}$.

The equilibrium $F_t + S_t = 1$ where all shares are held by both types of investors results in the equation:

$$E_{t-1}(R_t) = \begin{cases} \alpha + \mu(\sigma_t^2) - \rho^+ \mu(\sigma_t^2) R_{t-1} & \text{if } R_{t-1} \geq 0 \\ \alpha + \mu(\sigma_t^2) - \rho^- \mu(\sigma_t^2) R_{t-1} & \text{if } R_{t-1} < 0 \end{cases} \quad (4)$$

The difference between equations (2) and (4) is the additional terms that introduce the positive feedback traders into the CAPM equation and allows for negative serial correlation. This equation shows the relationship between the positive feedback trading and the returns. The terms $\rho^+ \mu(\sigma_t^2)$ and $\rho^- \mu(\sigma_t^2)$ induce negative autocorrelation between the index returns at period t-1 and the returns at period t. These terms also show that the higher the volatility $\mu(\sigma_t^2)$, the more negative the autocorrelation.

The rational expectations assumption states that the expectations of crowds would affect market movements and eventually concretize. This would allow for $R_t = E_{t-1}(R_t) + \varepsilon_t$ which results in the following equation:

$$R_t = \begin{cases} \alpha + \mu(\sigma_t^2) - \rho^+ \mu(\sigma_t^2) R_{t-1} + \varepsilon & \text{if } R_{t-1} \geq 0 \\ \alpha + \mu(\sigma_t^2) - \rho^- \mu(\sigma_t^2) R_{t-1} + \varepsilon & \text{if } R_{t-1} < 0 \end{cases} \quad (5)$$

For testing purposes, equation (5) needs to be transformed into a linear form. The linear form is more suitable for a regression equation. This conversion is achieved by setting $\rho^+ \mu(\sigma_t^2) = \rho_0^+ + \rho_1^+ \sigma_t^2$ and $\rho^- \mu(\sigma_t^2) = \rho_0^- + \rho_1^- \sigma_t^2$. As a result, equation (5) becomes

$$R_t = \begin{cases} \alpha + \lambda\sigma_t^2 - (\rho_0^+ + \rho_1^+ \sigma_t^2) R_{t-1} + \varepsilon & \text{if } R_{t-1} \geq 0 \\ \alpha + \lambda\sigma_t^2 - (\rho_0^- + \rho_1^- \sigma_t^2) R_{t-1} + \varepsilon & \text{if } R_{t-1} < 0 \end{cases} \quad (6)$$

In equation (6), the direct impact of feedback traders is given by (ρ_0^+) and (ρ_0^-) for the case of low volatility levels. As risk increases, the terms $\rho_1^+ \sigma_t^2$ and $\rho_1^- \sigma_t^2$ will have more influence on the return and the impact of feedback trading will be determined by ρ_1^+ and ρ_1^-. If $\rho_0^+, \rho_0^-, \rho_1^+, \rho_1^- < 0$, it is an indication that negative feedback traders are more active in the market. This outcome is more likely to occur during periods of low volatility knowing that negative feedback trading is only one of the hypotheses.

3.2. Non-Synchronous Trading and Serial Correlation

Another important factor that provokes positive serial correlation in return time series is non-synchronous trading. If two stocks are trading in different frequencies, one could react to news more quickly than the second. The lagged response of the second stock could manifest as a positive serial correlation between the two returns. However, the effects of non-trading may not be detectable in the returns of individual securities, but could be more pronounced in portfolio returns (Lo & MacKinlay, 1990). Perry (1985) documents that non-

synchronous trading is not the only cause of correlation in daily market indices, but it needs to be taken into account for the sake of the analysis.

The non-trading is typically associated with periods of low volatility but still present during high fluctuation periods. In order to account for positive feedback trading alone, returns are filtered using a linear autoregressive model LAR (p). The fitted returns have all the autocorrelation induced by non-synchronous trading removed (Koutmos & Saidi, 2001).

4. Strategy Testing and Empirical Results

4.1. Trading Nature and Frequency

Since the Moroccan regulation does not allow for short selling in the Casablanca stock exchange, the analysis is conducted based on the actual trading mechanisms and for the hypothetical scenario where short selling is permitted to measure the practice's impact on the different trading strategies. The strategies are also developed based on distinct trading frequencies. Each group of investors could trade on a daily basis or on a weekly basis. The daily trading allows the investor to decide on an action founded on the displayed closing price. The investor would act at the end of each trading day since the orders are assumed to be exercised instantaneous. In this method of trading, the investor obviously uses daily data (closing prices) as a basis for his decision making. For the weekly trades, the investors enters the markets on a given day of the week and keeps buying or selling (short selling) on that same day of every week. The weekly investor does not discard daily closing prices and bases his trades and forecasting tools on daily data. This investor also acts at the end of the trading exercise given the previous day's price or the next day's forecasted price (same day's Sharpe ratio for smart money investors).

Weekly trading based only on weekly data is also considered as one of the scenarios in this analysis. Traders would again trade on a given weekday, but would only look at previous week's closing price or next week's forecasted price. This trading frequency is only considered for the case of positive feedback trading.

4.2. Strategy Description

Passive Strategy

This is a simple buy and hold strategy that consists of keeping the index portfolio during the whole investment period. It involves entering a long position by buying the index on the first day and selling it on the last day of the study period. The transaction costs are accounted for in this paper and are considered to be equal to 0.22% of the transaction amount.

Smart Money Investors' Strategy

This is an active trading strategy followed by rational investors that rely on stock fundamentals to make their investment decisions. The investors decide on the basis of the Sharpe ratio; they buy (sell) when the Sharpe ratio is higher (lower) than 1. For the case where short selling is allowed, investors buy when the Sharpe ratio is higher than 1 and short sell the index when the Sharpe ratio is lower than 1.

$$\text{Sharpe ratio} = \frac{E_t(R_{t-1}) - \alpha}{\sigma_{t-1}}$$

Where α is the risk free rate and is considered to be equal to 3.27% for the seven years period, and σ_(t-1) is the annualized standard deviation for the previous year's return excluding transaction costs.

Positive Feedback Trading Strategy

This strategy is based on the trend chasing conduct. The investors base their decision on the price movement. So if the prices increase, investors buy the stocks and when the prices decrease they sell the stock. For the case of short selling, investors keep the same behavior in market increases and short sell the stocks in market declines.

Simple Hybrid Strategy

This strategy is adopted by positive feedback traders who integrate forecasting techniques in their trading conduct. Investors use GARCH in Mean (1,1) to model the volatility and a Linear Autoregressive Model to forecast the returns and the prices. Every year's volatility and return forecast is based on GARCH-M estimates derived from the preceding three years' data.

Return equation:

$$\widehat{R}_t = C_1 + C_2 * R_{t-1} + C_3 * \sigma_t^2 + \varepsilon$$

Variance equation:

$$\sigma_t^2 = C_4 + C_5 * \text{Residual}_{t-1}^2 + C_6 * \sigma_{t-1}^2$$

The residuals are assumed to follow a normal distribution and are computed based on actual returns and the forecasted values obtained from the return equation,

$$\text{Residual}_t = R_t - \widehat{R}_t$$

The forecasted returns $\widehat{R}_t$ and actual prices P_t are used to predict the prices $\widehat{P}_t$ as the following:

$$\widehat{P}_t = \widehat{R}_t * P_{t-1} + P_{t-1}$$

If $\widehat{P_{t+1}} \geq P_t$, the feedback traders would buy the stock, and if $\widehat{P_{t+1}} \leq P_t$ the feedback traders would sell the stock or short sell the stock in the relevant scenario.

Complex Hybrid Strategy

This strategy is also adopted by positive feedback traders who integrate GARCH modeling and forecasting in their trading. Investors estimate a GARCH model based on the Log Likelihood tests. The likelihood functions for the various sets of data are maximized using an Exponential GARCH (3,3,2) model with the Log of the variance in the mean equation.

The auto regressive order for the linear mean equation is determined using the tests conducted in the data analysis section and is set to be p=1. The leptokurtic property of the time series reveals that a generalized error distribution (GED) is more suitable for the residuals' estimation.

Each year's volatility and return forecast is based on E-GARCH-M estimates derived from the preceding three years' data.

Return equation:

$$\widehat{R}_t = C_1 + C_2 * R_{t-1} + C_3 * \mathrm{Log}(\sigma_t^2) + \varepsilon$$

Variance equation:

$$\mathrm{Log}(\sigma_t^2) = C_4 + C_5 * \mathrm{abs}\left(\frac{\mathrm{Residual}_{t-1}}{\sqrt{\sigma_{t-1}^2}}\right) + C_6 * \mathrm{abs}\left(\frac{\mathrm{Residual}_{t-2}}{\sqrt{\sigma_{t-2}^2}}\right) + C_7 * \mathrm{abs}\left(\frac{\mathrm{Residual}_{t-3}}{\sqrt{\sigma_{t-3}^2}}\right) + C_8 * \frac{\mathrm{Residual}_{t-1}}{\sqrt{\sigma_{t-1}^2}} + C_9 * \frac{\mathrm{Residual}_{t-2}}{\sqrt{\sigma_{t-2}^2}} + C_{10} * \mathrm{Log}(\sigma_{t-1}^2) + C_{11} * \mathrm{Log}(\sigma_{t-1}^2) + C_{12} * \mathrm{Log}(\sigma_{t-1}^2)$$

The feedback trader strategy is as in the previous simple hybrid strategy.

5. Numerical results

5.1. Results based on trading frequency

Daily Trading

The daily trading for all the strategies appears to be outperformed by the buy and hold strategy that leads to higher returns for the 2004-2010 period. This is mainly due to the high transaction costs incurred on a daily basis. The active traders however are better off during market downs since they benefit from smaller drawdown. Indeed, the buy and hold strategy suffered from a severe drawdown that other strategies did not go through.

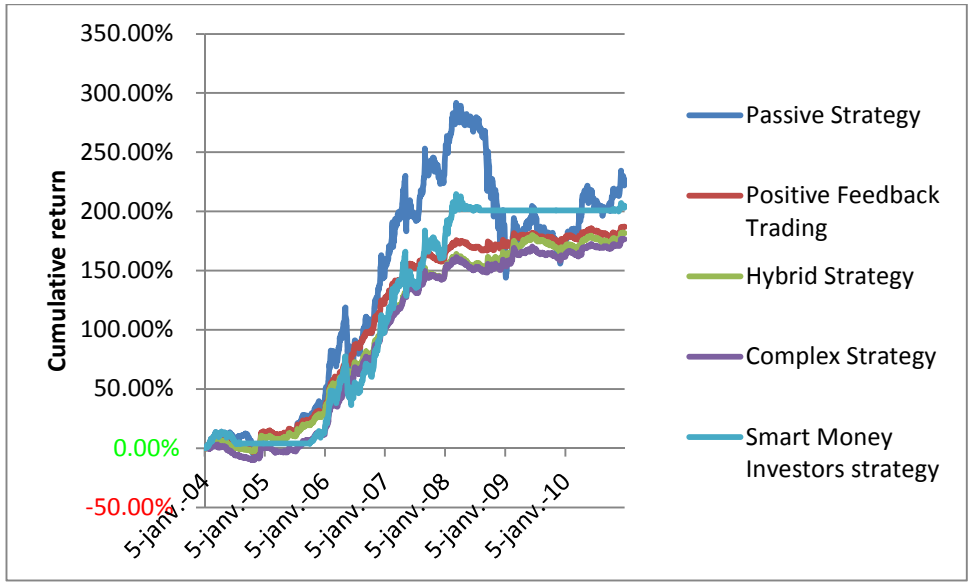


Figure 1: Trading Strategies (daily) with No Short Selling

The hypothetical short selling results in a quite different scenario; it allows traders to take advantage from the drawdown and succeed in beating the passive strategy. The herding traders profit from higher annualized returns and outpace other investors. The short selling does not boost smart money investors' returns since the drawdown caused the Sharpe ratio to fall below 1. The SMI were inactive during and after the down period because their signaling incorporates yearly data and hence ended up with similar annualized return.

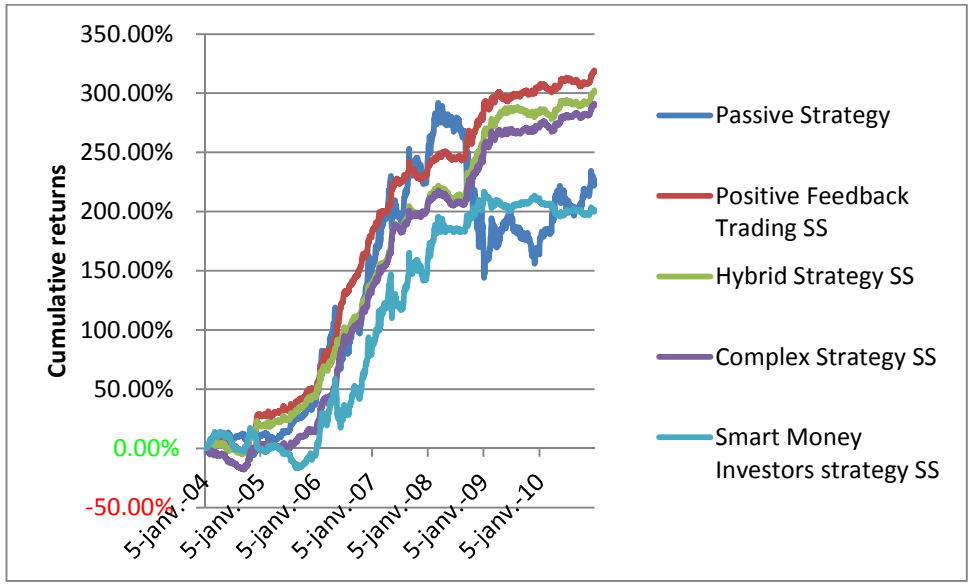


Figure 2: Trading Strategies (daily) with Short Selling

The impact of short selling is more visible on herding strategies in the way that the annualized returns increase significantly accompanied with an increase in volatility. The passive strategy that appears to be the most attractive one at first sight reveals to procure the lowest Sharpe ratio since its returns are subject to severe fluctuations and a high level of volatility. The positive feedback trading on the other hand seems to be the most stable strategy since it has the highest Sharpe ratio since it does not undergo from the sharp drawdown.

		Passive Strategy	Smart Money Investors Strategy	Positive Feedback Trading	Hybrid Strategy	Complex Hybrid Strategy
Annualized Return	No Short Selling	32.28%	29.2%	26.69%	25.93%	25.18%
	Short Selling	32.28%	28.74%	45.36%	42.9%	41.35%
Volatility	No Short Selling	39.28%	22.01%	11%	11.97%	11.55%
	Short Selling	39.28%	23.61%	15.43%	15.56%	15.53%
Sharpe Ratio	No Short Selling	0.82	1.33	2.43	2.17	2.18
	Short Selling	0.82	1.22	2.94	2.76	2.66

Table 1: Annualized Return, Volatility, and Sharpe Ratio for Daily Trading

Weekly Trading (using daily data)

The feedback strategies seem to be dominated by the passive and the smart money investors' strategies on a weekly basis. The gap between the feedback strategies is intensified and the forecasting techniques appear to be more lucrative for weekly trading. The reduction of transaction costs allows for higher annualized returns. The short selling

boosts those returns to increase even further and leads to the strategies having comparable payoffs in terms of the annualized returns.

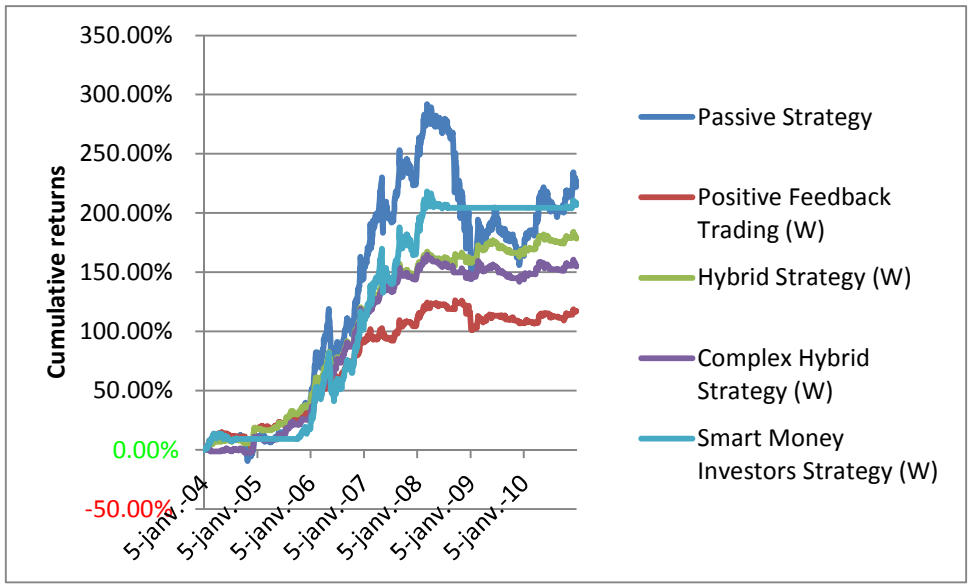


Figure 3: Trading Strategies (weekly) with No Short Selling

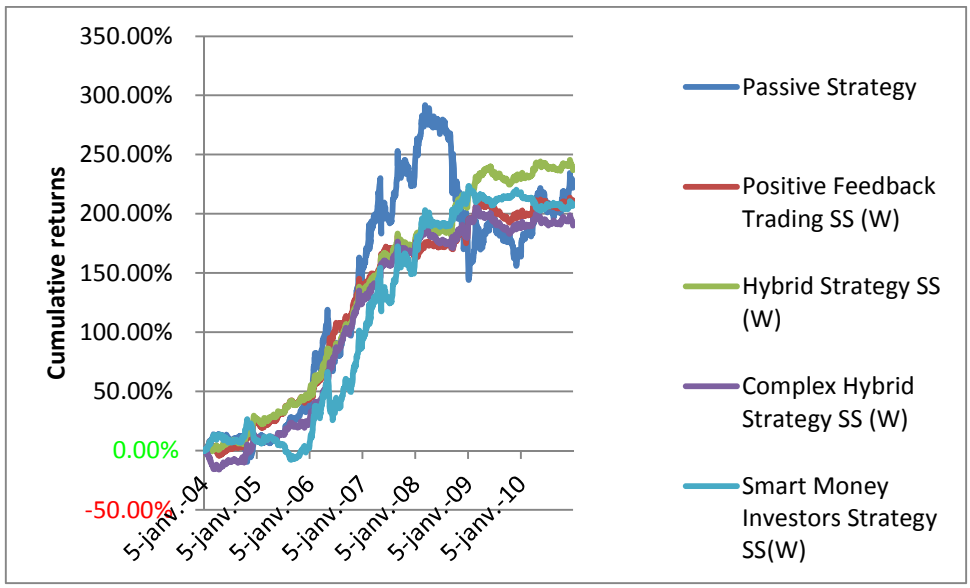


Figure 4: Trading Strategies (weekly) with Short Selling

The impact of the drawdown period is also visible on a weekly basis since it causes the passive strategy to be dominated. It leads to a stagnation of the smart money investors' trading activity who still endures a high volatility level. The feedback trading outperforms the other strategies for both short selling and no short selling scenarios.

