

# MODELLING FOR ENGINEERING AND HUMAN BEHAVIOUR 2011

*Instituto de Matemática Multidisciplinar*



**BOOK OF EXTENDED ABSTRACTS**

**L. Jódar, L. Acedo and J. C. Cortés, Editors**  
**Instituto Universitario de Matemática Multidisciplinar**

*im<sup>2</sup>*

Instituto de Matemática Multidisciplinar



UNIVERSIDAD  
POLITECNICA  
DE VALENCIA

# MODELLING FOR ENGINEERING, & HUMAN BEHAVIOUR 2011

Instituto de Matemática Multidisciplinar  
Universidad Politécnica de Valencia  
Valencia 46022, SPAIN

Edited by  
Lucas Jódar, Luis Acedo and Juan Carlos Cortés  
Instituto Universitario de Matemática Multidisciplinar  
I.S.B.N.: 978-84-695-2143-4

## CONTENTS

1. **E. Defez, J. Sastre, J. Ibáñez, P. Ruiz, J.C. Cortés**, Solving engineering models using matrix functions ..... Pag: 1-17
2. **L. Acedo, J. A. Moraño, J. C. Cortés, D. Martín**, A random network model for EEG Pag: 18-22
3. **J. C. R. Alcantud, R. de Andrés Calle, J. M. Cascón**, A unifying model to measure consensus solutions in a society ..... Pag: 23-26
4. **L. Flores, V. Vidal, P. Mayo, F. Rodenas, G. Verdú**, Contrast of Two Methods of Reconstruction of CT Images Using High Performance Computing ..... Pag: 27-30
5. **A. G. Antebas, F. D. Denia, A. M. Pedrosa, F. J. Fuenmayor**, A finite element approach for the acoustic modelling of perforated dissipative mufflers with non-homogeneous properties ..... Pag: 31-35
6. **A. J. Arenas, G. González-Parra, B. Melendez**, Numerical solution of a nonlinear equation in mathematical finance ..... Pag: 36-41
7. **D. Ayala-Cabrera, J. Izquierdo, I. Montalvo, R. Pérez-García**, Water supply system component evaluation from GPR radargrams using a multi-agent approach ..... Pag: 42-45
8. **C. Bachiller, H. Estebany, J.V. Morroy, V. Boriaz**, Hybrid mode matching method for the efficient analysis of rods in waveguided structures ..... Pag: 46-52
9. **L. Bayón, J.M. Grau, M.M. Ruiz and P.M. Suárez**, Mathematical modelling of the combined optimization of a pumped-storage hydro-plant and a wind park ..... Pag: 53-56
10. **R. Berenguer-Vidal, R. Verdú-Monedero, J. Morales-Sánchez**, Design of B-spline multidimensional deformable models in the frequency domain ..... Pag: 57-60
11. **C. Guardiola, B. Pla, D. Blanco-Rodriguez, P. Cabrera**, Adaptive modelling for automotive applications ..... Pag: 61-66
12. **P. Bader, S. Blanes, J. C. Cortés, E. Ponsoda**, Geometric integrators for solving optimal control problems appearing for quadrotor UAV ..... Pag: 67-75
13. **M. Caballer-Tarazona, L. Buigues-Pastor, I. Saurí-Ferrer, R. Usó-Talamantes, J. L. Trillo-Mata**, A standardization model by equivalent patient to control outpatient pharmaceutical expenditure ..... Pag: 76-79
14. **J. David Cabedo, A. Esteve, J. M. Tirado**, Differences between adult health and early adult / late adolescent health: an empirical analysis ..... Pag: 80-85
15. **C. Bartual Sanfeliu, R. Cervelló Royo, I. Moya Clemente**, Measuring performance of social and non-profit Microfinance Institutions (MFIS) ..... Pag: 86-92

16. **B. Chen-Charpentier and D. Stanescu**, Virus Propagation with Randomness .... Pag: 93-96
17. **L. Acedo, J.C. Cortés, E. Defez, E. Ponsoda, M.J. Rodríguez-Álvarez, A. Sánchez-Sánchez, M. M. Tung, R.J. Villanueva**, A dynamical model to study the academic underachievement in Spain: A epidemiological approach ..... Pag: 97-104
18. **R. Duato, L. Jódar**, Mathematical modeling of the spread of divorces in Spain ..... Pag: 105-111
19. **J. Galindo, A. Tiseira, P. Fajardo, R. Navarro**, Analysis of the Influence of Different Real Flow Effects on a CFD Anechoic Boundary Condition Based on the MoC Pag: 112-116
20. **C. García-Meca, M. M. Tung, A. Martínez, J. C. Cortés**, The Variational Principle in Transformation Optics Engineering and some Application ..... Pag: 117-124
21. **P. J. García Nieto, F.J. de Cos Juez, F. Sánchez Lasheras, J. A. Vilán Vilán**, Evolutionary support vector regression algorithm applied to the prediction of the thickness of the chromium layer in a hard chromium plating process ..... Pag: 125-130
22. **K. Gibert**, The benefits of hybrid methodologies in Data Mining and Knowledge Discovery. Applications to Human Behaviour ..... Pag: 131-136
23. **S. González-Pintor, D. Ginestar, G. Verdú** , Using the proper generalized decomposition to compute the dominant mode of a nuclear reactor ..... Pag: 137-140
24. **G. González-Parra, J. L. Herrera, A. J. Arenas**, Modeling adaptative social behaviour on epidemics with dynamical networks .....Pag: 141-146
25. **V. Giménez, F. García, F. Guijarro**, Credit risk management: A multicriteria approach to assess creditworthiness .....Pag: 147-153
26. **J. A. GutiérrezPérez, M. Herrera, R. PérezGarcía, E. RamosMartínez**, Application of graph-spectral methods in the vulnerability assessment of water supply networks ... Pag: 154-157
27. **D. Hinestroza, J. Peralta, L. E. Olivar**, Regularization Algorithm within Two-Parameters for one and two dimensions for the Identification of the Heat Conduction Coefficient in the Parabolic Equation ..... Pag: 158-165
28. **S.Hoyas, X. Margot, A. Gil, D. Khuong-Anh, F. Ravet**, Evaluation of the Eulerian-Lagrangian Spray Atomization (ELSA) model in spray simulations: 2D cases ..... Pag: 166-169
29. **A. Cordero, J.L. Hueso, E. Martínez, J.R. Torregrosa**, Optimal methods with approximated derivatives for nonlinear equations ..... Pag: 170-174

30. **J.M. Bru-Juanes, J.M. Bru-Lázaro, M. Herrera, J. Izquierdo**, Normal goniometric values to guide decision-making in lower-extremity rotational problems using support vector machine techniques .....Pag: 175-178
31. **J. J. López, D. Carnicero, N. Ferrando, J. Escolano**, Parallelization of the Finite-Difference Time-Domain Method for Room Acoustics Modelling based on CUDA ..... Pag: 179-182
32. **I. A. Mantilla, M. Leonardi, J. V. Balbastre, E. de los Reyes**, Singular value decomposition and Tikhonov regularization in hyperbolic passive location .....Pag: 183-188
33. **J. Cabrera, M.A. Castro, F. Rodríguez, J. A. Martín**, Difference schemes for some dual-phase-lag models of heat conduction ..... Pag: 189-193
34. **J. Martínez, M. López, J. M. Matías, J. Taboada**, Classifying Slate Tile Quality Using Automated Learning Techniques ..... Pag: 194-197
35. **S.Carlos, I. Marton, A. Sanchez, S. Martorell**, Onshore Wind Farms Maintenance Optimization Using a Stochastic Model ..... Pag: 198-202
36. **L. Bergamaschi, R. Bru, A. Martínez, J. Mas, M. Putti**, Low-rank Update of Pre-conditioners for the nonlinear Richards Equation ..... Pag: 203-206
37. **E. de la Poza, N. Guadalajara, L. Jódar, P. Merello**, Predicting the development of anxiolytic addiction: a discrete mathematical model approach ..... Pag: 207-210
38. **J. Cerdá, C. Montoliu, R. J. Colom**, LGEM: A Lattice Boltzmann Economic Model for income distribution and tax regulation .....Pag: 211-215
39. **A. Montorfano, F. Piscaglia, G. Ferrari**, Development of a Non-Reflective NSCBC for the Multi-Dimensional non-linear acoustic simulation of silencers for internal combustion engines .....Pag: 216-220
40. **J. Mora-Pascual, H. Mora-Mora, M. T. Signes-Pont, P. Soler-Javaloy**, Artificial Neural Model for Classification of Volunteers ..... Pag: 221-228
41. **F. Moreno, E. Ospina**, On Estimating the Position and Direction of a Leader of a Group of Entities ..... Pag: 229-232
42. **M. V. Sebastián, M. A. Navascués**, Electroencephalographic indices of interpolatory type .....Pag: 233-236
43. **J.J. López, R. Novella, A. García, J.F. Winklinger**, A computational investigation of the ignition and combustion processes of diesel-like dual-fuel sprays ..... Pag: 237-242
44. **P. Olmeda, V. Dolz, F.J. Arnau, M.A. Reyes-Belmonte**, A one dimensional lumped model for calculation of heat fluxes inside turbochargers ..... Pag: 243-247

45. **Š. Papáček, R. Kaňa, C. Matonoha**, Estimation of Diffusivity of Phycobilisomes on Thylakoid Membrane based on spatio-temporal FRAP images ..... Pag: 248-251
46. **R. Payri, J. Gimeno, P. Marti-Aldaravi, G. Bracho**, Study of the influence of the inlet boundary conditions in a LES simulation of internal flow in a diesel injector ..... Pag: 252-259
47. **F. Pedroche, F. Moreno, A. González, A. Valencia**, Leadership groups on Social Network Sites based on personalized PageRank ..... Pag: 260-263
48. **D. de Pereda, S. Romero, J. Bondia**, On the computation of output bounds on parallel inputs pharmacokinetic models with parametric uncertainty ..... Pag: 264-267
49. **J. R. Serrano, H. Climent, P. Piqueras, O. García-Afonso**, A chemical species transport model with variable gas properties for gas dynamic calculation of internal combustion engines ..... Pag: 268-273
50. **J. M. Rey**, Optimal control of long-term sentimental dynamics ..... Pag: 274-278
51. **F. Reyes-Santías, C. Cadarso-Suárez, A. Martínez-Calvo**, Applying a simulation model to manage waiting lists for hospital inpatient activity in a EU region .. Pag: 279-283
52. **R. Cantó, B. Ricarte, A. M. Urbano**, Modelling big game populations when hunting is age and sex selective ..... Pag: 284-287
53. **M. J. Rodríguez-Álvarez, F. Sánchez, A. Soriano, A. Iborra, J. C. Cortés**, Stability and convergence analysis of several approaches for a mathematical description of a small animal computed tomographic system ..... Pag: 288-292
54. **F. J. Salvador, J. Martínez-López, J. V. Romero, M. D. Roselló**, Study of the turbulence induced by cavitation phenomenon in diesel injector nozzles by Large Eddy Simulation (LES) ..... Pag: 293-299
55. **J. L. Sánchez-Romero, A. Jimeno-Morenilla, R. Molina-Carmona, J. Pérez-Martínez**, An Approach to the Application of Shift-and-add Algorithms on Engineering and Industrial Processes ..... Pag: 300-305
56. **V. Štumbauer, K. Petera, D. Štys**, Lattice Boltzmann method in bioreactor design and simulation ..... Pag: 306-309
57. **F. Chicharro, A. Cordero, J. R. Torregrosa**, Improvements of true anomaly iteration method for preliminary orbit determination ..... Pag: 310-313
58. **M. Peco, F. J. Santonja, A. C. Tarazona, R. J. Villanueva, J. Villanueva-Oller**, Modelling the attitude dynamics of the Basque population towards ETA before the passing of the Law of Political Parties (LPP) ..... Pag: 314-319
59. **R. Usó, M. Caballer, I. Barrachina, J. L. Trillo, D. Vivas**, Analysis and assessment of models of pharmaceutical expenditure prediction ..... Pag: 320-324

# Solving engineering models using matrix functions \*

E. Defez<sup>\*</sup> † J. Sastre<sup>†</sup>, J. Ibáñez<sup>‡</sup>, P. Ruiz<sup>‡</sup>, J.C. Cortés<sup>\*</sup>

(<sup>\*</sup>) Instituto de Matemática Multidisciplinar,

(<sup>†</sup>) Instituto de Telecomunicaciones y Aplicaciones Multimedia,

(<sup>‡</sup>) Instituto de Instrumentación para Imagen Molecular,

Universitat Politècnica de València, Camino de Vera s/n, 46022, Valencia, Spain.

November 30, 2011

## Abstract

Trigonometric matrix functions play a fundamental role in second order differential systems. In this article an efficient and highly-accurate Hermite algorithm is presented for the computation of matrix cosine. For an efficient evaluation of the polynomial approximation, the matrix is scaled and the Paterson-Stockmeyer's method is applied. The double angle formula  $\cos(2A) = 2\cos^2(A) - I$  is used to recover the cosine of the original matrix. A careful error analysis of the Hermite approximation is given, and an improved forward absolute error bound is obtained with respect to previous existing Hermite algorithms. The optimal values of its parameters are obtained, which allows to reduce the number of double angle steps and to obtain a more accurate computed cosine in floating point arithmetic. Also, a roundoff error study in the double angle phase and in the evaluation of polynomial approximation is performed. A MATLAB implementation of the developed algorithm is compared to other efficient state-of-the-art implementations on a large class of matrices for different

---

\* **Acknowledgements.** This work has been partially supported by the Spanish *Ministerio de Educación* grant MTM2009-08587 and the Universidad Politècnica de Valencia grant PAID-06-11-2020.

† e-mail: edefez@imm.upv.es

dimensions, obtaining higher accuracy and lower computational costs in the majority of cases.