		Smart Money Investors Strategy	Positive Feedback Trading	Hybrid Strategy	Complex Hybrid Strategy
Annualized Return	No Short Selling	29.71%	23.28%	25.46%	22.11%
	Short Selling	29.71%	29.97%	33.80%	27.14%
Volatility	No Short Selling	21.93%	11.35%	12.35%	11.78%
	Short Selling	23.54%	15.51%	15.55%	15.56%
Sharpe Ratio	No Short Selling	1.35	2.05	2.06	1.88
	Short Selling	1.26	1.93	2.17	1.74

Table 2: Annualized Return, Volatility, and Sharpe Ratio for Weekly Trading

5.2. Results Based on Trading Nature

Positive Feedback Trading

The positive feedback traders are better off trading on a daily basis during both market ups and downs. The short shelling enhances their annualized returns without having much of an effect on the volatility measures. It also widens the gap between the daily trading and weekly trading, showing that the drawdown is less for daily trading. The positive feedback trading is a stable and profitable strategy despite the high transaction costs the investors may incur.

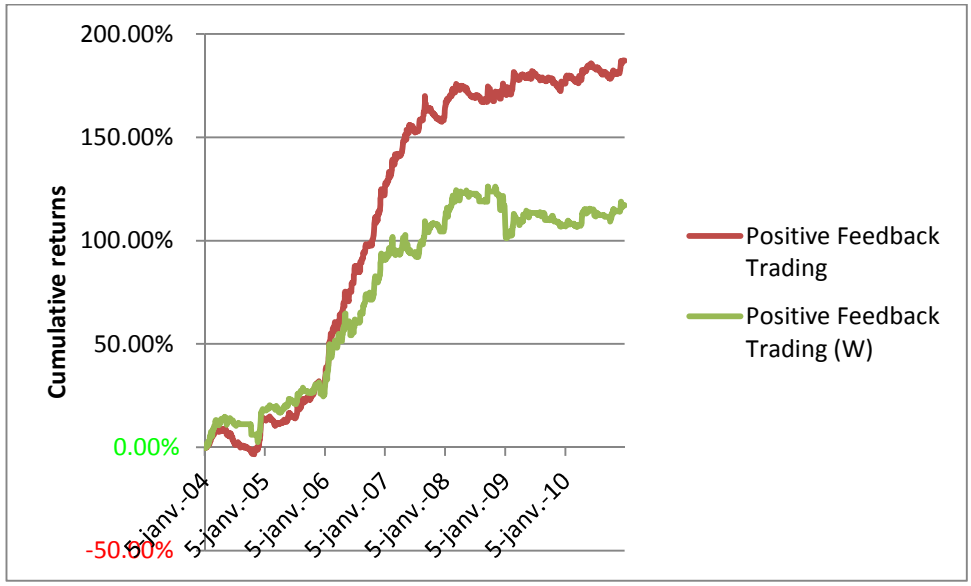


Figure 5: Positive Feedback Trading with No Short Selling

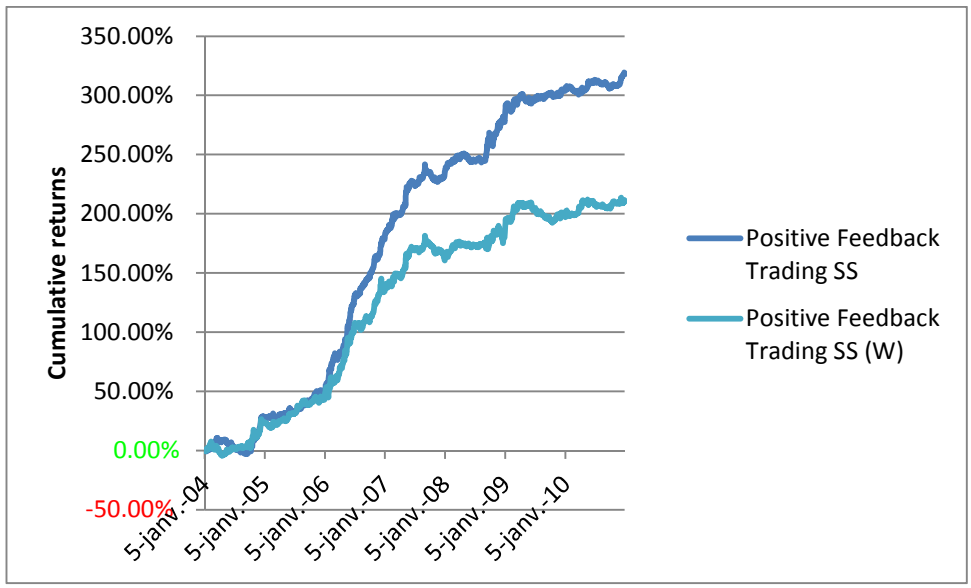


Figure 6: Positive Feedback Trading with Short Selling

		Daily	Weekly
Annualized Return	No Short Selling	26.69%	23.28%
	Short Selling	45.36%	29.97%
Volatility	No Short Selling	11%	11.35%
	Short Selling	15.43%	15.51%
Sharpe Ratio	No Short Selling	2.43	2.05
	Short Selling	2.94	1.93

Table 3: Annualized Return, Volatility, and Sharpe Ratio for Positive Feedback Trading Strategy

Hybrid Strategy

The hybrid strategy based on GARCH-M (1,1) forecasts is close to the positive feedback trading strategy in terms of outcome and profitability. They are also similar for daily and weekly trading and concerning the short selling effect. It is also a profitable strategy especially for the daily trades despite the fact that traders act based on forecasted prices instead of actual prices.

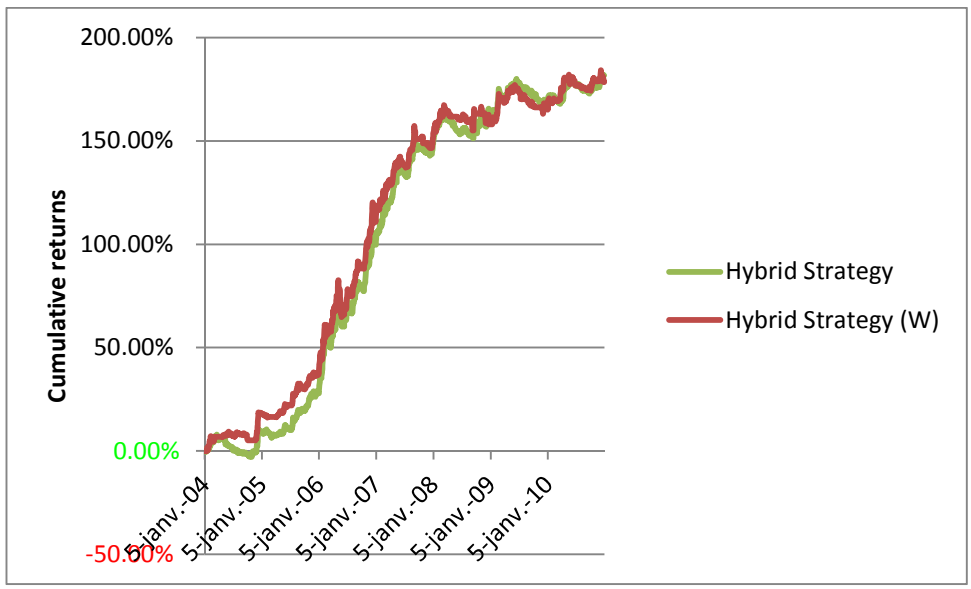


Figure 7: Hybrid Strategy with No Short Selling

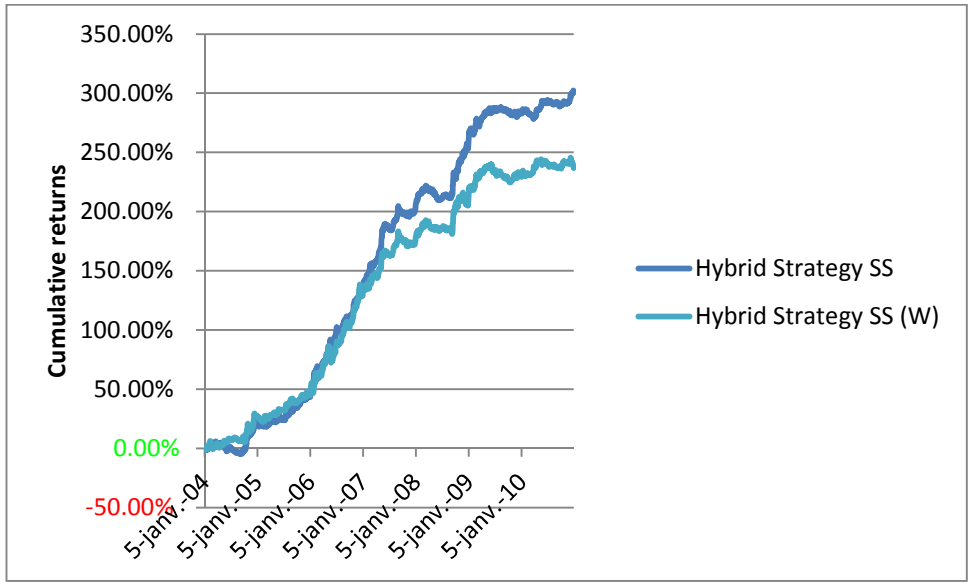


Figure 8: Hybrid Strategy with Short Selling

		Daily	Weekly
Annualized Return	No Short Selling	25.93%	25.46%
	Short Selling	42.9%	33.80%
Volatility	No Short Selling	11.97%	12.35%
	Short Selling	15.56%	15.55%
Sharpe Ratio	No Short Selling	2.17	2.06
	Short Selling	2.76	2.17

Table 4: Annualized Return, Volatility, and Sharpe Ratio for the Hybrid Strategy

Complex Hybrid Strategy

The complex hybrid strategy using E-GARCH-M (3,3,2) is very similar to the hybrid strategy in terms of the gains' characteristics, but the final outcome is lower compared to the other feedback strategies. This may be due to the fact that the forecasts incorporate with a higher precision the historical market dynamics, and result in more conservative prediction due to the sharp drawdown.

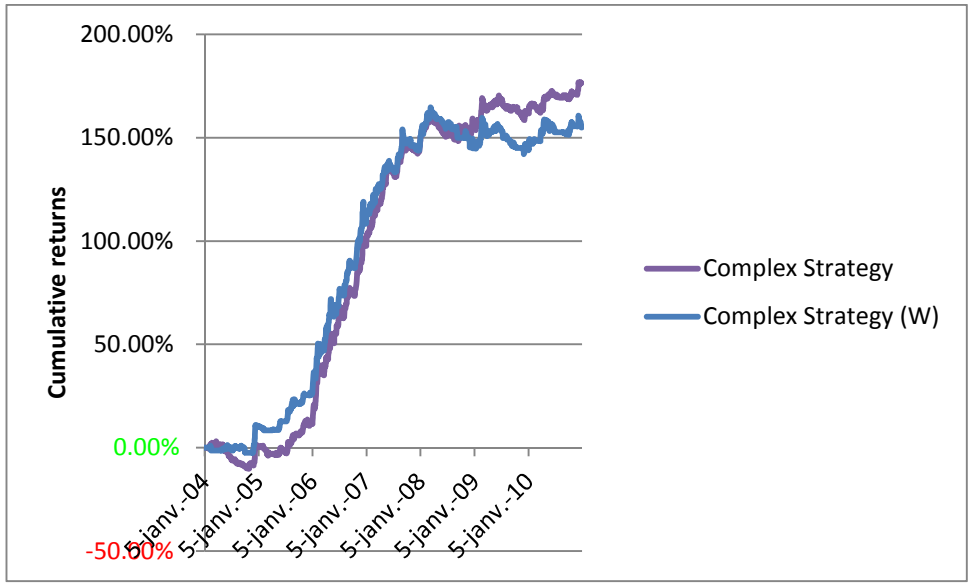


Figure 9: Complex Hybrid Strategy with No Short Selling

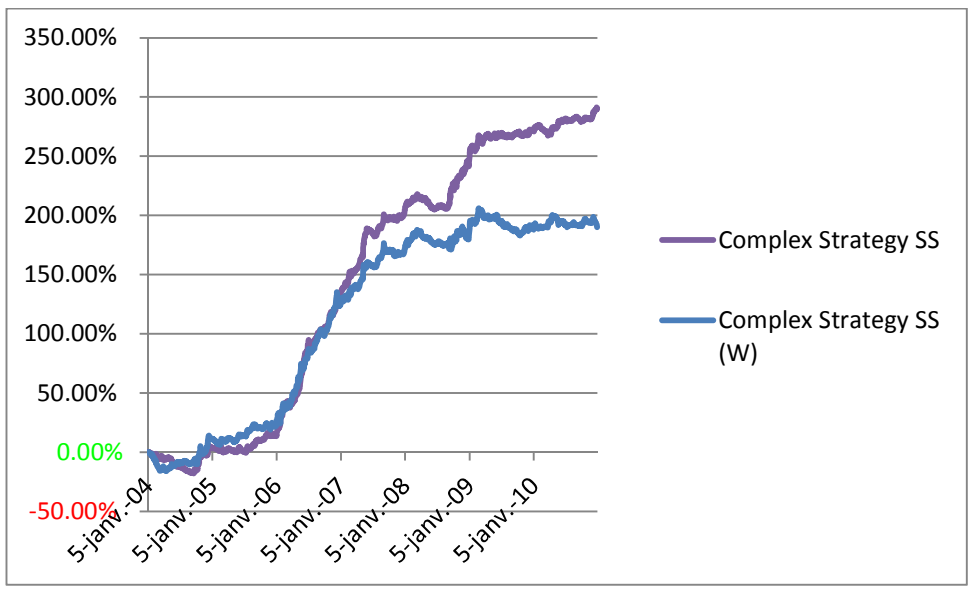


Figure 10: Complex Hybrid Strategy with Short Selling

		Daily	Weekly
Annualized Return	No Short Selling	25.18%	22.11%
	Short Selling	41.35%	27.14%
Volatility	No Short Selling	11.55%	11.78%
	Short Selling	15.53%	15.56%
Sharpe Ratio	No Short Selling	2.18	1.88
	Short Selling	2.66	1.74

Table 5: Annualized Return, Volatility, and Sharpe Ratio for the Complex Hybrid Strategy

5.3. Results Based on Day of Trade

Smart Money Investors Strategy

When trading on a weekly basis, the smart money investors do not seem to be affected by the day they actually enter the market. The steadiness of their payoffs suggests that their trading behavior is less spontaneous and impulsive than other types of trading. The same remarks apply to the weekly trading with short selling for this category of investors.

	Monday	Tuesday	Wednesday	Thursday	Friday
Annualized Return	29.74%	29.38%	29.71%	29.47%	29.31%
Volatility	22.02%	21.91%	21.93%	21.92%	21.91%
Sharpe Ratio	1.35	1.34	1.35	1.34	1.34

Table 6: Annualized Return, Volatility, and Sharpe Ratio Based on the Day of Trade for the Smart Money Investors Strategy

Positive Feedback Trading

The positive feedback traders' returns are highly influenced by the day they exercise their trades. The weekly outcomes of this strategy vary depending on the weekdays, with Wednesday being the best day to trade and Monday being the worst. The standard

deviation of the Sharpe ratios is quite high (47.91%) which accounts for the importance of the entry day effect. The influence of timing on annualized returns remains identical when short selling is introduced and leads to higher variations in the Sharpe ratio translated by 67.11% standard deviation.

	Monday	Tuesday	Wednesday	Thursday	Friday
Annualized Return	8.68%	13.93%	**23.28%**	16.70%	16.13%
Volatility	11.91%	11.88%	**11.35%**	11.63%	11.41%
Sharpe Ratio	0.73	1.17	**2.05**	1.44	1.41

Table 7: Annualized Return, Volatility, and Sharpe Ratio Based on the Day of Trade for the Positive Feedback Trading Strategy

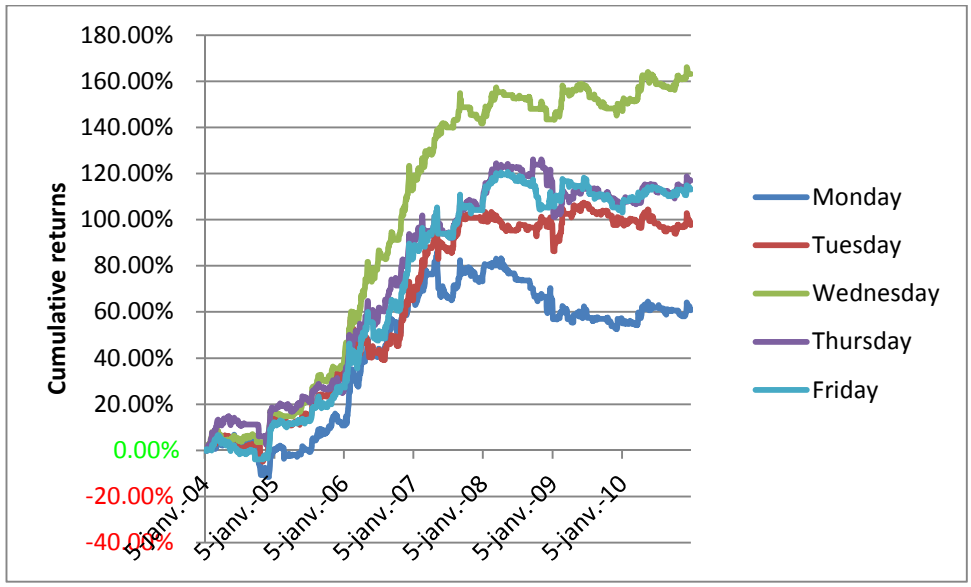


Figure 11: Weekly Positive Feedback Trading Using Daily Data Based on Day of Trade

Hybrid Strategy

The weekly outcomes of the hybrid strategy vary depending on the weekdays and the outcomes are different depending on the day of market entry, Wednesday is the best day to act and Monday is the worst with a Sharpe ratio lower than 1. The standard deviation of the Sharpe ratios is equal to 48.95% for the case of no short selling and 73.61% when short selling is simulated. The influence of timing on annualized returns remains identical when short selling is introduced; and Wednesday's Sharpe ratio undergoes some improvement and Monday's worsens.

	Monday	Tuesday	Wednesday	Thursday	Friday
Annualized Return	10.50%	13.51%	**25.46%**	15.62%	17.38%
Volatility	14.27%	12.66%	**12.35%**	13.01%	12.70%
Sharpe Ratio	0.74	1.07	**2.06**	1.20	1.37

Table 8: Annualized Return, Volatility, and Sharpe Ratio Based on the Day of Trade for the Hybrid Strategy Using Daily Data

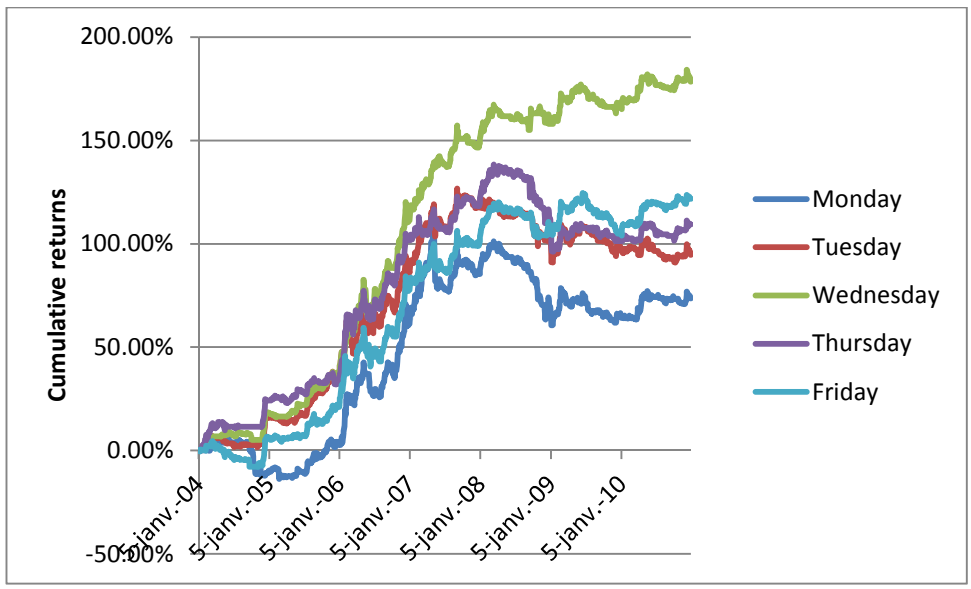


Figure 12: Weekly Hybrid Strategy Using Daily Data Based on Day of Trade

Complex Hybrid Strategy

The complex hybrid strategy also reveals Wednesdays as the best trading days and Mondays as the worst. The high importance of the entry and exit day is accounted for by the high standard deviation that is equal to 44.14% and 62.12% for the cases of no short selling and short selling respectively.

	Monday	Tuesday	Wednesday	Thursday	Friday
Annualized Return	9.72%	11.07%	**22.11%**	14.06%	17.14%
Volatility	12.58%	12.44%	**11.78%**	12.11%	12.33%
Sharpe Ratio	0.77	0.89	**1.88**	1.16	1.39

Table 9: Annualized Return, Volatility, and Sharpe Ratio Based on the Day of Trade for the Complex Hybrid Strategy Using Daily Data

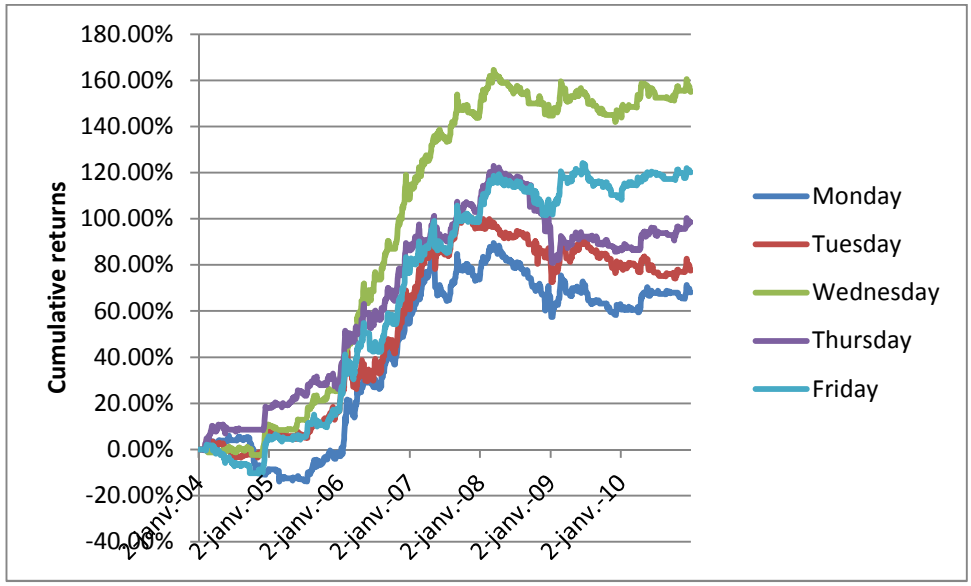


Figure 13: Weekly Complex Hybrid Strategy Using Daily Data Based on Day of Trade

5.4. Results Based on Data Frequency (for Positive Feedback Trading)

Effect of Data Frequency on the Return, Volatility, and Sharpe Ratio

Weekly Positive feedback traders can either rely on daily data or on weekly data to exercise their trades. Using daily data earns higher returns than using weekly prices and is subject to a lower volatility of returns. It is also more profitable for positive feedback traders to act based on daily data since it yields higher Sharpe ratio.

		Daily Data	Weekly Data
Annualized Return	No Short Selling	23.28%	17.32%
	Short Selling	29.97%	17.80%
Volatility	No Short Selling	11.35%	12.03%
	Short Selling	15.51%	15.67%
Sharpe Ratio	No Short Selling	2.05	1.44
	Short Selling	1.93	1.14

Table 10: Annualized Return, Volatility, and Sharpe Ratio for the Daily and Weekly Data (PFTS)

Effect of Data Frequency on the Day of Trade

Using weekly data, the outcome of the positive feedback trading strategy still varies depending of the day of the trade. This change however, is smaller in magnitude than the variation induced when using daily prices. The day of the trade does not matter as much for weekly trading with weekly data, and results in a standard deviation of the Sharpe ratio of 36.94% in the absence of short selling. It is rather high for the short selling scenario case (56.79%), which is foreseeable since the short selling intensifies the market dynamics. When trading with weeks data, the volatility increases on average and the trading day with the highest outcome differs. The Sharpe ratio keeps increasing throughout the week and reaches its peak on Friday for both short selling and no short selling scenarios.

		Monday	Tuesday	Wednesday	Thursday	Friday
	Daily Data	8.68%	13.91%	23.28%	16.70%	16.13%
Annualized Return	Weekly Data	13.96%	7.11%	8.74%	16.65%	17.32%
	Daily Data	11.91%	11.88%	11.35%	11.63%	11.41%
Volatility	Weekly Data	12.27%	11.60%	11.40%	12.00%	12.03%
	Daily Data	0.73	1.17	2.05	1.44	1.41
Sharpe Ratio	Weekly Data	1.14	0.61	0.77	1.39	1.44

Table 11: Annualized Return, Volatility, and Sharpe Ratio Based on the Day of Trade for the Daily and Weekly Data (PFTS)

6. Conclusions

The study of the 2004 -2010 Casablanca Stock exchange main MADEX reveals crucial facts about the characteristics of the returns and the nature of trading conducts in the market place. The significance of the serial autocorrelation in the daily data indicates the plausible presence of positive feedback traders and trends chasers in the market. It also implies that this category of traders could influence market movements and induce negative autocorrelation in the stock returns, trigger mean reversion phenomena, and allow trend predictability.

The analysis of the feedback traders or trend chasers' strategies profitability leads to four main results. First, positive feedback trading on the daily basis beats other herding types of trading and reveals to be steadier than the simple buy and hold strategy whencomparing Sharpe ratios. It allows the traders to go through the intense 2008-2010 drawdown and smooths the severity of the market fall. Positive feedback traders are subject to lower level of volatility even though they are pretty active in the market. The forecasting techniques however lead to inferior outcomes suggesting that there is a price to pay for information losses caused by forecasting.

Secondly, the GARCH forecasting techniques provide superior outcomes for weekly trading using daily data. It allows the trend chasers to incorporate three years of data in their trading conduct. The forecasts seem to be more suitable for less frequent trades since the signaling is unchanged for the hybrid strategy and the complex hybrid strategy, and daily trades only increase transaction costs incurred by the investors. The daily herd trading however results in higher annualized returns when compared to weekly trading, while the volatility does not seem to be affected by the frequency of trades.

Thirdly, short selling generally boosts the annualized returns and intensifies the volatility. The smart money investors are the only group that appears to be immunized from the influence of trading nature as well as frequency.

Finally, the comparison of annualized returns and Sharpe ratio based on the day of trading for the weekly activity reveals that smart money investors are not affected by the day they enter or exit the market, while the herd traders are. In fact, Wednesdays are characterized by lower volatility levels when compared to other week days and yield higher returns. Mondays however are the worst days for feedback traders to act on the stock exchange as they are earning much lower returns and endure higher volatility levels. This phenomenon could be explained by the fact that the market goes through an adjustment stage in the beginning of the week and reaches the "steady state" by the mid-week.

References

[1] D.M. Cutler, J.M. Poterba, L.H. Summers. Speculative Dynamics and the Role of Feedback Traders. The American Economic Review (1990), Vol.80 (2), pp. 63-68.
[2] R.J. Shiller. Stock Price and Social Dynamics. Brookings Papers on Economic Activity (1984), Vol. 2, pp. 457-498.
[3] E. Sentana, S. Wadhwani. Feedback traders and stock return autocorrelations: evidence from a century of daily data. The Economic Journal (1992), Vol. 102, pp. 415-25.
[4] G. Koutmos. Feedback Trading and the Autocorrelation Pattern of Stock Returns: Further Empirical Evidence. Journal of International Money and Finance (1997), Vol 16, pp.625 – 636.
[5] G. Koutmos, R. Saidi. Positive Feedback Trading in Emerging Capital Markets. Journal of Applied Financial Economics (2001), Vol. 19 (24), pp. 291-297.
[6] N. Squalli. Colorado: L'Effet Moutonnier. L'Economiste (2006, November 22nd).
[7] O. Drissi El Bouzaidi. Le Marché Boursier Prend Près de 30 % : Bulle Spéculative ou Pas? La vie économique (2006, February, 10th).
[8] S. Nhaili. Perte de Confiance et Manque de Visibilité à la Bourse de Casa. La vie économique (2009, January, 16th).
[9] A.W. Lo, A.C. MacKinlay. An Econometric Analysis of Non-Synchronous Trading. Journal of Econometrics (1991), Vol.45 (1-2), pp. 181-211.
[10] P.R. Perry. Portfolio Serial Correlation and Non-Synchronous Trading. The Journal of Financial and Quantitative Analysis (1985), Vol. 20 (5), pp. 517-523.

Design Considerations in Laminar Fluid Mixing with Unconventional Geometries

G. Ascanio[1], E. Brito-de la Fuente[2], R. Yatomi[3] and P.A. Tanguy[4]

[1]Centro de Ciencias Aplicadas y Desarrollo Tecnologico, Universidad Nacional Autonoma de Mexico, Circuito Exterior, Ciudad Universitaria, Mexico City, Mexico

[2]Fresenius-Kabi Deutschland GmbH, Oberursel, Germany

[3]Sumitomo Heavy Industries Process Equipment, Inc., Ehime, Japan

[4]Department of Chemical Engineering, Ecole Polytechnique, Montreal, Canada

Corresponding author E-mail: philippe.tanguy@polymtl.ca

Abstract. *Laminar mixing in stirred vessels is often inefficient due to the presence of flow structures generated during the process. Very long mixing times are then required for achieving a specific homogeneity level into the tank. Although the easiest way of removing such structures consist of increasing the impeller rotational speed, in many cases the turbulent regime cannot be achieved due to excessive media heating by viscous dissipation, torque limitations on the kinematic chain, and the costly power consumption. To alleviate these obstacles, alternate configurations specifically designed for operation in the laminar regime have been introduced. This paper deals with a review of these non-typical alternate configurations based mainly on off-centered impellers, multi-shaft equipment and planetary mixers. It will be shown that the performance of these systems paves the way to a new technological paradigm in laminar mixing.*

Key words: laminar, chaotic mixing, stirred vessels.

1. Introduction

Mixing is a key industrial process operation involving physical and/or chemical changes in the process media, which finds many applications in dispersions, mass transfer, chemical reactions and homogenization, among others. Some processes require mixing operations under laminar regime, which is frequently used with viscous fluids, fluids exhibiting complex rheology properties and shear-sensitive products. Many examples can be found in the food industry and polymer processing. Under such conditions, spurious flow structures such as caverns or islands are generated in the process vessel and they can remain for long (or even infinite) mixing times, so that the process become costly and inefficient. The presence of caverns, defined as well-mixed regions around the impeller was observed by Elson by agitating yield stress fluids (1). Dead flow or quasi-static flow regions also occur far from the agitator as a result of the poor momentum transfer due to viscous effects. These flow pathologies are produced mainly when agitating under symmetry conditions, i.e. with the agitator centered in the vessel. A typical image of such pathologies is show in Figure 1.

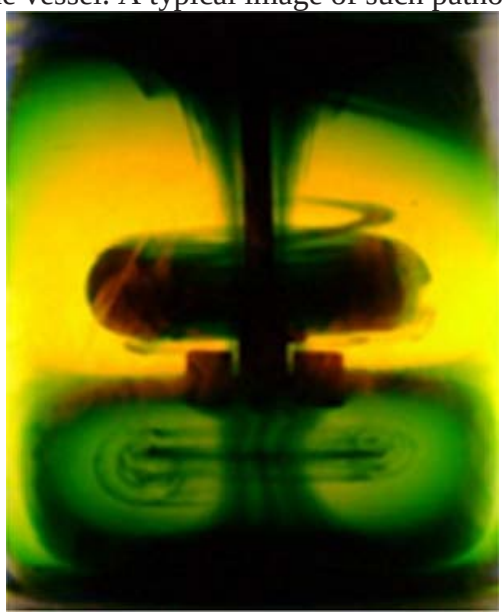


Figure 1: Flow structures generated under symmetric conditions (2).

This figure shows well-mixed regions (pseudo caverns) denoted by two ring vortices formed below and above the impeller. Stagnant regions surround the pseudo caverns and dead zones are formed far from the impeller in the top of the vessel (2). The formation of the two ring vortices pattern was first observed when agitating Newtonian fluids in the laminar regime (3). In a numerical study, the existence of such vortices was predicted founding that their size and position depend on the Reynolds number (4). Although such flow structures can be easily destroyed by increasing the rotational speed, this results in an increase of the power draw and the generation of intense shear stresses that can sometimes deteriorate the product quality (mechanical or thermal degradation). To alleviate these obstacles, alternate configurations specifically designed for operation in the laminar regime have been introduced. The objective of the present paper is to give a general overview of different mixing scenarios and configurations in stirred vessels based on of off-centered impellers, multi-shaft equipment and planetary mixers adapted to the case of inelastic non-Newtonian fluids. It will be shown that the performance of these systems paves the way to a new technological paradigm in laminar mixing.

2. Off-centered impellers

The usual practice with mixers is to design equipment based on symmetric characteristics. On one hand, vibrations are easier to control in symmetric mixers especially when agitating in the turbulent regime and on the other hand, processes requiring variable speed are also easier to implement. Flow structures generated in stirred vessels under symmetric conditions with the impeller operating steadily in the laminar regime have been extensively reported. Symmetry induces serious mixing malfunctions such as isolated well-mixed regions as well as quasi-static or stagnant flow zones. Such flow structures are the result of non-chaotic coherent zones without interaction with the entire flow domain. These mixing malfunctions can be avoided by increasing the kinetic forces. The use of dynamic perturbations has been proposed (5-7) but they imply the use of sophisticated speed control devices and protocols that cannot be readily used in the plant environment. Therefore other approaches based on geometrical perturbations by displacing the impellers from the vessel centerline seem to be attractive alternatives. To our knowledge, little information about the use of off-centered impellers operating in the laminar regime has been published in the open literature.
The formation of flow structures by agitating Newtonian fluids with three coaxial Rushton turbines in the laminar regime has been also reported (8, 9). One could expect that the mutual interaction of the discharged flows would promote good mixing characteristics. However, flow segregations, i.e. pseudo caverns, stagnant regions and isolated islands remain visible after mixing for long periods. Figure 2 shows the typical flow structures revealed by laser fluorescence induced in a stirred vessel equipped with three Rushton turbines coaxially placed. Islands regions are denoted by the dark regions below and above the impellers (9).

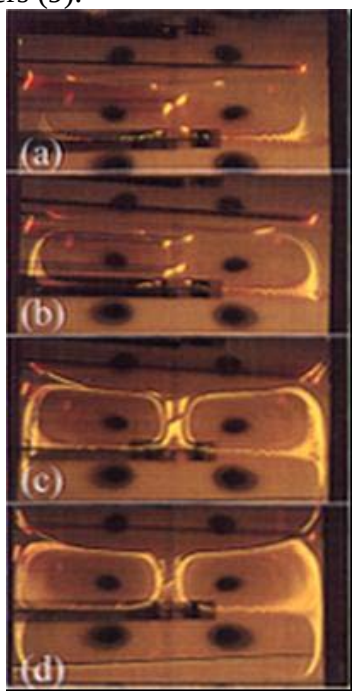


Figure 2: Flow structures observed in a stirred vessel with three coaxial Rushton turbines operating in the laminar regime (10).

Such regions are the result of the structure of the global unstable manifold characterized by periodic, symmetric flows that never fill the entire flow domain. These mixing pathologies can be easily removed by increasing the rotational speed. Because an increase of speed leads to higher shear stresses, such approach can yield poor performance when mixing shear sensitive media.

Under laminar flow conditions it has been demonstrated that these mixing inefficiencies can be overcome or avoided if the flow is continuously perturbed (11). Lamberto et al. (5) proposed the use of variable speed protocols to create chaotic flow required to achieve efficient mixing in the laminar regime. Another way of creating chaos in the laminar regime consists of varying the speed and the direction of rotation (6, 7).

The idea of using off-centered impellers is based on the seminal work by Aref (12), in which it was demonstrated that the formation of coherent regions in the vicinity of eccentric cylinder could be prevented by chaotic advection. This finding is depicted in Figure 3, in which a square array of tracers is initially stirred. After time t_6 the particles are distributed and then they are homogenized over a chaotic region. This finding was later confirmed by Aref and Balachandar (13) and Franjione et al. (14) and applied independently to stirred vessels by Alvarez et al. (9) and Ascanio et al. (7), who demonstrated that coherent regions are avoided or readily destroyed if the agitation shaft is displaced from the vessel centerline.

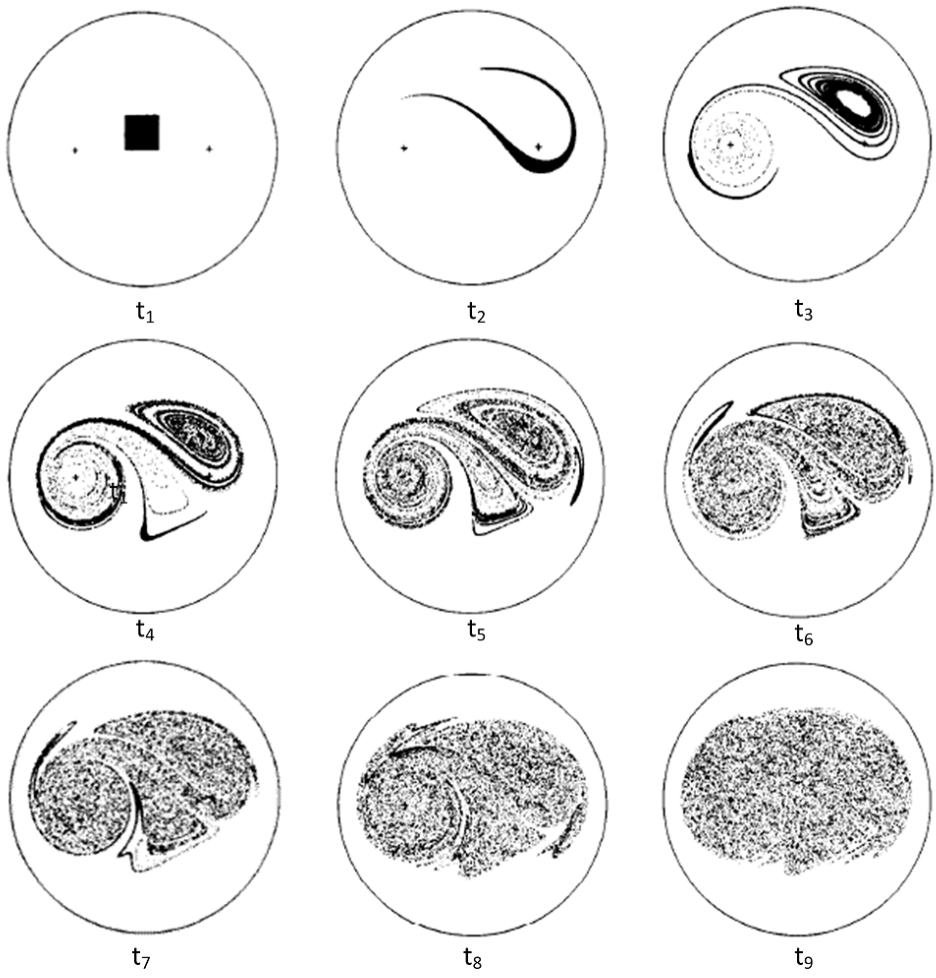
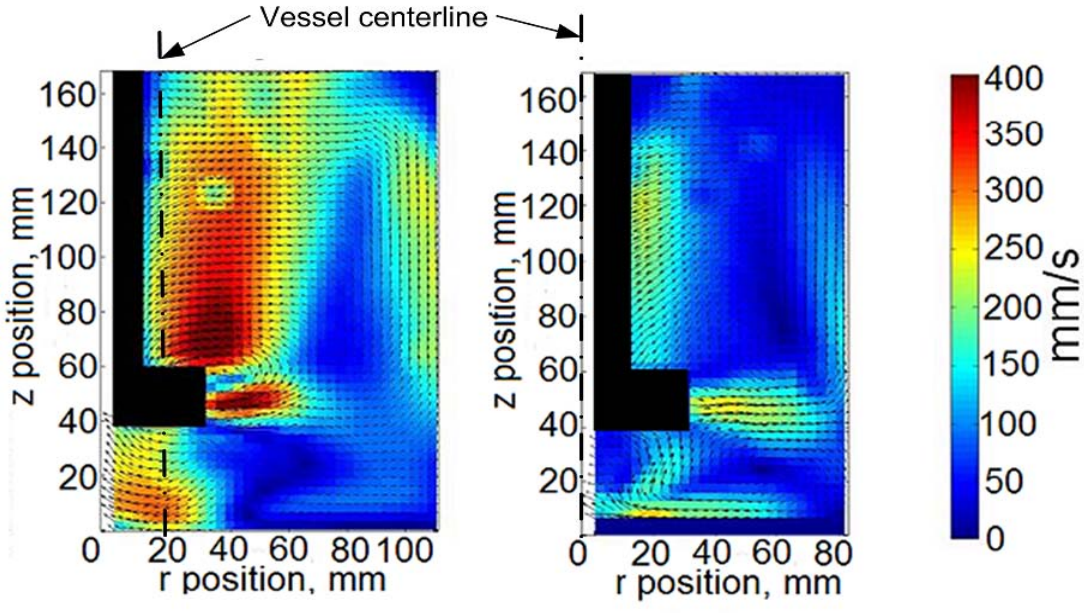


Figure 3: Phases in the stirring of initially square array of tracers.