## 1 Introduction

Matrix functions play a relevant role in different areas of science and technology. They arise most frequently in connection with the solution of differential systems and control theory. For example, it is well known that the wave equation

$$v^2 \frac{\partial^2 \psi}{\partial x^2} = \frac{\partial^2 \psi}{\partial t^2}, \quad (1)$$

plays an important role in many areas of engineering and applied sciences. When we use the spatially semi-discretization method of the wave equation (1), we obtain the matrix differential problem

$$Y''(t) + AY(t) = 0, \quad Y(0) = Y_0, \quad Y'(0) = Y_1, \quad (2)$$

where  $A$  is a square matrix and  $Y_0$  and  $Y_1$  are vectors, see [1] for details. Matrix problem (2) has the solution

$$Y(t) = \cos(\sqrt{A}t)Y_0 + (\sqrt{A})^{-1} \sin(\sqrt{A}t)Y_1, \quad (3)$$

where  $\sqrt{A}$  denotes any square root of a non-singular matrix  $A$  (see *e.g.* expression (1.2) of [2]). More general problems of type (2), with a forcing term  $F(t)$  on the right-hand side arise from mechanical systems without damping, and their solutions can be expressed in terms of integrals involving the matrix sine and cosine [3]. Thus, trigonometric matrix functions play an important role in second order differential systems, similar to matrix exponential  $e^{At}$  in first order differential systems [4].

Moreover, the matrix cosine is used in the method of Yau-Lu for reducing the symmetric eigenvalue problem (finding all the eigenvalues and eigenvectors of a dense symmetric matrix) to a number of matrix multiplications [5].

Computing the matrix sine reduces to computing the matrix cosine through  $\sin(A) = \cos(A - \frac{\pi}{2}I)$ . Thus we concentrate on the matrix cosine. Serbin and Blalock proposed a general algorithm for computing the matrix cosine in [1], which uses rational approximations and the double angle formula

$$\cos(2A) = 2 \cos^2(A) - I. \quad (4)$$

In [6] new methods for computing matrix exponential, sine and cosine based on Hermite matrix polynomial series were presented. A new bound for Hermite matrix polynomials was provided and it was used to give expressions for obtaining the number of Hermite series terms depending on the desired approximation error in exact arithmetic. Later, Higham, Smith and Hargreaves developed two algorithms based on (4) and Padé approximation in [7, 2], including truncation and rounding error analysis. Recently, in [8] the authors introduced a new parameter in the matrix cosine and sine Hermite series from [6] and new error bounds, improving both accuracy and efficiency.

In this paper we use the new matrix cosine Hermite series from [8], providing sharper bounds for Hermite matrix polynomials and the approximation error, and computing the optimal values of the series parameter to develop a competitive Hermite algorithm for computing the matrix cosine in IEEE double precision arithmetic that uses: matrix scaling based on (4), Paterson-Stockmeyer’s method for the evaluation of Hermite series [9, 10], and accuracy bound tests similar to those proposed in [10, p. 6456-6457]. A MATLAB implementation of this algorithm is made available online and it is compared with the MATLAB function `funm` [11] and a MATLAB implementation based on the Padé algorithm given in [2], i.e. function `cosm`, providing higher accuracy and efficiency than both methods in the majority of test matrices.

This paper is organized as follows. Section 2 summarizes previous results of Hermite matrix polynomial series expansion of  $\cos(A)$  and the development of a new error bound. In Section 3, an algorithm based on that error bound is described and a bound for the corresponding roundoff error is provided. Numerical experiments are presented in Section 4. Finally, conclusions are given in Section 5.

Throughout this paper,  $[x]$  denotes the integer part of  $x$ . To obtain the above mentioned error bound, we will use any subordinate matrix norm  $\|A\|$ ,  $A \in \mathbb{C}^{r \times r}$ , and in the subsequent error analysis, we will use the 1-norm  $\|A\|_1$ . If  $\mathcal{A}(k, n)$  is a matrix in  $\mathbb{C}^{r \times r}$  for  $n \geq 0, k \geq 0$ , the following identity holds [6]:

$$\sum_{n \geq 0} \sum_{k \geq 0} \mathcal{A}(k, n) = \sum_{n \geq 0} \sum_{k=0}^n \mathcal{A}(k, n - k). \tag{5}$$

## 2 Hermite matrix polynomial series expansions of matrix cosine. Error bound

For the sake of clarity in the presentation of the following results we recall some properties of Hermite matrix polynomials which have been established in [12, 6, 8]. From (3.4) of [12, p. 25] the  $n$ th Hermite matrix polynomial satisfies

$$H_n \left( x, \frac{1}{2}A^2 \right) = n! \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-1)^k (xA)^{n-2k}}{k!(n-2k)!}, \tag{6}$$

for an arbitrary matrix  $A$  in  $\mathbb{C}^{r \times r}$ . From [8], we have the following Hermite matrix polynomial series expansion of the matrix cosine  $\cos(Ay)$ :

$$\cos(Ay) = e^{-\frac{1}{\lambda^2}} \sum_{n \geq 0} \frac{(-1)^n}{\lambda^{2n}(2n)!} H_{2n} \left( y\lambda, \frac{1}{2}A^2 \right). \tag{7}$$

Denoting by  $C_N(\lambda, A^2)$  the  $N$ th partial sum of series (7) for  $y = 1$ , one gets the approximation

$$C_N(\lambda, A^2) = e^{-\frac{1}{\lambda^2}} \sum_{n=0}^N \frac{(-1)^n}{\lambda^{2n}(2n)!} H_{2n} \left( \lambda, \frac{1}{2}A^2 \right) \approx \cos(A), \quad \lambda \in \mathbb{C}. \tag{8}$$

Working similarly as in [13, pp. 1913], we can obtain a bound for Hermite matrix polynomials  $\|H_{2n}(x, \frac{1}{2}A^2)\|$  based on  $\|A^2\|$ , see [2], using the Taylor series for the hyperbolic cosine  $\cosh(y) = \sum_{n \geq 0} y^{2n}/(2n)!$ . Taking norms in (6), one gets

$$\left\| H_{2n} \left( x, \frac{1}{2}A^2 \right) \right\| \leq (2n)! \sum_{k=0}^n \frac{\left( \|A^2\|^{\frac{1}{2}} \right)^{2(n-k)}}{k!(2(n-k))!} |x|^{2(n-k)}. \tag{9}$$

On the other hand, using (5), it follows that

$$\begin{aligned} e \cosh \left( |x| \|A^2\|^{\frac{1}{2}} \right) &= \sum_{n \geq 0} \frac{\left( |x| \|A^2\|^{\frac{1}{2}} \right)^{2n}}{(2n)!} \sum_{k \geq 0} \frac{1}{k!} = \sum_{n \geq 0} \sum_{k \geq 0} \frac{\left( |x| \|A^2\|^{\frac{1}{2}} \right)^{2n}}{(2n)!k!} \\ &= \sum_{n \geq 0} \sum_{k=0}^n \frac{\left( \|A^2\|^{\frac{1}{2}} \right)^{2(n-k)}}{k!(2(n-k))!} |x|^{2(n-k)}, \end{aligned} \tag{10}$$

and as a consequence

$$\sum_{k=0}^n \frac{\left(\|A^2\|^{\frac{1}{2}}\right)^{2(n-k)}}{k!(2(n-k))!} |x|^{2(n-k)} \leq e \cosh\left(|x| \|A^2\|^{\frac{1}{2}}\right). \quad (11)$$