Hidalgo-Millán et al. (15) investigated the effect of the impeller eccentricity on the pumping capacity in unbaffled stirred vessels with Newtonian fluids by using the Particle Image Velocimetry technique (PIV). Figure 4 shows the flow patterns observed with an impeller placed at two different

eccentricity ratios, *e*: (a) e = 0.36 and; (b) e = 0 (centered). The eccentricity ratio, *e*, is defined as the ratio of radial displacement measured from the vessel centerline to the vessel radius (2).



(a) (b)

Figure 4: Flow patterns in stirred vessels in the laminar regime:
(a) off-centered, e = 0.36; (b) Centered, e = 0 (2).

An intensive mixing can be observed under off-centered conditions. Regions around the impeller in which the fluid reaches speeds of the order of 400 mm/s are obtained under non-symmetric conditions, while the centered case is mostly characterized by quasi-stagnant regions and zones having a maximum speed of about 250 mm/s. As a result of the intensive stirring with off-centered impellers the pumping capacity increases yielding shorter mixing times.
The benefits of off-centered impellers are well recognized in the literature. However, a conservative design approach based on centered systems is very often used because they are mechanically more stable and vibrations are easier to control. The use of eccentric impellers implies some modifications such as the redesign of reactor lids, couplings and supports for motors. Another important for considering a mixer system with off-centered impeller is the associated costs. The mechanical drive unit of off-centered mixers is the same as that of conventional mixers. However, in the case of mechanical drives directly mounted on the vessel direct, a flange is required and the thickness of the vessel wall must be increased in order to provide the structural strength and prevent vibrations. If an off-centered mixer system is considered, initial cost will rise up slightly about 10-30% of costs total (mechanical drive and vessel) in comparison with the conventional mixing system at the same power consumption. However, both operating and maintenance costs is the same as the ones when using conventional mixers.

3. Multi-shaft equipment

Some industrial processes require simultaneous operations such as dispersion and homogenization, which requires a priori very different mixing geometries, in particular in the laminar regime. Multiple open impellers fitted on the same shaft have been used for improving dispersion features

in stirred vessel with the drawback of large power consumption. Close clearance impellers such as anchors and helical ribbons are preferred when mixing viscous fluids. The combination of both approached into a single multi-shaft equipment offers an attractive alternative in applications where the two functions are needed.

Little information has been published in the literature regarding hybrid systems performing more than one mixing function. Figure 5 shows the experimental setup proposed by Espinosa-Solares et al. (16), who investigated the hydrodynamics of a coaxial mixer consisting of a Rushton turbine and helical ribbon impeller. The choice of such impellers was based on the well-known dispersion capacity of the Rushton turbine, while the helical ribbon proved to be a good option for bulk viscous mixing.

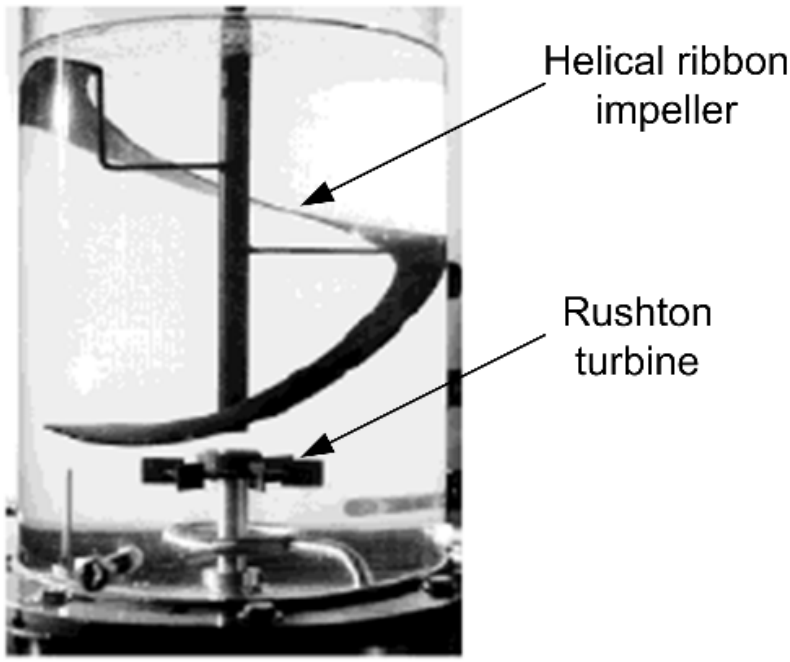


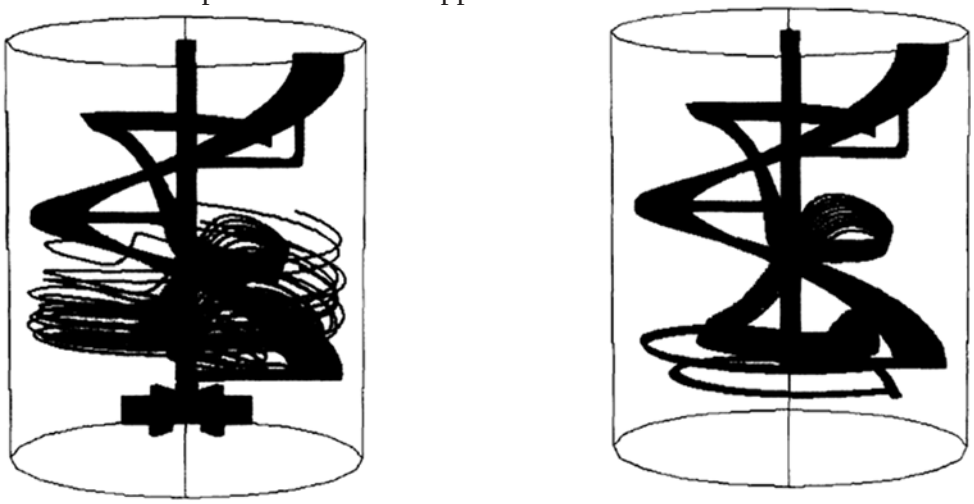
Figure 5: Independent coaxial mixer (16)

Thibault (17) compared the performance of an anchor-pitched blade turbine with rod wetting devices and a Cowles turbine during the make down of clay slurry. He found that a lower specific power is required when using the multi-shaft equipment.

Tanguy et al. (18) compared the mixing performance in terms of pumping, dispersion capabilities, and power consumption of a new dual impeller mixer composed by a disc turbine and a helical ribbon impeller mounted on the same axis but rotating at different speed. When the fluid rheology evolves during the process the dual impeller mixer better performs than the standard helical ribbon in terms of the top-to-bottom pumping. Figure 6 shows the dispersion patterns obtained by computing the trajectories of a series of tracers. As figure 6a shows, the effect of the disc turbine placed in the lower side of the tank is quite remarkable, which hampers the agglomeration of particles observed with the helical ribbon impeller only (Fig. 6b).

Mozaryn et al. (19) reported multi-shaft system consisting of a propeller and a helical ribbon impeller in a draught tube. Espinosa-Solares et al. (16) investigated the flow patterns of two dual mixers consisting of independently driven impellers. Rushton turbines or Smith turbines were used as high speed impellers, while a helical ribbon was chosen as low speed impeller. For low-viscosity fluids, if the helical ribbon is kept static, the vortex length is reduced, the ribbon acting as a baffle.

With the hybrid system with both the turbine and the helical ribbon in rotation, the vortex is eliminated when the impellers rotated in opposite directions.



(a) (b)

Figure 6: Dispersion pattern: (a) Induced by the dual impeller; (b) Induced by the helical ribbon only (18).

Foucault et al. (20) investigated the mixing efficiency using mixing times and power consumption as criteria to evaluate a coaxial system consisting of wall scraping impellers and dispersing impellers (see Figure 7). They found that operation in co-rotating mode was the best choice when using Newtonian and non-Newtonian fluids. They also reported an increase of the impeller speed on the anchor power consumption regardless the rotating mode.

Foucault et al. (21) characterized a coaxial mixer consisting of a wall-scraping anchor and different dispersion impellers (radial discharge) operating in co- and counter-rotating modes with Newtonian and non-Newtonian fluids. They observed that the anchor speed did not increase the power consumption of the dispersion turbines in co-rotating mode. However, higher power is required when turbines are operated in counter-rotating mode. Figures 8 and 9 show the dispersing turbines used by Foucault et al. (21) and a sketch of the rotating modes investigated.

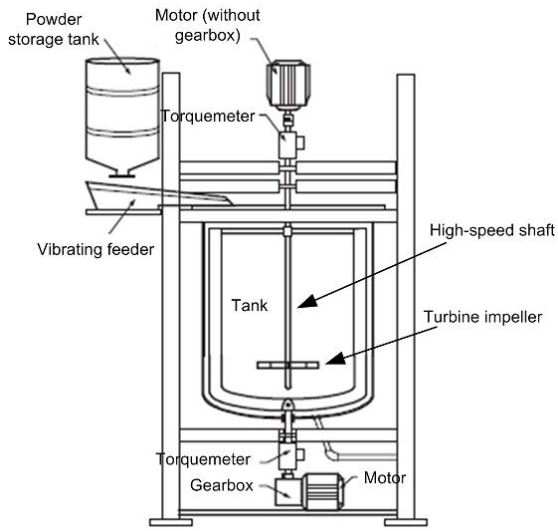


Figure 7: Coaxial mixer with scrapping and dispersing impellers (20).

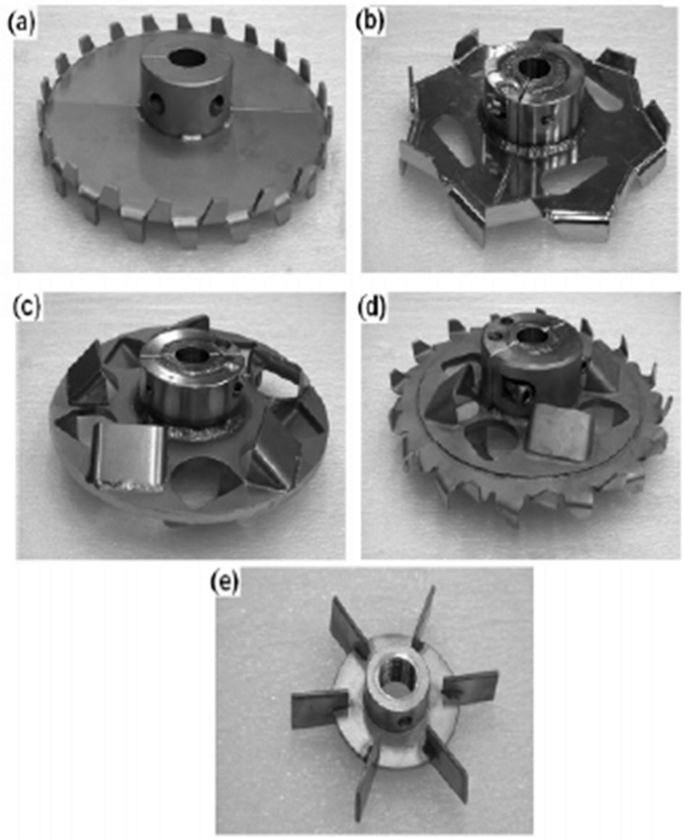


Figure 8: Dispersing turbines: (a) Cowles; (b) Deflo;
(c) Sevin; (d) Hybrid turbine; (e) Rushton turbine (21)

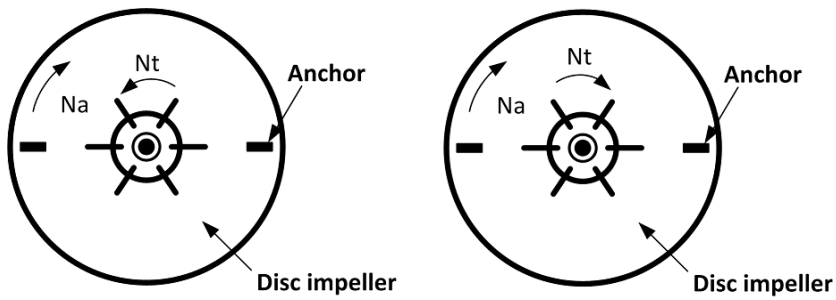


Figure 9: Rotating modes: (a) Counter-rotating; (b) Co-rotating. (21)

Khopkar et al. (22) investigated the emulsification capabilities of a dual shaft mixer comprising a Paravisc impeller and an off-centered rotor-stator (see Figure 10). For that purpose different oil to water ratios were used. They found the rotor-stator has little influence on the process time and the Paravisc has a dominant contribution in the overall energy consumption. Also they noticed that droplet size increases when increasing the oil concentration.

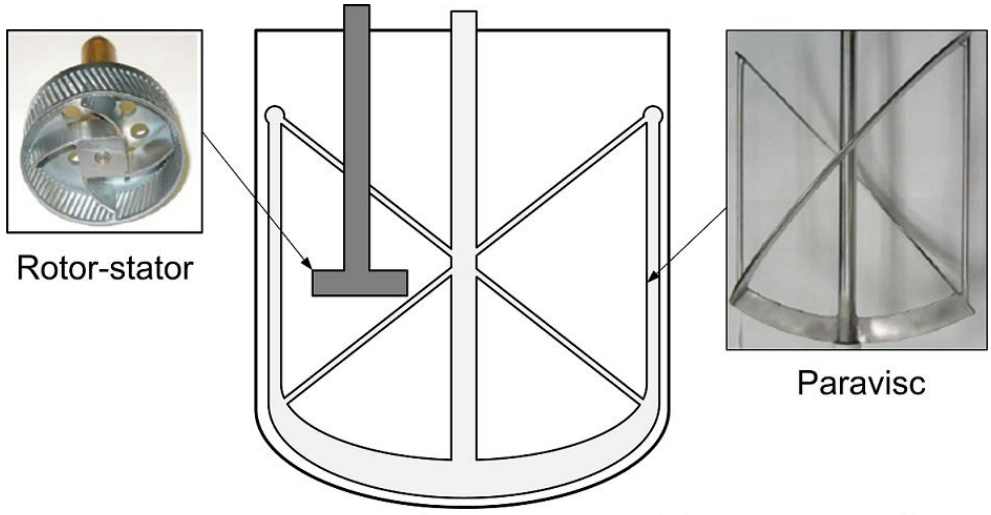


Figure 10: Mixing system with a rotor-stator and the Paravisc impeller.

Recently, a new technology combining a paddle-like impeller (Maxblend) with a double helical ribbon impeller, called as Superblend mixer, has been introduced by SHI Mechanical & Equipment (23). This system has been used for mixing applications where drastic viscosity changes in the medium result in a change of regime from the upper transition regime to deep laminar as in polymerization reactions. Rivera et al. (24) performed a numerical study of the Superblend coaxial mixer, which consists of a Maxblend impeller and a double helical ribbon agitator mounted on two independent coaxial shafts rotating at different speeds (see Figure 11).

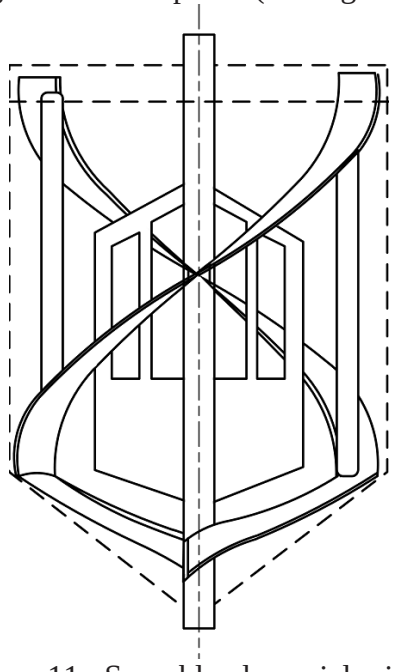


Figure 11 : Superblend coaxial mixer.

The numerical study was based on the solution of the classical Navier-Stokes equations with help of a parallel three-dimensional finite element solver. The rotation of agitators was modeled by using a hybrid approach based on a novel finite element sliding mesh and fictitious domain method. Based on the values of pumping and shearing, both the average pumping and shearing numbers were higher when operating the mixer in the counter-rotating mode. Figure 12 shows a comparison of the

rotating modes of the Superblend in the transition regime, in which a passive scalar was dispersed by the flow field into the vessel to determine the mixing time of the simulated scenarios.