Multiplying by  $(2n)!$  in (11) and using (9), we have the result:

$$\left\|H_{2n}\left(x, \frac{1}{2}A^2\right)\right\| \leq (2n)! e \cosh\left(x \|A^2\|^{\frac{1}{2}}\right), \quad \forall x \in \mathbb{R}, \quad n \geq 0, \quad \forall A \in \mathbb{C}^{r \times r}. \quad (12)$$

Taking into account approximation (8) and bound (12), it follows that

$$\begin{aligned} \|\cos(A) - C_N(\lambda, A^2)\| &\leq e^{-\frac{1}{\lambda^2}} \sum_{k \geq N+1} \frac{1}{\lambda^{2k} (2k)!} \left\|H_{2k}\left(\lambda, \frac{1}{2}A^2\right)\right\| \\ &\leq e^{1-\frac{1}{\lambda^2}} \cosh\left(\lambda \|A^2\|^{\frac{1}{2}}\right) \sum_{k \geq N+1} \frac{1}{\lambda^{2k}} \\ &= e^{1-\frac{1}{\lambda^2}} \cosh\left(\lambda \|A^2\|^{\frac{1}{2}}\right) \left[ \sum_{k \geq 0} \frac{1}{\lambda^{2k}} - \sum_{k=0}^N \frac{1}{\lambda^{2k}} \right]. \quad (13) \end{aligned}$$

Simplifying the geometric series in (13), we have finally the bound:

$$\|\cos(A) - C_N(\lambda, A^2)\| \leq \frac{e^{1-\frac{1}{\lambda^2}} \cosh\left(\lambda \|A^2\|^{\frac{1}{2}}\right)}{(\lambda^2 - 1)\lambda^{2N}}. \quad (14)$$

### 3 Algorithm

In this section we describe Algorithm 1 (`cosh`), based on Hermite series, which uses bound (14) for choosing the scaling factor for computing the matrix cosine.

---

**Algorithm 1** (*cosher*) Given a matrix  $A \in \mathbb{C}^{r \times r}$  and the order  $N$  of Hermite matrix approximation of the cosine function, this algorithm computes  $B \cong \cos(A)$ .

---

- 1: Preprocessing of matrix  $A$ :  $\check{A}$ .
  - 2: Compute optimal values of  $\lambda$  and  $s$ .
  - 3: SCALING PHASE:  $\tilde{A} = \check{A}/2^s$
  - 4: Compute  $\tilde{B} = C_N(\lambda, \tilde{A}^2)$ , where  $C_N$  is the Hermite approximation of the cosine function.
  - 5: **for**  $i = 1 : s$  **do**
  - 6:      $\tilde{B} = 2\tilde{B}^2 - I$
  - 7: **end for**
  - 8: Postprocessing of matrix  $\tilde{B}$ :  $B$ .
- 

The preprocessing and postprocessing are based on applying transformations to reduce the norm of matrix  $A$  and recover the matrix  $B \cong \cos(A)$  from the matrix  $\tilde{B}$  obtained in the Loop 5-7. The available techniques to reduce the norm of a matrix are argument translation and balancing [14, p. 299]. The argument translation is based on the formula

$$\cos(A - \pi j I) = (-1)^j \cos(A), \quad k \in \mathbb{Z},$$

and on finding the integer  $q$  such that the norm of matrix  $A - \pi q I$  is minimum. This value can be calculated by using Theorem 4.18 from [14]. Balancing is a heuristic that attempts to equalize the norms of the  $k$ th row and  $k$ th column, for each  $k$ , by a diagonal similarity transformation defined by a non singular matrix  $D$ . Balancing tends to reduce the norm, though this is not guaranteed, so we will use it only for matrices where the norm is really reduced. For those matrices, if  $\check{A} = D^{-1}(A - \pi q I)D$  is the obtained matrix in the preprocessing, the postprocessing consists of computing  $B = (-1)^q D \cos(\tilde{B}) D^{-1}$ .

For the evaluation of  $C_N(\lambda, \tilde{A}^2)$  the Horner and Paterson-Stockmeyer's method can be applied [9, 10], obtaining previously with MATLAB the symbolic expression of  $C_N(\lambda, \tilde{A}^2)$  as a polynomial of matrix  $\tilde{A}^2$  with degree  $N$  and coefficients depending on  $\lambda$ , for each considered value of  $N$ . The double angle formula (4) is used to recover  $\cos(\check{A})$  from the matrix  $\tilde{B}$  obtained in step 4, see [1] for details.

The optimal values of  $N$  with respect to cost of the evaluation of matrix polynomial  $C_N(\lambda, \tilde{A}^2)$  with Paterson-Stockmeyer method are included in the set  $S_N = \{1, 2, 4, 6, 9, 12, 16, 20, \dots\}$ , see [10, p. 6454]. The scaling factor  $s$

and the parameter  $\lambda$  is chosen as follows. If  $N_k$ , where  $k$  is the position of  $N$  in  $S_N$ , is the chosen order of Hermite approximation, the number of matrix product evaluations in Algorithm 1 is equal to  $k + s$ . Hence, the number of matrix products depends on the scaling factor  $s$ . For the evaluation of  $\cos(A)$  with IEEE double precision arithmetic, analogously to [2], we consider an absolute error-based algorithm. If we take the error in (14) to be lower than or equal to the unit roundoff in double precision floating-point  $u = 2^{-53}$ , i.e.,

$$\frac{e^{1-\frac{1}{\lambda^2}} \cosh\left(\lambda \sqrt{\|(\check{A}/2^s)^2\|_1}\right)}{(\lambda^2 - 1)\lambda^{2N}} \leq u, \quad (15)$$

then

$$s \geq \frac{1}{2} \log_2(\|\check{A}^2\|_1) + \log_2\left(\frac{\lambda}{\operatorname{arc\,cosh}\left(\frac{u(\lambda^2-1)\lambda^{2N}}{e^{1-1/\lambda^2}}\right)}\right). \quad (16)$$

We choose the parameter  $\lambda = \lambda_{\min}$ , where  $\lambda_{\min}$  is the value of  $\lambda$  such that the right-hand side of (16) reaches the minimum value. Hence, if we define

$$\Theta_N = \frac{1}{\lambda_{\min}} \operatorname{arc\,cosh}\left(\frac{u(\lambda_{\min}^2 - 1)\lambda_{\min}^{2N}}{e^{1-1/\lambda_{\min}^2}}\right) \quad (17)$$

then, using (16), it follows that

$$s = \left\lceil \log_2\left(\frac{\sqrt{\|\check{A}^2\|_1}}{\Theta_N}\right) \right\rceil. \quad (18)$$

We discard the values of  $N$  except for  $\{1, 2, 4, 6, 9, 12, 16, 20\}$ , because for  $N = 25, 30$  the corresponding values of  $\lambda_{\min}$  are negative. These values are independent of the norm of matrix  $\check{A}^2$ , see (16), and they have been computed by using the symbolic functions `diff` and `solve` from the Symbolic Math Toolbox 5 of MATLAB. The values of  $\lambda_{\min}$  and  $\Theta_N$  are listed in the Figure 1.

$\lambda_{\min}$		$\Theta_N$	
$N=1$	28614.3702451495925	$N=1$	$1.3988322173046763 \cdot 10^{-4}$
$N=2$	1304.99637514915918	$N=2$	$4.5977704110066707 \cdot 10^{-3}$
$N=4$	110.428178898694292	$N=4$	$9.0556596644120163 \cdot 10^{-2}$
$N=6$	38.3201292093300207	$N=6$	$3.6534325997941364 \cdot 10^{-1}$
$N=9$	17.3255806739152432	$N=9$	1.1543637495804793
$N=12$	11.2995380153548675	$N=12$	2.3009899711770276
$N=16$	8.08117035928883672	$N=16$	4.2073703112196084
$N=20$	6.56678564572528643	$N=20$	6.3959908727565082

(a) Optimal values of  $\lambda_{\min}$ . (b) Values of  $\Theta_N$  from (17).

Figure 1: Tables of  $\lambda_{\min}$  and  $\Theta_N$ .

In the final MATLAB implementation of `cosher`, available online in <http://personales.upv.es/~jorsasma/cosher.m>, we have applied similar accuracy bound tests to those in [10, p. 6456-6457] to the sum of the highest degree terms of the Hermite series  $C_N(\lambda, \tilde{A}^2)$ . The tests consist of verifying if the norm of the sum of several highest degree terms is lower than the unit roundoff. If that is the case, then the corresponding series terms can be neglected, permitting to save matrix products. This kind of tests were already performed with the matrix exponential Hermite series in [10], saving matrix products and therefore reducing the cost for some matrices. For a complete description of algorithm `cosher` see MATLAB implementation `cosher.m` mentioned above.

The analysis of rounding errors of Algorithm 1 are based on the results given in [14, p. 290-293]. Let us define  $\tilde{B}_i = \cos(2^{i-s}\tilde{A})$ . The positive integer  $s$  has been chosen such that we have obtained a good approximation of  $\tilde{B}_0 = \cos(2^{-s}\tilde{A})$ . The propagation error in the recurrence

$$\tilde{B}_{i+1} = 2\tilde{B}_i^2 - I, \quad i = 0 : s - 1,$$

can be obtained by applying [14, Th. 12.5]:

If  $\tilde{B}_i = \tilde{B}_i + E_i$ ,  $\tilde{B}_{i+1} = fl(2\tilde{B}_i^2 - I)$  and we assume that  $\|E_i\|_1 \leq 0.05\|\tilde{B}_i\|_1$ , then

$$\begin{aligned} \|E_i\|_1 &\leq (4.1)^i \|E_0\|_1 \|\tilde{B}_0\|_1 \|\tilde{B}_1\|_1 \cdots \|\tilde{B}_{i-1}\|_1 + \\ \tilde{\gamma}_{r+1} &\sum_{j=0}^{i-1} (4.1)^{i-j-1} \left( 2.21 \|\tilde{B}_j\|_1^2 + 1 \right) \|\tilde{B}_{j+1}\|_1 \cdots \|\tilde{B}_{i-1}\|_1, \end{aligned}$$

where the error constant  $\tilde{\gamma}_{r+1}$  is defined by

$$\tilde{\gamma}_{r+1} = \frac{c(r+1)u}{1 - c(r+1)u},$$

where  $c$  is a small integer constant whose precise value is unimportant [14, p. 332]. Since  $\|\tilde{B}_{i+1}\|_1 \leq 2\|\tilde{B}_i\|_1^2 + 1$ , the error  $\|E_i\|_1$  fundamentally depends on the norms of the matrices  $\|\tilde{B}_0\|_1$  and  $\|E_0\|_1$ .

From [14, Th. 12.3] and taking into account (14), (15) and (18), we obtain the following bounds

$$\|E_0\|_1 \leq \frac{e^{1 - \frac{1}{\lambda_{m_N}^2}} \cosh(\lambda_{m_N} \Theta_N)}{(\lambda_{m_N}^2 - 1)\lambda_{m_N}^{2N}} \leq u, \quad (19)$$

$$\|\tilde{B}_0\|_1 = \|\cos(2^{-s}A)\|_1 \leq \cosh\left(\sqrt{\|4^{-s}A^2\|_1}\right) \leq \cosh(\Theta_N), \quad (20)$$

where  $\lambda_{m_N}$  is the minimum  $\lambda$  corresponding to  $N$ .

Table 1 shows the values of bound (20) for different orders. For all considered  $N$ , the bounds of  $\|\tilde{B}_0\|_1$  are not much larger than 1, and given that  $\|E_0\|_1$  is bounded by the unit roundoff in double precision floating-point  $u$ , the absolute error criterion allows significant accuracy in the double angle phase for these values.

$N=1$	1.0000000097836579
$N=2$	1.0000105697649961
$N=4$	1.0041030513703506
$N=6$	1.0674834828383193
$N=9$	1.7436313211414083
$N=12$	5.0421105683031042
$N=16$	33.597265612110441
$N=20$	299.71933153364563

Table 1: Bounds of  $\|\tilde{B}_0\|_1$ .

The effects of rounding errors on the evaluation of  $\tilde{B} = C_N(\lambda, \tilde{A}^2)$  can be deduced from [14, Th. 4.5]:

Let  $p_N(\tilde{A}^2) = \sum_{k=0}^N b_k \tilde{A}^{2k}$  be the polynomial expansion of  $C_N(\lambda, \tilde{A}^2)$  (see (8)). If the Paterson-Stockmeyer method is used for computing  $p_N(\tilde{A}^2)$ , then

$$\left\| p_N(\tilde{A}^2) - fl\left(p_N(\tilde{A}^2)\right) \right\|_1 \leq \tilde{\gamma}_{Nr} \hat{p}_N\left(\|\tilde{A}^2\|_1\right), \quad (21)$$

where  $\tilde{\gamma}_{Nr} = \frac{cNru}{1-cNru}$ ,  $\hat{p}_N\left(\|\tilde{A}^2\|_1\right) = \sum_{k=0}^N |b_k| \|\tilde{A}^2\|_1^k$ .

Taking account the rounding error in forming  $\tilde{A}^2$ , we found that the bound (21) is multiplied by a term that is approximately equal to

$$\mu(\tilde{A}) = \frac{\|\tilde{A}^2\|_1}{\|\tilde{A}\|_1^2} \geq 1.$$

As

$$\mu(\tilde{A}) \leq \left( \frac{\|\tilde{A}\|_1}{\sqrt{\|\tilde{A}^2\|_1}} \right)^2,$$

if  $\mu(\tilde{A})$  increases, then the rounding errors in the evaluation of  $p_N(\tilde{A}^2)$  increase. However, a smaller  $s$  is also obtained, and therefore the rounding errors in the double angle phase decrease. Hence, the rounding errors in the two phases are balanced.

## 4 Numerical examples.

In this section we compare MATLAB implementation `cosh` with functions `funm` and `cosm`. `funm` is a MATLAB function that computes matrix cosine and other matrix functions at square matrices using the Schur-Parlett algorithm from [11]. Function `cosm` is a MATLAB implementation of Algorithm 5.1 proposed in [2] which uses Padé approximants of cosine function (<http://www.maths.manchester.ac.uk/higham/mftoolbox>). MATLAB 7.9 (R2009b) implementations were tested on an Intel Core 2 Duo processor at 2.52 GHz with 4 GB main memory. Algorithm accuracy was tested by computing the relative error

$$E = \frac{\|\cos(A) - \tilde{Y}\|_1}{\|\cos(A)\|_1},$$

where  $\tilde{Y}$  is the computed solution and  $\cos(A)$  the exact solution. In the tests we did not use any preprocessing/postprocessing in the implemented algorithms. Analogously to the experiments in [15], we found that turning on preprocessing in this algorithm provided similar results to those presented in this section without preprocessing.