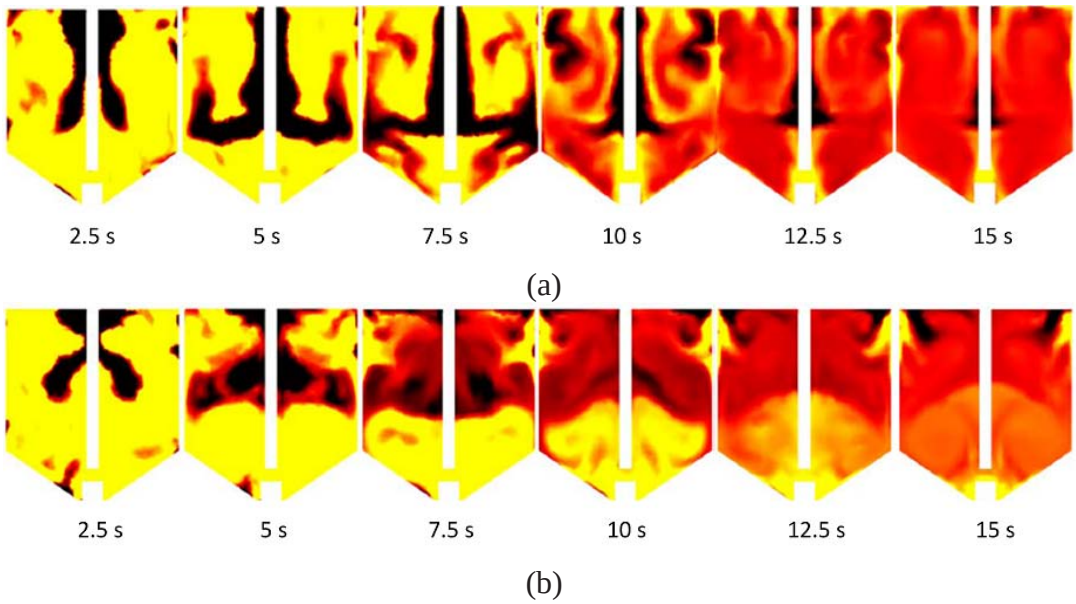


Figure 12: Tracer dispersion mechanisms at Re = 520: (a) co-rotating; (b) counter-rotating. (24)

As Figure 12 shows, the tracer disperses faster in the first seconds in the counter-rotating mode; however, after 15 s the tracer has been almost dispersed when operating in the co-rotating mode resulting in shorter mixing times. With respect to the power consumption, due to the presence of high-shear vortices between the coaxial impellers, high energy is required in the counter-rotating mode. These results were experimentally confirmed by Farhat et al. (25), who demonstrated also that the Superblend exhibit higher performance than other coaxial configurations based on mixing efficiency criteria such as mixing time and energy consumption.

Although cylindrical vessels are commonly used for coaxial mixers, other geometries such as conical mixers have also used for dispersing purposes. Conical mixers were introduced in the 50´s to ensure an adequate axial pumping of viscous suspensions. To our best knowledge, the only published investigation dealing with conical mixers is the contribution of Dubois et al. (26), who reported physical and numerical experiments with a bench-scale mixer consisting of two counter-rotating intermeshing helical ribbons in a conical housing.

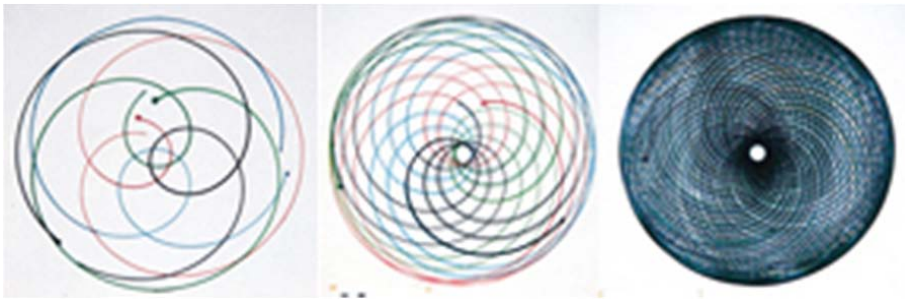
Multi-shaft equipment is clearly capable of delivering efficient mixing with a high degree of homogenization. However, one must take into account that this kind of technology is suitable mainly for very viscous and time-dependent rheology fluids. When used at industrial scales, the vessel volume is usually limited.

From a manufacturing point of view, coaxial equipment requires sophisticated technologies, not only for the mechanical design but also for precision machining, fabrication, assembly and shaft alignment. In particular, the use of coaxial impellers implies very advanced mechanical seals with a double pipe configuration. Moreover, the mechanical drive is also more complex especially in applications requiring a variable speed ratio. As a consequence initial costs drastically rise up about 1.5 - 2 times of total cost in comparison with conventional technology (the cost of the mechanical drive unit itself is 2 to 3 times of conventional one). Because of the difficulties involved during the clean-up, disassembling and re-assembling operating and maintenance costs will rise up too. On the other hand, in some cases the impellers are driven separately, one from the top and the other from

the bottom. Although, the speed ratio is relatively easy to vary, mechanical sealing in the bottom could increase the initial and operating costs.

4. Planetary mixers

A planetary mixer is a device used to provide homogeneous and intensive mixing due to the off-centered movement of the product, which is created by the eccentric position of the mixer arm in relation to the vessel centerline. They are designed to provide complete and effective dead spot free mixing due to the planetary motion and find applications in the makeup of complex liquid formulations, powders and pastes. Figure 13 shows the pathlines patterns observed at different times (27).



(a) (b) (c)

Figure 13: Cross-section flow fields observed with planetary mixers:
(a) In the beginning; (b) after one revolution; (c) after several revolutions. (27)

Planetary mixers are not only designed to provide intensive mixing but also for dispersing, kneading and deareating products such as adhesives, pastes, coatings, granulations and media of moderate to high viscosity. Figure 14 shows an image of a typical industrial double planetary mixer.

Figure 14: Double planetary mixer (28)

Little information regarding planetary mixing has been found in the literature. Tanguy et al. (29) performed a numerical study of viscous materials exhibiting yield stress and non-Newtonian properties such as slurries, whose rheology can be described by the Bingham model. Flow structures and power number as a function of the Bingham number allowed the Metzner-Otto shearing constant, *Ks*, to be obtained leading to a generalized power curve. Such results were later confirmed (30, 31) with double planetary mixer. Figure 15 shows an example of the typical dispersion observed with a planetary mixer.

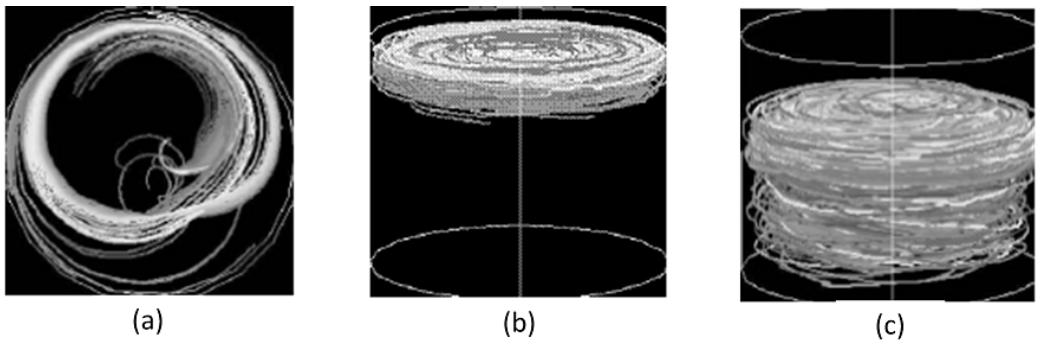


Figure 15: Typical dispersion observed with a planetary mixer:
(a) radial; axial near the free surface; (c) axial at the bottom (30).

Clifford et al. (32) carried out a numerical and experimental study of a simple planetary mixer at very low Reynolds numbers. They found that numerical simulations in the Stokes flow regime may be used as parameter to selecting good mixing protocols at low Re. Delaplace et al. (33) performed a dimensional analysis for a planetary mixer equipped with a vertically and centrally mounted impeller in the tank with Newtonian and non-Newtonian fluids with the aim of clarifying the definition of the characteristic speed of the mixer. Connelly and Valenti-Jordan (34) performed a numerical study of a planetary pin mixer with viscous Newtonian fluids. They observed that planetary pin mixer did not experience as much axial mixing as cross-sectional mixing over the same time span; however, the mixing efficiency is highly dependent on the spin position.
Because planetary mixing promotes the motion of fluids in the radial direction while it is pumped from the bottom to the top, it is commonly used from laboratory scale to industrial level but limited for very viscous fluids and solid-like medium such as pastes and concentrated slurries. Its use for low-viscosity fluids can result in long mixing times and energy consumption. It is important to point out that planetary mixers are mostly used at atmospheric pressure or under slight vacuum, for instance in the food industries. At atmospheric pressure, as no mechanical sealing is required, the cost of the special gearbox and impellers are the main issue to be considered as initial investments.

5. Conclusion

Table 1 summarizes the main features of the non-typical mixer configurations that have been reviewed, which can help to select a mixing system. Although dynamic perturbations appear to be an efficient alternative for improving mixing in terms of mixing times and the energy required to achieve the desired level of energy, they are hard to implement in industrial mixers, therefore other approaches based on other geometries should be taken into account.

Mixer configuration	Viscosity level	Phases	Typical applications
Off-centered impellers	Low to medium	Liquid-liquid	• Liquid blending
Multi-shaft impeller	Medium to high	Liquid-liquid Liquid-solid	• Suspension • Dispersion • Emulsion
Planetary mixer	Very high	Liquid-solid	• Pastes • Concentrated suspensions

Table 1: Main features of mixing configurations

From an industrial standpoint, mixing is a very complex unit operation in which many factors are involved. The fluid rheology and the type of applications are indeed the most important factors to consider when selecting a mixing system. From an economic standpoint, the costs associated with unconventional geometries tend to increase with respect to the conventional mixing system at the power consumption. However, the use of such geometries can result in a higher performance mixing system especially in the laminar regime providing better quality products.

Nowadays, most of the technical and scientific information reported in the literature deals with the hydrodynamics in stirred vessel focusing mainly to measure mixing times and power consumption as mixing efficiency criteria for selecting a system. Although, some works with non-Newtonian fluids have been reported, there is still a significant knowledge gap with fluids exhibiting complex properties and time-dependent rheology. It is clear that more research is needed looking to understand the role of the material properties of the products to be mixed on the process efficiency. Laminar mixing is still an open field where a new technological paradigm is needed to ensure operating efficiency, consistency and performance.

References

[1] T.P. Elson, D.J. Cheesman and A.W. Nienow, A.W. X-ray studies of well-mixed region sizes and mixing performance with fluids possessing a yield stress. Chem. Eng. Sci., 41 (1986), 2555-2562.

[2] A. Hidalgo-Millán. Geometric perturbations in mechanically agitated vessels. Ph.D. thesis in Chemical Engineering, National Autonomous University of Mexico (in Spanish) (2010).

[3] J. Solomon, T.P. Elson, A.W. Nienow and G.W. Pace. Cavern Sizes in Agitated Fluids with a Yield Stress, Chem. Eng. Comm., 11 (1981), 143-164

[4] D.J. Lamberto, M.M. Alvarez and F.J. Muzzio. Experimental and Computational Investigation of the Laminar Flow Structure in a Stirred Tank. Chem. Eng. Sci., 54 (1999), 919-942

[5] D.J. Lamberto, F.J. Muzzio, P.D. Swanson and A.L. Tonkovich. Using Time-dependent RPM to Enhance Mixing in Stirred Vessels", Chem. Eng. Sci., 51, (1996), 733-741

[6] G. Ascanio, M. Brito-Bazán, E. Brito-de la Fuente, P.J. Carreau, P.A. Tanguy. Mixing Enhancement of non-Newtonian Fluids Using Unconventional Configurations. ASME Fluids Eng. Div. Summer Meeting, Montreal, Canada, July 14-18, (2002a), paper FEDSM2002-31359

[7] G. Ascanio, M. Brito-Bazán, E. Brito-de la Fuente, P.J. Carreau, P.A. Tanguy. Unconventional Configuration Studies to Improve Mixing Times in Stirred Tanks, Can. J. Chem. Eng., 80, (2002b), 558-565

[8] M.M. Alvarez. Using Spatio-Temporal Asymmetry to Enhance Mixing in Chaotic Flows: From Maps to Stirred Tanks Doctoral Thesis, The State University of New Jersey, (2000)

[9] M. Alvarez, P.E. Arratia, F.J. Muzzio. Laminar mixing in eccentric stirred tank systems, Can. J. Chem. Eng., 80, (2002), 546-557

[10] M. Alvarez, J. M. Zalc, T. Shinbrot, P. E. Arratia, and F. J. Muzzio. Mechanisms of Mixing and Creation of Structure in Laminar Stirred Tanks, AIChE Journal, 48, No. 10, (2002), 2135-2148.

[11] E. Brito-de la Fuente, S., Hernández, A. Segura, L. Medina, P.A. Tanguy, G. Ascanio. Laminar Mixing in Stirred Tanks using Dynamic Perturbations and Non-Symmetric Geometry Conditions, in Applied Mechanics in the Americas, 6, (1999), 155-158.

[12] H. Aref. Stirring by chaotic advection, J. Fluid Mech., 143, (1984), 1-21

[13] H. Aref, S. Balachandar, Chaotic advection in a Stokes flow, Phys. Fluids, 29, (1986), 3515-3521

[14] J.G. Franjione, C.W. Leong and J.M. Ottino. Symmetries within Chaos: A Route to Effective Mixing, Phys. Fluids, 1, (1989), 1772-1783

[15] A. Hidalgo-Millán, E. Soto, R. Zenit and G. Ascanio. Effect of eccentricity on the pumping capacity in an unbaffled vessel, Can. J. Chem. Eng., 89, (2011), 1051-1058

[16] T. Espinosa-Solares, E. Brito-de la Fuente, A. Tecante and P.A. Tanguy, P. A. Flow Patterns in Rheologically Evolving Model Fluids Produced by Hybrid Dual Mixing Systems. Chem. Eng. Tech., 24, (2001), 913–918.

[17] F. Thibault. Analyse du procédé de mélange solide-liquide : application à la préparation des sauces de couchage du papier, Ph.D. Thesis, Ecole Polytechnique of Montreal, (1999).

[18] P.A. Tanguy, F. Thibault, E. Brito-de la Fuente, T. Espinosa-Solares, A Tecante. Chem. Eng. Sci., 52, (1997), 1733-1741.

[19] W. Mozaryn, L.P.B.M. Janssen, L.P.B.M. and L.L. van Dierendonck, L.L. Design rule for pumping devices in a loop polymerization reactor, Chem. Eng. Res. Des, 72 (A3), (1994), 461-464

[20] S. Foucault, G. Ascanio and P.A. Tanguy. Power Characteristics in Coaxial Mixing: Newtonian and non-Newtonian fluids, Ind. Eng. Chem. Res., 44, (2005), 5036-5043.

[21] S. Foucault, G. Ascanio and P.A.Tanguy. Coaxial Mixer Hydrodynamics with Newtonian and non-Newtonian Fluids, Chem. Eng. Tech. 27, (2004), 324-329.

[22] A.R. Khopkar, L. Fradette and P.A. Tanguy. Emulsification capability of a dual shaft mixer, Chem. Eng. Res. Des., 87, (2009), 1631-1639.

[23] M. Kuratsu, H. Nishimi, R. Yatomi, H. Sato and M. Mishima. Mixing Reactor "Superblend" Applied to Wide Range of Viscosity, Sumitomo Heavy Industries Technical Review 42, (1994), 82–85.

[24] C. Rivera, M. Heniche, K. Takenaka and P.A. Tanguy. Finite element modeling of the laminar and transition flow of the Superblend dual shaft coaxial mixer on parallel computers, Chem. Eng. Sci., 64, (2009), 4442-4456.

[25] M. Farhat, L. Fradette, H. Horiguchi, R. Yatomi and P.A. Tanguy. Experimental Investigation of Superblend Coaxial Mixer, J. Chem. Eng. Japan 42 (2009), 797–803.

[26] C. Dubois, F. Thibault , P.A. Tanguy and A. Ait-Kadi. Characterization of viscous mixing in a twin intermeshing conical helical mixer, Proc. of Fluid Mixing 5, 140, (1996), 249-258.

[27] http://www.mixers.com/ (visited on December 2011)

[28] http://www.scottmixer.com/mixers_double.php (visited on April 2011)

[29] P.A. Tanguy, F. Bertrand, R. Labrie, and E. Brito-de la Fuente. Numerical modelling of the mixing of viscoplastic slurries in a twin-blade planetary mixer, Chem. Eng. Res. Des., A4 74, (1996), 499-504.

[30] P.A. Tanguy, F. Thibault, C. Dubois, and A. Aït-Kadi. Mixing hydrodynamics in a double planetary mixer. Chem. Eng. Res. Des. 77, (1999), 318–324.

[31] G. Zhou, P.A. Tanguy and C. Dubois. Power consumption in a double planetary mixer with non-Newtonian and viscoelastic materials, Chem. Eng. Res. Des., 78, (2000), 445-453.

[32] M.J. Clifford, S.M. Cox, M.D. Finn. Reynolds number effects in a simple planetary mixer, Chem. Eng. Sci., 59, (2004), 3371-3379

[33] G. Delaplace, R.K. Thakur, L. Bouvier, C. Andre and C. Torrez. Dimensional analysis for planetary mixer: Mixing time and Reynolds numbers, Chem. Eng. Sci., 62, (2007), 1442-1447.

[34] R.K. Connelly, J. Valenti-Jordan. Mixing analysis of a Newtonian fluid in a 3D planetary pin mixer, Chem. Eng. Res. Des., 86, (2008), 1434-1440.

A new stochastic approach for advection-diffusion problems with uncertain parameters

Mofdi El-Amrani[1] **and** **Mohammed Seaid**[2] **and** **Nourdine Lanjri Zaïdi**[3]

[1] Université Abdelmalek Essaadi, Faculté des Sciences et Téchniques, Tanger, Morocco
[2] School of Engineering and Computing Sciences, University of Durham, DH1 3LE, UK
[3] Université Abdelmalek Essaadi, Faculté des Sciences, BP 2121 Tetouan, Morocco
Corresponding Author E-mail: mofdi.elamrani@gmail.com

Abstract. *We propose a new numerical method for solving advection-diffusion equations with uncertainty. The stochastic effects are introduced either in the velocity field or in the initial data resulting into a class of stochastic partial differential equations. Chaos polynomials are used to represent the stochastic processes in the considered model. This procedure, known also by spectral decomposition, results into a system of deterministic advection-diffusion equations. For each chaos coefficient, the method of characteristics is used to integrate in time its associated advection-diffusion equation. Central finite differencing is implemented for the space discretization. The proposed method is verified for an advection-diffusion equation with known analytical solution. We also apply the method for simulation of the one-dimensional viscous Burgers equation. In both examples, the method demonstrates its ability to better maintain the shape of the solution in the presence of uncertainty and shocks.*

Key words: Stochastic advection-diffusion equations, chaos polynomials, method of characteristics, stochastic simulation.

1. Introduction

In many practical problems from engineering and mathematical sciences, the involved coefficients are not precisely defined in a deterministic way. For instance, in advection-diffusion models for ground water flows where exact knowledge of the permeability of the soil, magnitude of source terms, inflow or outflow conditions are usually not known. The presence of uncertainties in these problems can be conveniently described by random fields, whose statistics are mainly inferred from experiments, compare for example [3]. This requires to include, in the partial differential equations

modeling the problem under consideration, a rational assessment of uncertainty. Consequently, this leads to the notion of stochastic partial differential equations. Recently, stochastic partial differential equations have also been used in the literature to model fluid flows [2, 7, 15], heat transfer [13], reaction-diffusion problems [12] among others.

In recent years there is a growing interests in developing efficient and accurate numerical solvers for partial differential equations with uncertainty. For the quantification of uncertainty these solvers used probabilistic modelling using either a statistical approach or a non-statistical approach. The statistical approach such as Monte Carlo simulation performs repetitive tests over a sufficiently large body of sampling, see [10]. However, this approach is the method of last resort since the attendant computational cost can be prohibitive for problems modelled using a large number of degrees of freedom. On the other hand, the non-statistical approach such as polynomial chaos representation is based upon an analytical treatment of the uncertainty. This later approach is considered in our study. It should be stressed that polynomial chaos representation has been extensively used for efficient uncertainty quantification in engineering problems of solid and fluid mechanics for example, in turbulent flows [2], flow in porous media [7], combustion [15], thermal flows [13], and elastic structures [14]. In the framework of finite element methods, the authors in [5] proposed a class of spectral approximations for the stochastic convection-diffusion problems.