We used a set of 102 test matrices:

- Matrices 1-47: Forty-seven  $10 \times 10$  matrices obtained from the function `matrix` of the Matrix Computation Toolbox [16]. Four of 52 matrices from above Toolbox were excluded (matrices 17, 20, 42 and 44) because their matrix cosine can not be represented in double precision due to overflow errors. Also, the matrix 49 has been excluded because this matrix is equal to matrix 43.
- Matrices 48-71: Twenty four  $9 \times 9$  or  $10 \times 10$  matrices from the Eigtool MATLAB package [17]. This package allows large-scale eigenvalue computations for the pseudospectra of dense and sparse matrices.
- Matrices 72-75: Four  $3 \times 3$  and one  $10 \times 10$  real matrices from [18].
- Matrices 76-79: They correspond to the above excluded matrices of Matrix Computation Toolbox, but considering other dimensions. The dimensions considered were  $2 \times 2$  for matrices 17 and 20,  $3 \times 3$  for matrix 42 and  $6 \times 6$  for matrix 44.
- Matrix 80:  $2 \times 2$  matrix from [4, p. 10].
- Matrix 81:  $3 \times 3$  real diagonal matrix from [19].
- Matrix 82:  $2 \times 2$  real matrix from [11, example 2].
- Matrix 83:  $5 \times 5$  real matrix from [20, p. 655].
- Matrix 84:  $10 \times 10$  real matrix from [21, example 11].
- Matrix 85-86: Two symmetric  $10 \times 10$  real matrices from [22, examples 1 and 2].
- Matrix 87:  $10 \times 10$  Vandermonde real matrix from MATLAB 7.6 function `vander` whose columns are powers of the vector  $v = 1 : 0.2 : 3$  (see notation of MATLAB vectors).

- Matrix 88:  $10 \times 10$  Hankel real matrix from MATLAB 7.6 function `hankel`. This matrix was generated as `A=hankel(1:10,1:10)`.
- Matrix 89:  $10 \times 10$  Toeplitz matrix from MATLAB 7.6 function `toeplitz`. This matrix was generated as `A=toeplitz([1 2 3 4 5], [1.5 2.5 3.5 4.5 5.5])`.
- Matrix 90:  $10 \times 10$  nonsymmetric Wilkinson matrix from MATLAB 7.6 function `wilkinson`. This matrix was generated as `A=wilkinson(10)`.
- Matrix 91: Symmetric  $10 \times 10$  real matrix obtained from  $A = (A_0 + A_0^T)/2$ , where  $A_0$  is a matrix generated with uniform random values in  $[-50, 50]$ .
- Matrix 92: Nondefective  $10 \times 10$  real matrix with separate eigenvalues  $\sigma_i = 10^{i-8}$ ,  $i = 1, 2, \dots, 10$ .
- Matrix 93: Nondefective  $10 \times 10$  real matrix with a multiple eigenvalue  $\sigma = 10$ .
- Matrix 94: Nondefective  $10 \times 10$  real matrix with two separate eigenvalues  $\sigma = 10^{-4}$  and  $\sigma = 10^2$  of multiplicities equal to 5.
- Matrix 95: Defective  $10 \times 10$  real matrix with eigenvalues 2,  $-3$  and 4 of multiplicities equal to 3, 2 and 5 respectively.
- Matrices 96-102: Seven  $10 \times 10$  real matrices from [23, example 3.10].

The “exact” matrix cosine was calculated analytically when possible, and otherwise using MATLAB’s Symbolic Math Toolbox with high precision arithmetic.

Tables 2a and 2b show the comparatives `cosher-cosm` and `cosher-funm` for  $N \in \{9, 12, 16, 20\}$ . The first three rows show the percentages of times that the relative error of the first function is lower, equal or greater than the relative error of the second function. The fourth row shows the ratio of matrix products needed for computing the matrix cosine for over all the test matrices. In the same way as in [10, p. 6459] we have considered that the asymptotic cost in terms of matrix products for solving the multiple right-hand side linear system that appears in Padé algorithm is  $4/3$ . The computational cost of `funm` depends greatly on the eigenvalue distribution of



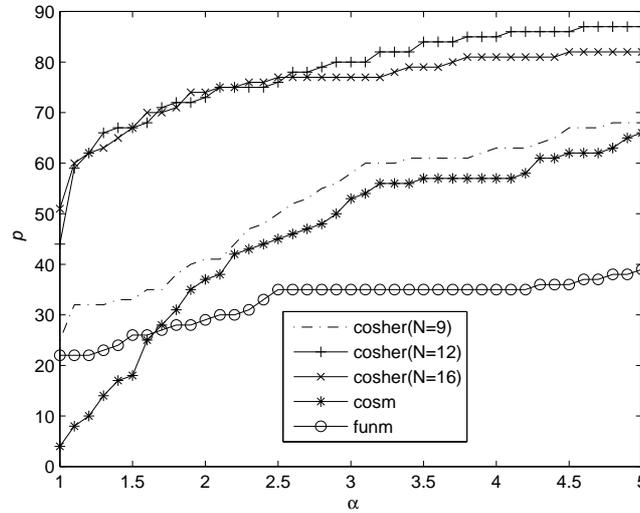


Figure 3: Performance profile of *cosher*, *cosm* and *funm* for the set of test matrices.

The following conclusions can be emphasized:

- Computational cost of *cosher* is lower than computational costs of *cosm* and *funm* for all considered orders.
- Function *cosher* is more accurate than *cosm* for  $N = 12, 16, 20$ , and it is more accurate than *funm* for all considered orders.
- In the performed numerical tests, the optimal order of Hermite algorithm is  $N = 16$ .

## 5 Conclusions.

In this work an efficient algorithm to compute the matrix cosine based on Hermite matrix polynomial expansions has been presented. This algorithm improves the algorithms proposed by the authors in [6, 8]. The new algorithm uses the scaling technique based on the double angle formula, the Horner and Paterson-Stockmeyer’s method for computing the Hermite matrix polynomial approximation, a new bound of the absolute error, and accuracy bound

tests for neglecting higher order series terms. A careful analysis of rounding errors of the developed algorithm has also been given. The MATLAB function `cosh` is available online and it has been compared with the built-in MATLAB function `funm` and the MATLAB function `cosm` based on the Padé algorithm given in [2]. Numerical tests show that `cosh` has lower computational cost and higher accuracy than `funm` and `cosm` for several orders of the Hermite approximation, reaching its best performance when the Hermite matrix polynomial approximation of order  $N = 16$  is used.