Our objective in this work is to develop an efficient numerical method for solving stochastic advection-diffusion equations. The uncertainty in these equations may be included in the velocity field or diffusion coefficient or initial conditions. The key idea in our approach is to combine the polynomial chaos expansion with a method of characteristics. Using the polynomial chaos representation to treat the randomness in the equations, the method transforms the stochastic advection-diffusion problem into a system of deterministic advection-diffusion equations to be solved for each chaos coefficient in the solution expansion. Discretizations in time and space of the resulted deterministic system can be implemented in a standard manner. In the current study, we formulate a method of characteristics for the time integration along with finite difference method for the spatial discretization. The characteristic method treats the advection by using a Lagrangian tracking algorithm along the characteristic curves while keeping the convenience of a fixed computational mesh. The numerical information from the previous time level is projected from the background Eulerian mesh into the Lagrangian mesh and the required quantities in the present time level are calculated by interpolation. The significant advantage of this method is that, due to the Lagrangian treatment of advection, the CFL restriction is relaxed and the time truncation errors are reduced in the Eulerian methods. Indeed, the characteristic methods perform the temporal discretization on the total derivative by tracking fictitious fluid particles during each time step for each chaos coefficient.

Numerical results are presented for a linear advection-diffusion problem and a nonlinear Burgers equation in one space dimension. In the first case, analytical solution is available and thus it can be used for verification of convergence rates and accuracy of the proposed approach. In the other case, comparison to deterministic solutions is shown to illustrate uncertainty effects. Our method highly approximates numerical solution to these linear and nonlinear problems. The obtained results demonstrate good shock resolution without any nonphysical oscillations near the shock areas or extensive numerical dissipation and without relying on statistical methods.

This paper is organized as follows. We first give a brief description of the stochastic advection-diffusion equations considered. In section 3., we then formulate the spectral stochastic method of characteristics. This section includes the reconstruction of chaos polynomials in the spectral decomposition stage and the method of characteristics for time steeping stage. Numerical results and examples are presented in section 4.. Conclusions are summarized in section 5..

2. Stochastic Advection-Diffusion Equations

Let (Ω, Θ, μ) denotes a probability space with Ω is the set of all possible events, Θ is a σ-algebra on Ω, and μ is a probability measure in (Ω, Θ). We also use the symbol ω to indicate the dependence of a quantity on the random dimension of the problem, and $\xi_i(\omega)$ are orthogonal random variables. As shown in [1], a solution process can be expressed in terms of polynomial functions of ξ_i as

$$u = a_0\Gamma_0 + \sum_{i_1=0}^{\infty} a_{i_1}\Gamma_1(\xi_{i_1}) + \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{\infty} a_{i_1}a_{i_2}\Gamma_2(\xi_{i_1}, \xi_{i_2}) + \ldots, \tag{2.1}$$

where $\Gamma_n(\xi_{i_1}, \ldots, \xi_{i_n})$ denote the chaos polynomials of order n in variables $(\xi_{i_1}, \ldots, \xi_{i_n})$, see [22] among others. These polynomials are usually generalized multi-dimensional Hermite polynomials of independent variables that are measurable functions with respect to the Wiener measure. In particular, when independent variables are identified as the Gaussian vector $\xi = (\xi_{i_1}, \ldots, \xi_{i_n})$, one recovers the familiar expression of the expectation as

$$\langle f \rangle = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{-\infty}^{\infty} f(\xi) e^{-\frac{|\xi|^2}{2}}\, d\xi, \tag{2.2}$$

where $|\xi| = \sum_{i=1}^{n} \xi_i^2$. For simplicity in presentation, it is also convenient to introduce a one-to-one mapping between the set of indices appearing in (2.1) and the set of ordered indices, and rewrite the sum (2.1) in single-index form as

$$u = \sum_{i=0}^{\infty} u_i\psi_i, \tag{2.3}$$

where ψ_i denotes the polynomial chaos in single-index notation. The set $\{\psi_i\}_{i=1}^{\infty}$ forms a complete basis in the space of second-order random variables, see for instance [1]. These polynomials are orthogonal in the sense that

$$\langle \psi_i\psi_j \rangle = 0, \qquad \forall\, i \neq j. \tag{2.4}$$

In this paper, we restrict our study to the one-dimensional stochastic advection-diffusion equation

$$\frac{\partial u}{\partial t} + v(t, x, \omega)\frac{\partial u}{\partial x} - \nu\frac{\partial^2 u}{\partial x^2} = 0, \quad \text{in} \quad (0, T] \times \mathcal{D} \times \Omega, \tag{2.5}$$

where $\mathcal{D} \subset \mathbb{R}$ is a bounded spatial domain and $(0, T]$ is the time interval. Here, $u(t, x, \omega)$ is the required solution, $v(t, x, \omega)$ is the velocity field and ν is the diffusion coefficient. The equation (2.5) is equipped with the initial condition

$$u(0, x, \omega) = u^0(x, \omega), \quad \text{in} \quad \mathcal{D} \times \Omega, \tag{2.6}$$

with $u^0(x,\omega)$ is a given initial data. Moreover, in order to formulate a well-posed mathematical problem, boundary conditions are required for the equation (2.5). These conditions are problem dependent and their discussion is postponed for section 4. where numerical examples are discussed. Note that, the stochastic functions $v(t,x,\omega)$, $\nu(\omega)$ and $u^0(x,\omega)$ can be independent white Gaussian noises with bounded and continuous expectations or simply,

$$v(t,x,\omega) = \bar{v}(t,x) + v'(t,x)\xi, \tag{2.7}$$

with a similar formula for $\nu(\omega)$ and $u^0(x,\omega)$. In (2.7), ξ is a Gaussian variable with unit variance, $\bar{v}(t,x)$ and $v'(t,x)$ represent the mean and standard deviation of the velocity field, respectively. It should be stressed that the velocity v can be the solution itself such as the case of the Burgers equation.

3. Spectral Stochastic Method of Characteristics

A solution procedure for the advection-diffusion equations (2.5)-(2.6) requires discretization of random, space and time variables. In the current work, we discretize the random variable using the polynomial chaos expansion, the spatial discretization is carried out using a second-order finite difference method, and the time integration is performed using a TVD Runge-Kutta scheme along the characteristic curves. In what follows, we discuss each discretization stage separately.

3.1. Spectral Decomposition

For computational purposes, the representation (2.2) is truncated by retaining polynomials of order $\leq p$, with p is a prescribed value as

$$u(t,x,\omega) = \sum_{i=0}^{P} u_i(t,x)\psi_i(\xi), \tag{3.1}$$

with $P+1$ is the total number of polynomial chaos of order $\leq p$. For a space with n stochastic dimensions,

$$P + 1 = \frac{(p+n)!}{p!n!}.$$

Due to the orthogonality property (2.4), the coefficients of the polynomial chaos expansion u_i satisfy

$$u_i = \frac{\langle u\psi_i \rangle}{\langle \psi_i^2 \rangle}, \qquad i = 0, 1, \ldots, P. \tag{3.2}$$

It is clear that, the accuracy of the polynomial chaos expansion depends on the number P such that high accuracy is obtained for large values of P. However, large values of P can lead to high computational cost and may limit the efficiency of the overall procedure. For one-dimensional problems (*i.e.* $n = 1$), the expectation with respect to the Gaussian measure (2.2) reduces to

$$\langle f \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(\xi) e^{-\frac{\xi^2}{2}}\, d\xi, \tag{3.3}$$

Table 1: The first five polynomial chaoses and their corresponding variances.

Order i	Polynomial $\psi_i(\xi)$	Variance $\langle\psi_i^2\rangle$
$i=0$	$\psi_0(\xi)=1$	$\langle\psi_0^2\rangle=1$
$i=1$	$\psi_1(\xi)=\xi$	$\langle\psi_1^2\rangle=1$
$i=2$	$\psi_2(\xi)=\xi^2-1$	$\langle\psi_2^2\rangle=2$
$i=3$	$\psi_3(\xi)=\xi^3-3\xi$	$\langle\psi_3^2\rangle=6$
$i=4$	$\psi_4(\xi)=\xi^4-6\xi^2+3$	$\langle\psi_4^2\rangle=24$

and the five polynomials ψ_i and their corresponding variance are listed in Table 1. Other higher-order chaos polynomials can be derived in a similar manner using a recurrence formula.

Hence, by representing the solution $u(t,x,\omega)$, the velocity field $v(t,\mathbf{x},\omega)$, the diffusion coefficient $\nu(\omega)$ and the initial data $u^0(x,\omega)$ as

$$u(t,x,\omega)=\sum_{i=0}^{P}u_i(t,x)\psi_i(\xi),\qquad v(t,x,\omega)=\sum_{i=0}^{P}v_i(t,x)\psi_i(\xi),$$

$$\nu(\omega)=\sum_{i=0}^{P}\nu_i\psi_i(\xi),\qquad u^0(x,\omega)=\sum_{i=0}^{P}u_i^0(x)\psi_i(\xi),$$

the advection-diffusion problem (2.5) transforms to

$$\sum_{i=0}^{P}\frac{\partial u_i}{\partial t}\psi_i+\sum_{i=0}^{P}\sum_{j=0}^{P}v_j(t,x)\psi_j\frac{\partial u_i}{\partial x}\psi_i-\sum_{i=0}^{P}\sum_{j=0}^{P}\nu_j\psi_j\frac{\partial^2 u_i}{\partial x^2}\psi_i=0, \tag{3.4}$$

subject to the initial condition

$$\sum_{i=0}^{P}u_i(0,x)\psi_i=\sum_{i=0}^{P}u_i^0(x)\psi_i. \tag{3.5}$$

Multiplying the equations (3.4) and (3.5) by ψ_k, taking their expectation and using the orthogonality relation (2.4), the problem statement becomes:

For each $k=0,1,\ldots,P$ solve the following system of advection-diffusion equations

$$\begin{aligned}\frac{\partial u_k}{\partial t}+\sum_{i=0}^{P}\sum_{j=0}^{P}\frac{\langle\psi_k\psi_j\psi_i\rangle}{\langle\psi_k^2\rangle}v_j(t,x)\frac{\partial u_i}{\partial x}-\sum_{i=0}^{P}\sum_{j=0}^{P}\frac{\langle\psi_k\psi_j\psi_i\rangle}{\langle\psi_k^2\rangle}\nu_j\frac{\partial^2 u_i}{\partial x^2} &= 0,\\ u_k(0,x) &= u_k^0(x).\end{aligned} \tag{3.6}$$

Once the chaos coefficients $u_k(t,x)$ are obtained from the equations (3.6), the solution u of the original advection-diffusion problem (2.5) is recovered by

$$u(t,x,\omega) = \sum_{i=0}^{P} u_i(t,x)\psi_i(\xi). \tag{3.7}$$

The main advantage of the polynomial chaos expansion is the fact that the stochastic advection-diffusion problem (2.5) is replaced by a finite sequence of deterministic system of coupled advection-diffusion equations. From a computational view point, the chaos coefficients $v_i(t,x)$, ν_i, $u_i^0(x)$, and the expectations appeared in (3.6) can be computed in advance and stored to be used in the deterministic solver.

3.2. Method of Characteristics

The first equation in the polynomial chaos expansion (3.6) can be rewritten in an advective form as

$$\frac{\partial u_k}{\partial t} + V_k(t,x)\frac{\partial u_k}{\partial x} = \sum_{i=0}^{P} D_{ik}\frac{\partial^2 u_i}{\partial x^2} - \sum_{i=0}^{P} C_{ik}(t,x)\frac{\partial u_i}{\partial x}, \quad \text{in} \quad (0,T]\times\mathcal{D}, \tag{3.8}$$

where

$$V_k(t,x) = \sum_{j=0}^{P} \frac{\langle\psi_k\psi_j\psi_k\rangle}{\langle\psi_k^2\rangle} v_j(t,x), \qquad C_{ik}(t,x) = \sum_{\substack{j=0\\ j\neq k}}^{P} \frac{\langle\psi_k\psi_j\psi_i\rangle}{\langle\psi_k^2\rangle} v_j(t,x),$$

$$D_{ik} = \sum_{j=0}^{P} \frac{\langle\psi_k\psi_j\psi_i\rangle}{\langle\psi_k^2\rangle}\nu_j. \tag{3.9}$$

The method we consider in this paper consists of two fractional steps. The first step is the Lagrangian interpretation for the advection part in (3.8) by the method of characteristics, while the second step uses the Eulerian coordinates for discretization of the diffusion part in (3.8). We first consider the homogeneous advective part of the problem (3.8)

$$\frac{Du_k}{Dt} := \frac{\partial u_k}{\partial t} + V_k(t,x)\frac{\partial u_k}{\partial x} = 0. \tag{3.10}$$

Recall that $\frac{Du}{Dt}$ measures the rate of change of a function u following the trajectories of the flow particles. The fundamental idea of the method of characteristics is to impose a regular grid at the new time level and to backtrack the flow trajectories to the previous time level. At the old time level, the quantities that are needed are evaluated by interpolation from their known values on a regular grid.

Let the time interval $[0,T]$ be divided into N subintervals $[t_n, t_{n+1}]$ of length Δt such that $t_n = n\Delta t$ and $T = N\Delta t$. Following [16, 4], the characteristic curves of the equation (3.10) are

solutions of the initial-value problem

$$\begin{aligned} \frac{dX_k(\tau;t_{n+1},x)}{d\tau} &= V_k\left(X_k(\tau;t_{n+1},x)\right), \quad \tau\in[t_n,t_{n+1}], \\ X_k(t_{n+1};t_{n+1},x) &= x\,. \end{aligned} \tag{3.11}$$

Notice that $X_k(\tau;t_{n+1},x)$ is the departure point at time τ of a particle that will arrive at point x in time t_{n+1}. The method of characteristics does not follow the flow particles forward in time, as the Lagrangian schemes do, instead it traces backwards the position at time t_n of particles that will reach the points of a fixed mesh at time t_{n+1}. By doing so, the method avoids the grid distortion difficulties that the conventional Lagrangian schemes have.

The solutions of (3.11) can be expressed as

$$X_k(t_n;t_{n+1},x) = x - \int_{t_n}^{t_{n+1}} V_k\left(X_k(\tau;t_{n+1},x)\right)d\tau, \quad k=0,1,\ldots,P. \tag{3.12}$$

For a velocity field given explicitly independent of the solution u_k the integral in (3.12) can be determined analytically. In other cases, this integral can be calculated using a fourth-order explicit Runge-Kutta scheme which is accurate enough to maintain a particle on its curved trajectory.

Once the characteristics curves $X_k(t_n;t_{n+1},x)$ are known, the method of characteristics advects the solution of (3.10) at instant t_{n+1} as

$$\hat{u}_k(t_{n+1},x) := u_k\left(t_n, X_k(t_n;t_{n+1},x)\right)\,. \tag{3.13}$$

In general, the departure points $X_k(t_n;t_{n+1},x)$ do not coincide with the spatial position of a grid-point. A requirement is then that the scheme to compute $X_k(t_n;t_{n+1},x)$ be provided with a search-locate algorithm to find the host element where such point is located. For structured grids this step can be as simple as index checking or *ad hoc* searching. Assuming a suitable approximation is made for $X_k(t_n;t_{n+1},x)$, the solution $\hat{u}_k(t_{n+1},x)$ in (3.13) should be obtained by interpolation from known values at the gridpoints in the host cell of the departure points. The interpolation procedure we used in this paper is the cubic spline interpolation most commonly used in practice. Other interpolation procedures can also applied.

To differentiate in the characteristic direction $s=s(x)$ associated with the total derivative $\frac{D}{Dt}$, we first define

$$\gamma(x) = \sqrt{1+V_k^2}.$$

Then,

$$\frac{\partial u_k}{\partial s} = \frac{1}{\gamma(x)}\frac{Du_k}{Dt},$$

and the equation in (3.8) is transformed to

$$\gamma(x)\frac{\partial u_k}{\partial s} = \sum_{i=0}^{P} C_{ik}\frac{\partial u_i}{\partial x} - \sum_{i=0}^{P} D_{ik}\frac{\partial^2 u_i}{\partial x^2}. \tag{3.14}$$

Based on the approach presented in [4], the characteristic derivative is approximated by

$$\begin{aligned}\gamma(x)\frac{\partial u_k}{\partial s} &\approx \gamma(x)\frac{u_k(t_{n+1},x)-u_k(t_n,x)}{\sqrt{(x-X_k)^2+(\Delta t)^2}},\\ &= \frac{u_k(t_{n+1},x)-u_k(t_n,x)}{\Delta t},\\ &\approx \frac{\partial u_k}{\partial t}.\end{aligned}$$

Note that the main convective term in (3.8), which causes many difficulties in most of Eulerian-based discretizations, has been moved from the left-hand side in (3.14) using the Lagrangian step in our formulation.

3.3. Solution of the Diffusion Stage

In the current work, we formulate a second-order finite difference method for the spatial discretization. Thus, the spatial domain $\mathcal{D}$ is discretized in cells $[x_j, x_{j+1}]$ with uniform dimensions $\Delta x = x_{j+1} - x_j$. We use the notations $u_{k,j} = u_k(t, x_j)$ and $u^n_{k,j} = u_k(t_n, x_j)$, and we discretize the diffusion operator in (2.5) using the central difference formula. Hence, we obtain the semi-discrete problem

$$\frac{du_{k,j}}{dt} = \sum_{i=0}^{P} D_{ik}\frac{u_{i,j+1}-2u_{i,j}+u_{i,j-1}}{(\Delta x)^2} - \sum_{i=0}^{P} C_{ik}\frac{u_{i,j+\frac{1}{2}}-u_{i,j-\frac{1}{2}}}{\Delta x}. \tag{3.15}$$

A second-order upwind scheme is used for the reconstruction of numerical fluxes in (3.15)

$$u_{k,j+\frac{1}{2}} = \frac{u^R_{k,j+\frac{1}{2}}+u^L_{k,j+\frac{1}{2}}}{2}, \qquad u_{k,j-\frac{1}{2}} = \frac{u^R_{k,j-\frac{1}{2}}+u^L_{k,j-\frac{1}{2}}}{2}, \tag{3.16}$$

where the left and right numerical fluxes are reconstructed as

$$\begin{aligned}u^L_{k,j+\frac{1}{2}} &= u_{k,j} + \Phi\left(r_{k,j}\right)\frac{u_{k,j+1}-u_{k,j}}{2},\\ u^L_{k,j-\frac{1}{2}} &= u_{k,j-1} + \Phi\left(r_{k,j-1}\right)\frac{u_{k,j}-u_{k,j-1}}{2},\\ u^R_{k,j+\frac{1}{2}} &= u_{k,j+1} - \Phi\left(r_{k,j+1}\right)\frac{u_{k,j+2}-u_{k,j+1}}{2},\\ u^R_{k,j-\frac{1}{2}} &= u_{k,j} - \Phi\left(r_{k,j}\right)\frac{u_{k,j+1}-u_{k,j}}{2},\end{aligned}$$

with the slope $r_{k,j}$ is given by

$$r_{k,j} = \frac{u_{k,j}-u_{k,j-1}}{u_{k,j+1}-u_{k,j}},$$

and Φ is the Van-Leer slope limiter function [11]

$$\Phi(r) = \frac{|r|+r}{1+|r|}.$$

The solution procedure for equations (3.6) is complete when a time integration of semi-discrete equations (3.15) is selected. This stage can be handled by any implicit ordinary differential equations (ODE) solver, since they are computationally without risk by virtue of their accuracy and linear unconditionally stability. This allows for larger time steps in the integration process. However, due to the large set of linear system of algebraic equations at each time step, these methods may be computationally inefficient. As an alternative, we use an explicit Runge-Kutta method studied in [19]. Let us rewrite the equations (3.15) in a compact ODE form as

$$\begin{aligned} \frac{du_{k,j}}{dt} &= F\left(u_{k,j}\right), \qquad t \in (0,T), \\ u_{k,j}(0) &= u^0_{k,j}, \end{aligned} \tag{3.17}$$

The procedure to advance the solution from the time t_n to the next time t_{n+1} can be carried out as

$$\begin{aligned} U^{(1)}_{k,j} &= u^n_{k,j} + \Delta t F(u^n_{k,j}), \\ U^{(2)}_{k,j} &= \frac{3}{4}u^n_{k,j} + \frac{1}{4}U^{(1)}_{k,j} + \frac{1}{4}\Delta t F(U^{(1)}_{k,j}), \\ u^{n+1}_{k,j} &= \frac{1}{3}u^n_{k,j} + \frac{2}{3}U^{(2)}_{k,j} + \frac{2}{3}\Delta t F(U^{(2)}_{k,j}). \end{aligned} \tag{3.18}$$

This class of explicit time integration schemes has become popular in computational fluid dynamics, see for example [8]. The main feature of this method lies on the fact that (3.18) is a convex combination of first-order Euler steps which exhibit strong stability properties. Therefore, the scheme (3.18) is TVD, third-order accurate in time, and stable under the usual deterministic hyperbolic and parabolic CFL conditions

$$\max_{i,k} \max_{x} \left\{C_{ik}(t_n, x)\right\} \frac{\Delta t}{\Delta x} \leq 1, \tag{3.19}$$

and

$$\max_{i,k} D_{ik} \frac{\Delta t}{(\Delta x)^2} \leq \frac{1}{4}, \tag{3.20}$$

respectively.

4. Numerical Examples

We examine the performance of the proposed method for a class of stochastic advection-diffusion problems in one space dimension. For each test example we perform $100,000$ realizations and statistical moments such as mean and standard deviation are computed. In all our simulations we have used variable time steps Δt adjusted at each step according to the CFL conditions (3.19) and (3.20).

Table 2: Spatial error-norms for the pure advection problem using two values of σ.

	$\sigma = 0.01$			$\sigma = 0.5$		
M	L^∞-error	L^1-error	L^2-error	L^∞-error	L^1-error	L^2-error
21	0.643E-01	0.818E-01	0.613E-01	0.323E-01	0.314E-01	0.246E-01
41	0.233E-01	0.206E-01	0.158E-01	0.141E-01	0.797E-02	0.705E-02
81	0.815E-02	0.461E-02	0.389E-02	0.116E-01	0.360E-02	0.395E-02
161	0.240E-02	0.832E-03	0.847E-03	0.404E-02	0.200E-02	0.185E-02
321	0.681E-03	0.303E-03	0.257E-03	0.506E-02	0.170E-02	0.155E-02

4.1. Pure Advection Problem

We solve the stochastic advection-diffusion equation (2.5) with $\nu = 0$ and

$$v(t, x, \omega) = \overline{v} + \sigma\xi, \tag{4.1}$$

where $\overline{v} = 1$ is the mean velocity and σ is a parameter to control the amplitude of the stochastic perturbation. Periodic boundary conditions are used and the initial condition is set to

$$u(t, x, \omega) = \sin\big((x + 1)\,\pi\big). \tag{4.2}$$

It is easy to verify that

$$\overline{u}(t, x) = \sin\big((x + 1 - \overline{v}t)\,\pi\big)e^{-\frac{\pi^2\sigma^2t^2}{2}}, \tag{4.3}$$

is the expected analytical solution for the pure advection problem (2.5)-(4.1). The exact solution (4.3) is used to evaluate the expected error function at time t_n as

$$e_i^n = \langle u_i^n \rangle - \overline{u}(t_n, x_i), \tag{4.4}$$

where $\overline{u}(t_n, x_i)$ and $\langle u_i^n \rangle$ are respectively, the exact and expectation of numerical solutions at grid-point x_i and time t_n. The following discrete error-norms are defined

$$\|e\|_{L^\infty} = \max_{1\le i\le M} |e_i^n|, \qquad \|e\|_{L^1} = \Delta x \sum_{i=1}^{M} |e_i^n|, \qquad \|e\|_{L^2} = \left(\Delta x \sum_{i=1}^{M} |e_i^n|^2\right)^{\frac{1}{2}},$$

with M is the total number of gridpoints. In this example, the total number of chaos polynomials is $P = 5$, the spatial domain $\mathcal{D} = [-1, 1]$ and the time period $T = 1$.

In Table 2 we display the error-norms for two values of σ using different numbers of gridpoints M. Observe that larger stochastic perturbations are expected for larger values of σ. It is clear that increasing the number of gridpoinds in the space domain results in a decrease of all error-norms for both values of σ. A slower decrease in these error-norms has been detected in the case of $\sigma = 0.5$

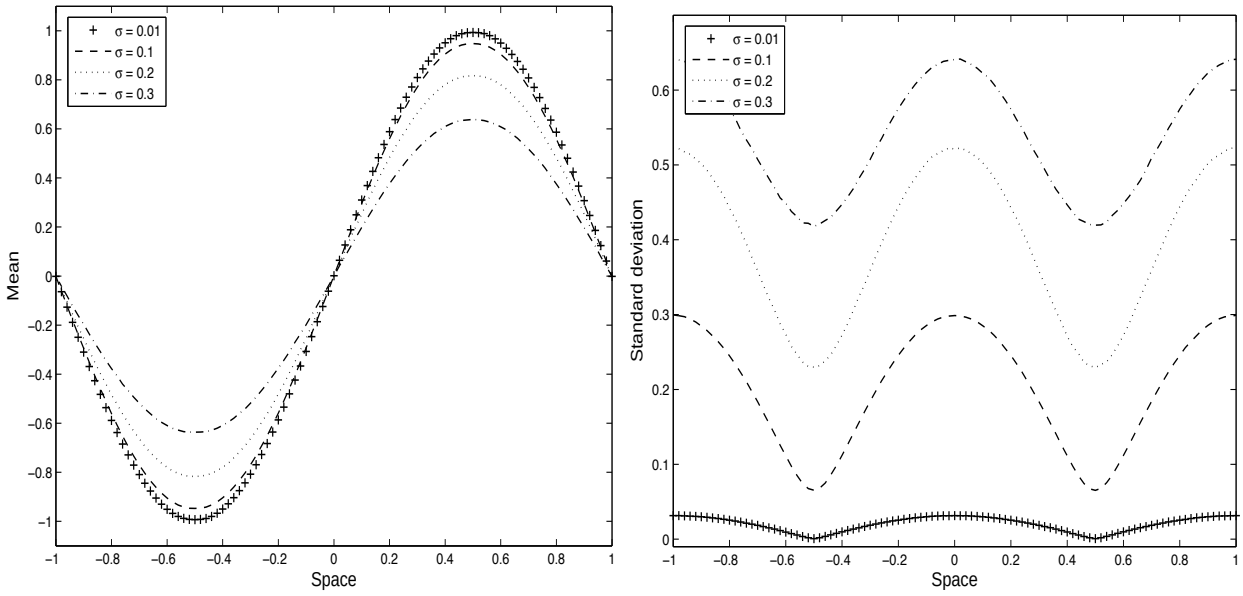


Figure 1: Expectation (left) and standrad deviation (right) for numerical solutions of the pure advection problem at different values of σ.

than those obtained using $\sigma = 0.01$. It is worth remarking that for large values of σ, higher order chaos polynomials have to be used in order to model uncertainties.

In Figure 1 we show the expectation and the standrad deviation of numerical solutions obtained using different values of σ at time $t = 1$ and $M = 100$. As can be seen from this figure, smaller standard deviations are obtained for smaller values of stochastic magnitudes σ. Our method accurately resolves this pure advection problem without exhibiting non-physical oscillations. It is evident that as $t \to \infty$ the expected exact solution tends to zero. This behavior has also been detected for the numerical expected solution and the larger σ, the faster $\langle u_i^n \rangle$ goes to zero.

4.2. Advection-Diffusion Problem

This example considers the stochastic advection-diffusion problem (2.5) in the space domain $\mathcal{D} = [0, 2]$ and subject to the following deterministic initial condition

$$u^0(t,x) = \begin{cases} 1, & \text{if} \quad x \in [a,b], \\ \\ 0, & \text{otherwise} \end{cases} \tag{4.5}$$

and homogeneous Dirichlet boundary conditions

$$u(t, x = 0, \omega) = u(t, x = 2, \omega) = 0. \tag{4.6}$$

Here, the velocity field and the viscosity coefficient are stochastic given by

$$\nu(t,x,\omega) = \overline{\nu} + \sigma_1 \xi, \qquad v(t,x,\omega) = \overline{v} + \sigma_2 \xi, \tag{4.7}$$

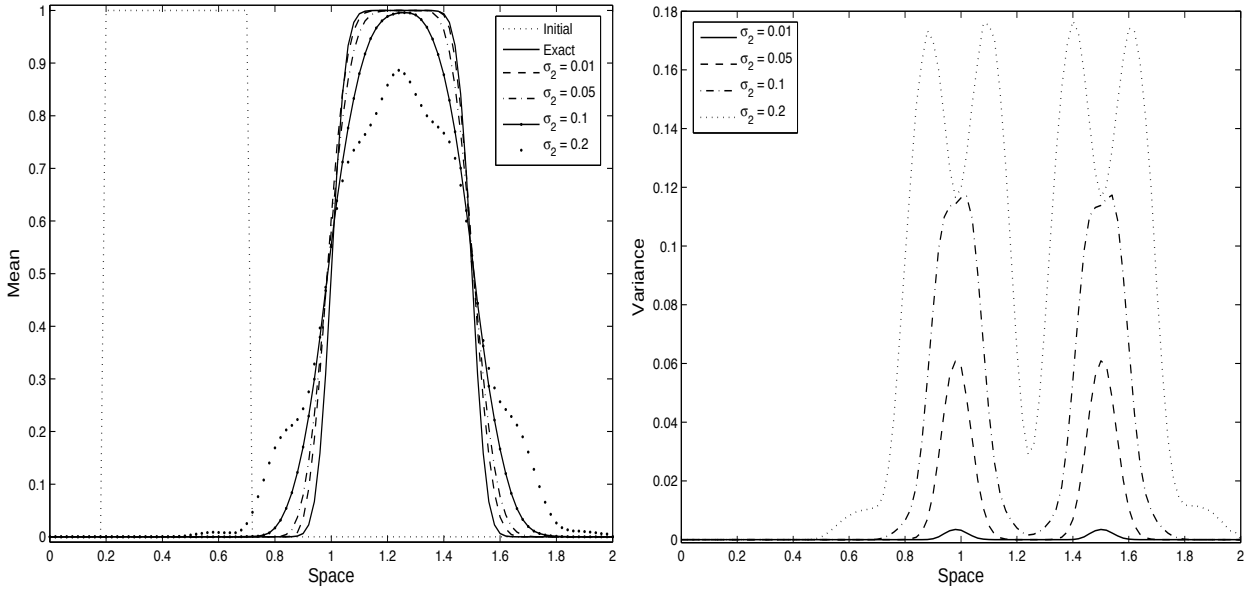


Figure 2: Expectation (left) and variance (right) for numerical solutions of the advection-diffusion problem at $t = 0.8$ and different values of σ_2 using $\overline{\nu} = 10^{-3}$.

where $\overline{\nu}$ is the viscosity coefficient and $\overline{v} = 1$ is the mean velocity field. The parameters σ_1 and σ_2 are used to control the amplitude of the stochastic perturbations. The analytical expected solution of this problem is given by

$$\overline{u}(t,x) = \frac{1}{2}\left(\mathrm{erf}\left(\frac{x-\overline{v}t-a}{\sqrt{4\overline{\nu}t}}\right) - \mathrm{erf}\left(\frac{x-\overline{v}t-b}{\sqrt{4\overline{\nu}t}}\right)\right), \tag{4.8}$$

with $\mathrm{erf}(x)$ denotes the error function defined as

$$\mathrm{erf}(x) = \frac{1}{\sqrt{\pi}}\int_0^x e^{-s^2}\, ds.$$

In this example we set $\sigma_1 = 0$, $a = 0.2$, $b = 0.7$, and the mean diffusion coefficients $\overline{\nu} = 10^{-3}$ and $\overline{\nu} = 10^{-2}$ are used in our simulations. The spatial domain is discretized in 100 gridpoints and numerical results are displayed at time $t = 0.8$. Figure 2 shows the numerical expectation and the variance obtained using $\overline{\nu} = 10^{-3}$ and different values of σ_2. The results obtained using $\overline{\nu} = 10^{-2}$ are presented in Figure 3. Along these figures we have included the initial condition and the expected analytical solutions.

For small values of σ_2, the exact and numerical expected solutions are in good agreement. Increasing the value of σ_2, the numerical expected solutions deviate from the exact expected solution. The uncertainty in this advection-diffusion problem seems to play an extra diffusion role in the system. However, in contrast to the physical diffusion which has a smoothing effects, the stochastic inputs do not illustrate these features, compare the results for variances obtained using $\overline{\nu} = 10^{-3}$ and $\overline{\nu} = 10^{-2}$ in Figure 2 and Figure 3. The proposed method performs well for this

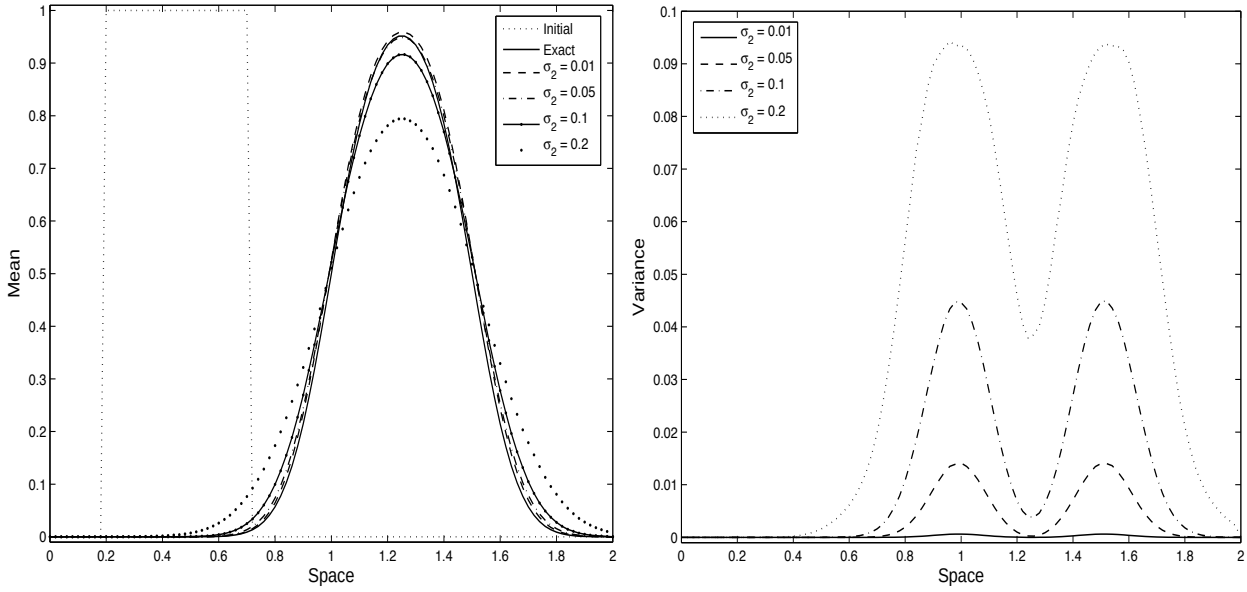


Figure 3: Expectation (left) and variance (right) for numerical solutions of the advection-diffusion problem at $t = 0.8$ and different values of σ_2 using $\overline{\nu} = 10^{-2}$.

unsteady advection-diffusion example and reproduces accurate solutions without requiring special treatment of the source terms or complicated representation of the uncertainty.

4.3. Viscous Burgers Problem

Our final test example consists on solving the following viscous Burgers problem with a random force term

$$\frac{\partial u}{\partial t} + u(t,x,\omega)\frac{\partial u}{\partial x} - \nu\frac{\partial^2 u}{\partial x^2} = f(\omega), \quad \text{in} \quad (0,T]\times\mathcal{D}\times\Omega, \tag{4.9}$$

where $\mathcal{D} = [0,1]$ and the force term f is defined as

$$f(\omega) = \sigma\xi, \qquad \text{with} \quad \sigma = \frac{1}{2}\cos(4\pi x)\,. \tag{4.10}$$

Periodic boundary conditions are imposed and the initial condition is given by

$$u(0,x) = \frac{1}{2}\left(e^{\cos(2\pi x)} - 1.5\right)\sin\Big(2\pi(x+0.37)\Big).$$

A similar problem has been studied in [9] using Eulerian-based techniques. In Figure 4 we present the numerical results at time $t = 0.8$ using a mesh of 80 gridpoints and chaos polynomials with different degrees. The influence of the truncation degree P on the numerical solutions can clearly be seen from these results. For instance, results obtained using $P = 8$ show large deviation from those obtained using $P = 2$. A smaller deviation is observed between results obtained using $P = 8$ and $P = 4$. For the stochastic force considered, no difference has been detected between numerical

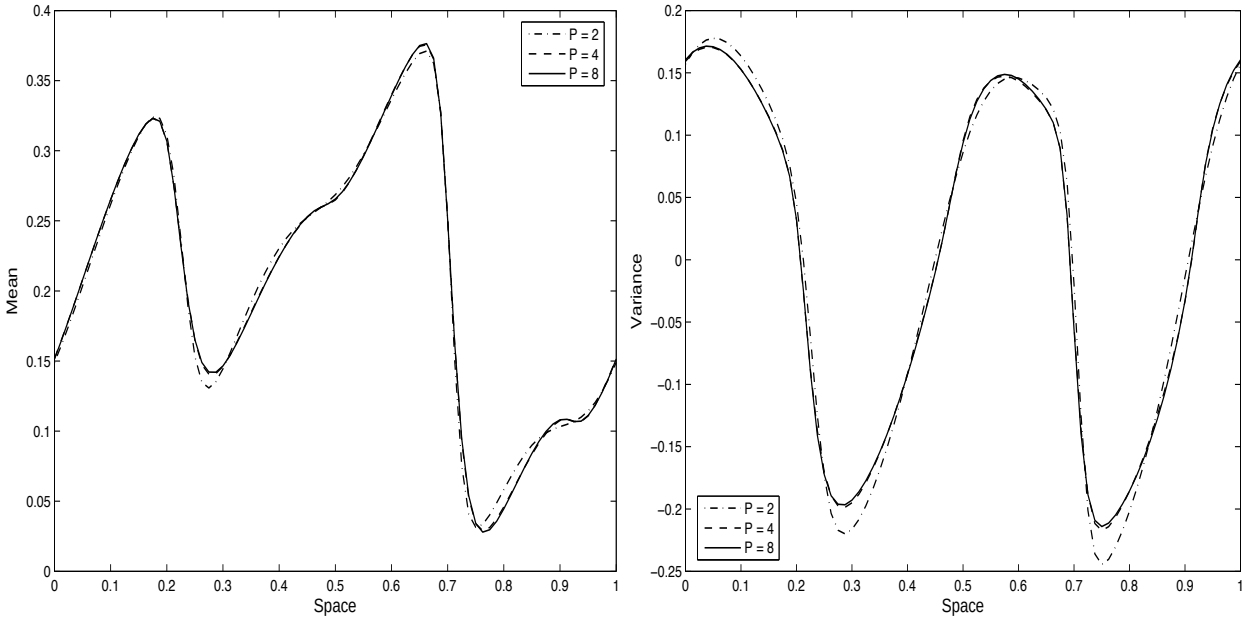


Figure 4: Expectation (left) and variance (right) for numerical solutions of the viscous Burgers problem at $t = 0.8$ using chaos polynomials with different degrees.

results obtained using $P = 8$ and chaos polynomials with degree higher than 8. However, for large values of the amplitude of random forcing σ, higher order chaos polynomials have to be used in order to accurately model uncertainties.