## References

- [1] S. M. Serbin, S. A. Blalock. An algorithm for computing the matrix cosine, *SIAM J. Sci. and Stat. Comput.*, 1(2): 198–204, 1980.
- [2] G. I. Hargreaves, N. J. Higham. Efficient algorithms for the matrix cosine and sine, *Numer. Algorithms*, 40: 383–400, 2005.
- [3] S. Serbin. Rational approximations of trigonometric matrices with application to second-order systems of differential equations, *Appl. Math. Comput.*, 5(1): 75–92, 1979.
- [4] C. B. Moler, C. V. Loan. Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later\*, *SIAM Rev.*, 45: 3–49, 2003.
- [5] S.-T. Yau, Y. Y. Lu. Reducing the symmetric matrix eigenvalue problem to matrix multiplications, *SIAM J. Sci. Comput.*, 14(1): 121–136, 1993.
- [6] E. Defez, L. Jódar. Some applications of Hermite matrix polynomials series expansions, *J. Comput. Appl. Math.*, 99:105–117, 1998.
- [7] N. J. Higham M. I. Smith. Computing the matrix cosine, *Numer. Algorithms*, 34: 13–26, 2003.
- [8] E. Defez, J. Sastre, J. J. Ibáñez, P. A. Ruiz. Computing matrix functions solving coupled differential models, *Math. Comput. Model.*, 50(5-6): 831–839, 2009.

- [9] M. S. Paterson, L. J. Stockmeyer. On the number of nonscalar multiplications necessary to evaluate polynomials, *SIAM J. Comput.*, 2(1): 60–66, 1973.
- [10] J. Sastre, J. J. Ibáñez, E. Defez, P. A. Ruiz. Efficient orthogonal matrix polynomial based method for computing matrix exponential, *Appl. Math. Comput.*, 217(14): 6451–6463, 2011.
- [11] P. I. Davies, N. J. Higham. A Schur–Parlett algorithm for computing matrix functions, *SIAM J. Matrix Anal. Appl.*, 25(2): 464–485, 2003.
- [12] L. Jódar, R. Company. Hermite matrix polynomials and second order matrix differential equations, *J. Approx. Theory Appl.*, 12(2): 20–30, 1996.
- [13] E. Defez, M. M. Tung, J. Sastre. Improvement on the bound of hermite matrix polynomials, *Linear Algebra Appl.*, 434(8): 1910–1919, 2011.
- [14] N. J. Higham, *Functions of Matrices: Theory and Computation*. Philadelphia, PA, USA, Society for Industrial and Applied Mathematics, 2008.
- [15] N. J. Higham. The scaling and squaring method for the matrix exponential revisited, *SIAM J. Matrix Anal. Appl.*, 26(4): 1179–1193, 2005.
- [16] N. J. Higham. *The Test Matrix Toolbox for MATLAB*, Numerical Analysis Report No. 237, Manchester, England (Dec. 1993).
- [17] T. Wright. Eigtool, version 2.1, <http://web.comlab.ox.ac.uk/pseudospectra/eigtool/>, (2009).
- [18] R. C. Ward, Numerical computation of the matrix exponential with accuracy estimate, *SIAM J. Numer. Anal.*, 14(4): 600–610, 1977.
- [19] I. Najfeld, T. F. Havel. Derivatives of the matrix exponential and their computation, *Adv. Appl. Math.*, 16: 321–375, 1995.
- [20] C. S. Kenney, A. J. Laub. A Schur–Fréchet algorithm for computing the logarithm and exponential of a matrix, *SIAM J. Matrix Anal. Appl.*, 19(3): 640–663, 1998.

- [21] D. Westreich. A practical method for computing the exponential of a matrix and its integral, *Commun. Appl. Numer. M.*, 6 (1990) 375–380.
- [22] Y. Y. Lu. Computing a matrix function for exponential integrators, *J. Comput. Appl. Math.*, 161: 203–216, 2003.
- [23] L. Dieci, A. Papini. Padé approximation for the exponential of a block triangular matrix, *Linear Algebra Appl.*, 308: 183–202, 2000.
- [24] E. D. Dolan, J. J. Moré. Benchmarking optimization software with performance profiles, *Math. Program.*, 91: 201–213, 2002.

# A random network model for EEG

L. Acedo\*, J.-A. Morano\*, J.-C. Cortés\* and  
D. Martín†

(\*) Instituto de Matemática Multidisciplinar,  
Universitat Politècnica de València,  
Edificio 8G, Piso 2, 46022 Valencia, España.

(†) CES Felipe II,  
Universidad Complutense de Madrid,  
Edificio Isabel de Farnesio, 28300 Aranjuez, Madrid, España.

November 30, 2011

## 1 Introduction

The history of electroencephalography (EEG) begins with Hans Berger who discovered in 1929 that between electrodes attached to the human scalp a potential difference in the mV range could be detected by means of a precision galvanometer [1]. He observed that a rhythmic pattern with a frequency of 8-12 Hz was recorded from subjects with their eyes closed (Alpha rhythm or Berger's wave). After opening their eyes the frequency increased to 12-30 Hz (Beta rhythm). Alpha power was larger with eyes closed than with eyes open, and it was associated with a relaxed brain.

In his work, Berger was inspired by the findings of the surgeon Richard Caton who in 1875 measured electrical potentials on the cortex of laboratory animals. The discovery of intracranial measures of electrical activity preceded by half a century the epoch-making Berger's discovery. The importance of Berger's method is that, being an extracranial and non-invasive technique, could develop into a very useful monitoring and diagnosing tool for neurologists and psychiatrists.

EEG discovery was ignored for almost a decade. Most researchers thought that these small currents were artifacts of the experimental apparatus and the human body. The work of Berger was eventually recognized and EEG developed into a field with vast applications in Neurological Diagnosis, Experimental Psychology and Psychiatry [2]. In 1937, A. Lee Loomis and his collaborators classified the different stages of sleep in relation with the

---

\*e-mail: luiacrod@imm.upv.es

EEG signals [3]. At that time, it was clear that during sleep the EEG pattern suffered changes that could be recognized by counting the number of zero crossings (as a qualitative measure of frequency) and its amplitude. These transitions are:

- Stage 1: As observed by Berger during awareness the brain produces the alpha wave with a mixture of frequencies in the range 8-12 Hz. This is a fast rhythm with low amplitude. When the individual becomes drowsy the pattern changes to the theta wave with lower frequencies in the range 4-7 Hz.
- Stage 2: In this stage the basal EEG is sporadically interrupted by K-complexes (high-voltage peaks around 100  $\mu V$  occurring with a periodicity of 1.0-1.7 minutes) and the so-called sleep spindles or sigma waves with a frequency 13-17 Hz appearing every 0.5-1.5 seconds.
- Stage 3: This is the most deep stage of sleep. It is characterized by delta waves ranging from 0.5 to 2 Hz and an amplitude from the negative to the positive peak around 75  $\mu V$ . The small frequency of these waves has suggested the name “slow-wave” sleep for this stage.

In experiments with insects it has been found that intrinsic oscillations around 20 Hz are induced by different odours. These oscillations are generated as a global pattern involving feedback with the mushroom bodies structures of the insect protocerebrum and they are not observed as local field potentials recordings in the antennal lobes [4]. These oscillations have also been observed in bees and they are regarded as the encoding in the neural assembly of the odour stimulus. These oscillations are different for very similar odours (1-hexanol and 1-octanol), although the discriminant capacity is impaired by the action of antagonists of the  $\gamma$ -aminobutyric acid receptors such as picrotoxin [5]. In humans neural oscillations are related with cognitive processes: memory encoding (theta wave), attention (alpha and gamma) and conscious awareness of meaningful visual patterns (synchronous gamma oscillations through separate brain areas).