Next, we compare the numerical solution obtained for the stochastic viscous Burgers equation (4.9) and its deterministic counterpart (*i.e.* $\sigma = 0$). The obtained results are displayed in Figure 5 at $t = 0.8$. It is evident that the stochastic solution exhibits different behavior than that illustrated for the deterministic solution. For example, the shock is better captured in the deterministic solution than its stochastic counterpart, and diffusion is more pronounced in the stochastic solution than the deterministic one. Our method accurately resolves this nonlinear stochastic Burgers problem.

5. Conclusions

We have developed a new numerical method based on combining the polynomial chaos expansion with the method of characteristics for solution of stochastic advection-diffusion equations. This method exploits the interesting features offered by both techniques to construct an efficient algorithm for numerical treatment of stochastic advection-diffusion problems. The important advantage of the new method is that it transforms the stochastic advection-diffusion equation to a sequence of deterministic system of advection-diffusion equations which can numerically be solved using method of characteristics without relying on Monte Carlo approach. In addition, the convective term that has to be treated carefully in most of Eulerian-based methods has been moved from the new method by using the method of characteristics to interpret the transport nature of the equation.

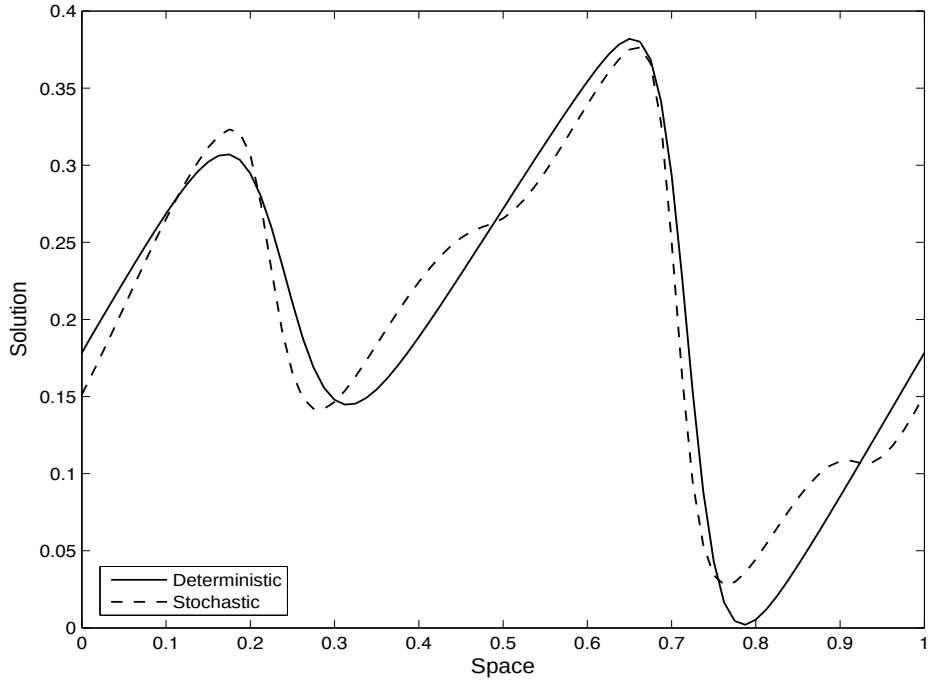


Figure 5: Comparison between the deterministic and stochastic solutions for the viscous Burgers problem at $t = 0.8$.

Several numerical examples with uncertainty in advection field or diffusion rate were considered to test the accuracy of the method. Statistical moments of the computed solutions are illustrated for the considered test examples. The obtained results show high accuracy, good shock resolution and less numerical dissipation. The efficiency of the method comes from the fast convergence of the polynomial chaos expansion. In all examples presented, we have found that a few terms in the chaos expansion are enough to ensure an accurate representation. Finally, we should point out that the implementation of our method has been restricted to stochastic advection-diffusion problems involving the Gaussian distribution as representation of randomness. For this class of problems, the Hermite polynomials are suitable for its chaos expansion. However, in more general applications involving different random distributions, other polynomial chaos representations have to be used. Extension of our method for multi-dimensional problems using generalized chaos expansions is currently underway.

References

[1] R.H. Cameron, W.T. Martin, The Orthogonal Development of Nonlinear Functionals in Series of Fourier-Hermite Functionals, Ann. Math. **48** (1947) 385–392.

[2] A.J. Chorin, Gaussian Fields and random Flow. J. Fluid Mech. **63** (1947) 21–32.

[3] G. Dagan, S.P. Neuman. *Subsurface Flow and Transport: A Stochastic Approach*. International Hydrology Series, Canbridge University Press, 1997.

[4] J. Douglas, T.F. Russell, Numerical Methods for Convection Dominated Diffusion Problems Based on Combining the Method of Characteristics with Finite Elements or Finite Differences, SIAM J. Numer. Anal. **19** (1982) 871–885.

[5] M. El-Amrani, M. Seaïd, Spectral Stochastic Semi-Lagrangian Method for Convection-Diffusion Equations with Uncertainty, J. Scientific Computing. **39** (2009) 317–393.

[6] R. Ghanem and P. Spanos. *Stochastic Finite Elements: A Spectral Approach*. Springer-Verlag, New York, 1991.

[7] R. Ghanem, Probabilistic Characterization of Transport in Herterogeneous Media. Comput. Methods Appl. Mech. Eng. **158** (1998) 199–220.

[8] S. Gottlieb, C.W. Shu and E. Tadmor. Strong Stability Preserving High Order Time Integration Methods. *SIAM Rev.*, 43:89–112, 2001.

[9] T.Y. Hou, W. Luo, B. Rozovski, H.M. Zhou: "Wiener Chaos expansions and numerical solutions of randomly forced equations of fluid mechanics". *J. Comput. Physics.*, Vol. **216**. pp:687-706, 2006.

[10] J.E. Hurtado, A.H. Barbat: "Monte Carlo Techniques in Computational Stochatsic Mechanics". *Arch. Comput. Methods Engrg.* Vol. **5**. pp:3-29, 1998.

[11] J. LeVeque Randall: "Numerical Methods for Conservation Laws", *Lectures in Mathematics ETH Zürich*, 1992.

[12] H. Manouzi, M. Seaïd, M. Zahri: "Wick-Stochastic Finite Element Solution of Reaction-Diffusion Problems". *J. Comp. Applied Math.*, Vol. **203**. pp:516-532, 2007.

[13] O. P. Le Maître, O. M. Knio, B. J. Debusschere, H. N. Najm and R. G. Ghanem: "A multigrid solver for two-dimensional stochastic diffusion equations". *Comput. Methods Appl. Mech. Eng.*, Vol. **192**. pp:41-42, 2003.

[14] H.G. Matthies, C.E. Brenner, C.G. Bucher and C.G. Soares: "Uncertainties in probabilistic numerical analysis of structures and solids-Stochastic finite elements". *Structural Safety.*, Vol. **19**. pp:283-336, 1997.

[15] B.D. Phenix, J.L. Dinaro, M.A. Tatang, J.W. Tester, J.B. Howard, G.J. McRae: "Incorporation of parametric Uncertainty into Complex Kinetic Mechanisms: Application to Hydrogen Oxidation in Supercritical Water". *Combust. Flame.*, Vol. **112**. pp:132-146, 1998.

[16] A. Robert, A stable Numerical Integration Scheme for the Primitive Meteorological Equations, Atmos. Ocean **19** (1981) 35–46.

[17] M. Seaïd, On the Quasi-Monotone Modified Method of Characteristics for Transport-Diffusion Problems with Reactive Sources, Comp. Methods in App. Math. **2** (2002) 186–210.

[18] M. Seaïd, Semi-Lagrangian Integration Schemes for Viscous Incompressible Flows, Comp. Methods in App. Math. **4** (2002) 1–18.

[19] C.W. Shu. Total Variation Diminishing Time Discretizations. *SIAM J. Sci. Stat. Comput.*, 9:1073–1084, 1988.

[20] A. Staniforth, J. Côté, Semi-Lagrangian Integration Schemes for the Atmospheric Models: A Review, Wea. Rev. **119** (1991) 2206–2223.

[21] C. Temperton, A. Staniforth, An Efficient Two-Time-Level Semi-Lagrangian Semi-Implicit Integration Scheme, Quart. J. Roy. Meteor. Soc. **113** (1987) 1025–1039.

[22] N. Wiener. The Homogeneous Chaos. *Am. J. Math.*, 60:897–936, 1938.

Instructions for authors

Frontiers in Science and Engineering an International Journal edited by Hassan II Academy of Science and Technology uses author-supplied PDFs for all online and print publication.

1. Type of Manuscripts
The FSE Journal publishes the following article types :
Reviews/State of the art, usually through Academy invitation and organized into themed issues, report on recent advances in science and technology. 20 pages maximum according to the format given.
Original Research Papers contain innovative and hypothesis-driven research; supported by sound experimental design, methodology, proofs, and data interpretation. 16 pages maximum according to the format given.
Letters to the Editor: May be submitted by readers commenting articles already published by the Journal. 1 to 2 pages maximum according to the format given.

2. Organization of the Manuscript
The entire manuscript, including mathematical equations, flow-sheets, chemical structures, tables, and figures must be prepared in electronic form and submitted as pdf files. Use Times New Roman size 12. For all special characters (e.g., Greek characters) use the font Symbol. Use carriage returns only to end headings and paragraphs, not to break lines of text. Automatic hyphenation should be turned off. Do not insert spaces before punctuation. Verify the correct spelling for the final version with the Spelling and Grammar function of Word.

3. Language
Manuscripts should be written in English.

4. Submission of the Manuscript
The manuscript and supplementary material must be send by the corresponding author to the Editor in an electronic form as pdf files.
You may be required to register as a new user within the Publication System Manager upon your first visit. Straightforward login and registration procedures can be found on the website. Editorial Manager allows authors to track the progress of manuscript review in real time. Detailed, step-by-step instructions for submitting manuscripts can be found on the website. All correspondence regarding your manuscript must go through Publication System Manager.
Authors are asked to prepare their papers, and PDF files, according to the templates and guidelines provided below.

The author(s) need to supply the following items :
- the PDF file of the paper
- a signed Assignment of Copyright Form (once the paper accepted)

We highly recommend that authors prepare their papers/PDFs using the following Microsoft Word or LaTeX templates, which can be downloaded from the following links :
- Microsoft guidelines and templates
- Latex guidelines and class file

A copyright license form will be provided to the corresponding author only when a paper is accepted for publication.

5. Peer Review
All submissions will be reviewed anonymously by at least two independent referees, and a referee should never communicate directly with an author. A referee must treat as confidential material the manuscript and any supplementary material. Authors may suggest names and email addresses of expert reviewers, but selection remains a prerogative of the Editors. Authors may include supplementary notes to facilitate the review process. If an accepted paper is cited

that has not yet appeared in print and is required for evaluation of the submitted manuscript, authors should provide an electronic version for use by the Reviewers. Authors are responsible for all statements in their work, including changes made by the copy editor after a manuscript is accepted.

6. Revision of the Manuscript
All comments made by referees must be addressed. A letter describing all changes that were made should be attached with the revised version of the manuscript. A copyright license form must accompany the final version of the manuscript.

7. Special Features, Appendices and Supplementary Material
Special features containing highly interactive features or large databases can be included. All authors are encouraged to take advantage WEB online publishing capabilities (i.e. 3-D, video, and interactive graphics). All special features must be created by the Author›s).
Authors who wish to publish electronic supplementary material to their article (Excel files, images, audio/video files) must submit the supplementary files/materials with their manuscript submission via our online peer review tracking Publication System Manager. Note that supplementary files are not automatically included in the reviewer PDF. Please therefore note in the cover letter if these materials should be evaluated by reviewers.

8. Preprint Option
Before making a PDF file of your article, please check the following tips in the next section.

Article checklist
There are a number of essential basic requirements which must be followed during preparation of your article. If article PDFs are prepared without following these essential requirements, publication may be delayed until a usable and compatible PDF is received.
Articles must not contain page numbers, headers or footers
This is extremely important. Page numbers, copyright details etc are added by the Publication System Manager Publishing during the production and publication process. If you put page numbers on your paper we will have to contact you for a replacement PDF, which could delay publication.

Article margins must be adequate
We recommend a minimum 15mm all round. The Microsoft Word template or Latex templates automatically provide the correct margins so their use is highly recommended.
All articles must have an abstract
When readers are searching for information online, an abstract of an article is the first thing they see. Your abstract needs to be concise but convey as much information as possible about the content of your article. In addition, our Publication System Manager Publishing will supply your abstract to many other database systems used by researchers to find papers.

Addresses should be complete and include the country and a contact name
The title of the article, author names with full first name (no degrees), each author's affiliation, and a suggested running head (of less than 50 characters, including spaces). The affiliation should comprise the department, institution (usually university or company), city, and country and should be typed as a footnote to the author's name. For the corresponding author designated to correspond with the Editorial Office and review proofs, indicate his/her complete mailing address, office/cellular telephone number, fax number, and e-mail address. During production of the electronic paper we may need to contact you if there is something to check or for a request a replacement PDF file.

References should be complete and carefully formatted
Online versions of all reference lists will, wherever possible, be electronically linked to the articles that you cite. Reference lists containing many links direct to the cited paper are a valuable research tool. The time and effort spent in preparing your references, so that they can be linked, will be very appreciated by readers of your paper. Please, notice also that all the citations should be justified according to the contents of the proposed article. Abusive citations of the same author may induce a certain delay in the overall review process

PrePrint
Any manuscript received for publication in FSE can be published on the Web as preprint. All authors submitting a manuscript must clearly indicate that they wish to publish it as a preprint. The referees appreciate if the manuscript meets the basic requirements for publication and recommend its publication as preprint. A preprint not accepted for publication by the referees will be immediately removed from the preprint collection. A published paper which was previously available as a preprint will have clearly indicated the date when it was first published on the Web. A work published as preprint can benefit from comments from the readers which can eventually improve the manuscript. Revised versions that incorporate corrections from reviewers and suggestions from readers can be also published as preprints.

9. Publication of the Manuscript
Accepted papers are published as PDF files available at the Web site of the Academy.

Transfer of Copyright Form
A signed copy of the Transfer of Copyright must be submitted online as part of the manuscript submission process (FSECopyright.pdf).

Astract
Reviews/State of the Art, Original Research Articles, require an abstract. The abstract is limited to 300 words or less. For Research Articles, the abstract should include a brief statement for each of the sections related to Introduction, Methods/Approaches/Materials and Discussion, and Conclusion written in paragraph form. All abstracts must be written in one paragraph, with no subheadings, equations, tables, reference citations or graphics.

Keywords
Provide a list of no more than 5 key words.
Introduction
Required for Reviews/State of the art and Original Research Articles.

Main Text Body
For Original Research Articles, organize the main text as follows: Introduction, Approach/Materials and Methods, Results, Discussion, and Conclusion. The use of subheadings to divide the text is encouraged. Primary, Secondary, and Third level headings should be clearly defined, but do not use numbers or letters.
Recommended word counts are as follows: Reviews/State of the art: 8000, Original Research Articles : 6000.
Use abbreviations sparingly, and define them at the first insertion in the text. Use the metric system for all measurements. Express metric abbreviations in lowercase letters without periods (cm, ml, sec). Define all symbols used in equations and formulas. When symbols are used extensively, the authors may include a list of all symbols in a table.

Conclusion
The conclusion should be a brief paragraph, containing 3 to 4 sentences, that summarizes the findings presented.

Acknowledgments
Include funding source(s) and other contributions. If the work has been funded by any organisation please provide name(s) of funding institute(s) and grant number(s).

References
References should conform to Vancouver style and be numbered consecutively in the order in which they are cited in the text. Cite in the text by the appropriate Arabic numeral enclosed in parentheses, e.g., (1), (2-5), etc.
It is advisable to limit the maximum number of references as really needed only.
References to unpublished peer-reviewed, personal communications, including conference abstracts, and papers in preparation or in review, cannot be listed, but can be notated parenthetically in the text.
Abbreviations for journal names should conform to those of Vancouver style (as depicted in http://www.library.uq.edu.au/training/citation/vancouv.pdf). The style and punctuation of the references should conform to conventional referencing.
Whenever, the paper is not yet published officially but accepted, please write down the corresponding DOI within the reference.
Authors may identify uniform resource locators (URLs) for websites that provide the reader with additional information on the topic addressed in the manuscript. Although URLs are an important feature of electronic publishing, authors are encouraged to be very selective in their choice of sites to include. Do not include links to sites that are not accessible without a password.
All on-line documents should contain author(s), title, On-line document/ Web /FTP /organisation /On-line database/ Supplementary material/ Private homepage , and Accessed Day Month Year, so that readers can refer to.

Tables
Tables must be created in Microsoft Word /Latex table format. Tables should be numbered (with Roman numerals) and referred to by number in the text. Center the title above the table, and type explanatory footnotes (indicated by superscript lowercase letters) below the table. Data must be placed in separate cells of the table to prevent text and numbers from shifting when the table is converted for publication on the Internet. Empty cells may be inserted to create spacing. Tables should not duplicate information provided in the text. Instead, tables should be used to provide additional information that illustrates or expands on a specific point the author wishes to make. Each table should be self-explanatory.

Figures
The FSE offers authors the use of color figures in online published manuscripts. Figures (as well as photographs, drawings, diagrams, and charts) are to be numbered in one consecutive series of Arabic numerals in the order in which they are cited in the text. All Electronic artwork must be submitted online via our online peer review tracking system, Publication System Manager.
The maximum combined count for tables and figures for papers should not exceed 15 to 20.

Footnotes
Footnotes should be avoided. When their use is absolutely necessary, footnotes should be numbered consecutively using Arabic numerals and should be typed at the bottom of the page to which they refer. Place a line above the footnote, so that it is set off from the text. Use the appropriate superscript numeral for citation in the text.

Guidelines

We highly recommend that authors prepare their papers/PDFs using the following Microsoft Word or LaTeX templates, which can be downloaded from the following links:

• Latex guidelines and class file http://www.academie.hassan2.sciences.ma/fse/FSE%20Sample%20cls.tex.txt

• PDF guidelines and templates http://www.academie.hassan2.sciences.ma/fse/FSE%20Sample%20cls.pdf

• Microsoft word guidelines and templates

Contractual issues

1. Full Disclosure
During the manuscript submission process, all authors will be required to confirm that the manuscript has not been previously published in any language anywhere and that it is not under simultaneous consideration by another journal.

2. Conflicts of Interest
Authors must declare all conflicts of interest (or their absence) in their cover letter upon submission of a manuscript. This conflict declaration includes conflicts or potential conflicts of all listed authors. If any conflicts are declared, FSE publish them with the paper. In cases of doubt, the circumstance should be disclosed so that the editors may assess its significance.
Conflicts may be financial, academic, commercial, political or personal. Financial interests may include employment, research funding (received or pending), stock or share ownership, patents, payment for lectures or travel, consultancies, nonfinancial support, or any fiduciary interest in a company.

3. Copyright Transfer
The Copyright Revision Act requires that Authors transfer their copyrights to the Publisher, HIIAST, in order to provide for the widest possible dissemination of professional and scientific literature. A signed Transfer of Copyright form must be submitted online with the manuscript. The Transfer of Copyright form for an accepted manuscript must be on file with the HIIAST Editorial Office prior to production for publication. Corresponding Authors may print and sign the form on behalf of all authors The Transfer of Copyright form can be found at fsecopyright.pdf.

4. Use of Copyrighted Tables and Figures
A copy of the granted permission to use copyrighted figures and tables must be included with the submitted manuscript.