It is our objective to develop a random network model of the brain in order to study the generation of these oscillatory patterns of activity both in humans and in animals. In our model we consider a network with a large number of artificial neurons simulated as automata with three states: firing, quiescent or resting and refractory. The connections among these neurons mimic the synapses in real brains. The average number of firing neurons is the mathematical measure of activity and it can be studied as a function of time.

## 2 The Brain as a Random Network Model

The cytoarchitecture of the human cortex is characterized, in general, by stratified layers of neurons. This basic structure is already present at birth but dendritic arbors develop and grow during the first two years. It is known that the details of this fine substructure develops throughout many years [6, 7]. In our approach, we are more interested in

the topological properties of these networks and we should ignore the stratified architecture. We should choose the Erdős-Rényi random network model characterized by a Poisson distribution of contacts among nodes with a mean value  $k$ .

Random networks are characterized by the number of sites or nodes ( $N = 10^6$  in our simulations) and the average number of contacts of every individual,  $k$  (called the degree of this node). Consequently, the number of links in the network is given by  $Nk/2$ . These links are randomly assigned to pairs of nodes with the obvious rule that, at most, only a link can connect two nodes.

Once the random network is generated we apply an evolution algorithm in order to analyze the number of active neurons as a function of time. As the initial state we consider only a small fraction of firing neurons. All quiescent neurons are then checked iteratively and they start to fire at time step  $t + 1$  with probability  $\alpha$  for every contact with a firing neuron at time  $t$ . The average time a neuron remains in the firing state is  $1/\nu$  whereas the average refractory time is  $1/\gamma$ . These firing and refractory times are intrinsic to the neuron structure and we assume that they follow a Poisson distribution with the aforementioned averages  $1/\nu$  and  $1/\gamma$ . BOINC computing architecture has been used in order to obtain results in a reasonable time [8].

In our stochastic model we are simulating the transmission of action potentials from neuron to neuron as a stochastic process. In practice, neurons integrate the signals received from the axon projections of their neighbors, this is the basic idea of the so-called integrate-and-fire models. If a certain threshold is reached the quiescent neuron start to fire. In our model this collective effect is captured by the independent stochastic interactions of a quiescent neuron with their firing networks in its neighborhood. If we have  $F(t)$  firing neurons in contact with a given quiescent neuron at time  $t$  the probability for this neuron to start firing at  $t + 1$  is  $F(t)\alpha$  and, consequently, increase linearly with the number of firing neurons in their topological vicinity.

It is difficult to ascribe a definite value to  $\alpha$  according to our present understanding of brain physiology. However, some recent physiological studies have revealed that the levels of glutamic acid, the main excitatory neurotransmitter in the cerebral cortex, increase after sleep deprivation in rats [9, 10, 11, 12]. The measurements of cerebral glucose utilization by means of the positron emission tomography technique also reveal a decline in several areas of the brain during sleep [13]. These observations are consistent with a scenario in which the homeostatic equilibrium of neurotransmitters is different between sleep states and the awoken state and could be mimicked by adjusting the values of  $\alpha$ .

In the case of external stimuli that promote the opening of sodium channels in sensory neurons the role of  $\alpha$  is clearer as a probability that measures the increase of firing probability of the stimulated neurons. Furthermore, the mean firing time,  $1/\nu$ , and the mean refractory time,  $1/\gamma$ , can, in principle, be obtained from the physiology of a single neuron. Values of the refractory period around 0.8-1.1 ms were reported early in the literature [14]. The firing phase should be comprised between the polarization and depolarization stages where the spike train is generated and should be much shorter.

In our study we should consider the values  $1/\nu = 10$  and  $1/\gamma = 200$ . Similar results are obtained for other proportions between  $\nu$  and  $\gamma$ . We have analyzed the behavior found

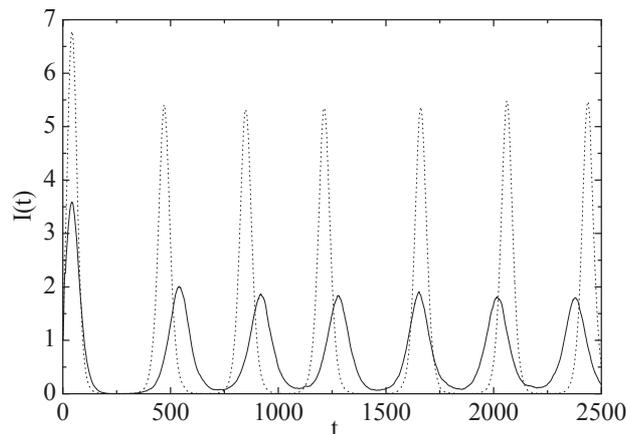


Figure 1: Number of firing neurons in the random network,  $I(t)$ , measured in units of 10,000 for an average degree of each node  $k = 500$  and  $\alpha = 2.6 \times 10^{-4}$  (solid line),  $\alpha = 2.9 \times 10^{-4}$  (dotted line). Oscillatory behaviour is observed in these two cases.

with special interest in the periodic oscillations of the brain activity.

Some typical oscillatory behaviour is plotted in Fig. 1 for a network with one million neurons and an average connectivity of  $k = 500$  among them. The sensitivity of this behaviour is very high. A small perturbation of  $\alpha$  causes that the oscillations is replaced by damping behaviour. This fine tuning simulates the role of neural oscillations to discriminate among different but similar stimuli.

## Acknowledgements

This work has been partially supported by the Spanish M.C.Y.T. grant MTM2009-08587 and the Universitat Politècnica de València grant PAID06-09-2588.

## References

- [1] H. Berger, Über das Elektroenkephalogramm des Menschen, *Archiv für Psychiatrie und Nervenkrankheiten*, 87 (1929) 527–570.
- [2] S. Tong and N. V. Thakor (editors), *Quantitative EEG Analysis: Methods and Clinical Applications*, Artech House, Norwood, MA, 2009.
- [3] A. L. Loomis, E. N. Harvey and G. A. Hobart, Cerebral states during sleep, as studied by human brain potentials, *Journal of Experimental Psychology*, 21(2) (1937) 127–144.
- [4] G. Laurent and M. Naraghi, Odorant-induced oscillations in the mushroom bodies of the locust, *The Journal of Neuroscience* 14(5) (1994) 2993-3004.

- [5] M. Stopfer, S. Bhagavan, B. H. Smith and G. Lauren, Impaired odour discrimination on desynchronization of odour-encoding neural assemblies, *Nature* 390 (1997) 70–74.
- [6] J. M. Fuster, *The Prefrontal Cortex*, Elsevier Academic Press, Amsterdam, Fourth Edition, 2008.
- [7] M. Marín-Padilla, *The Human Brain: Prenatal Development and Structure*, Springer-Verlag, Berlin, 2011.
- [8] Berkeley Open Infrastructure for Network Computing, <http://boinc.berkeley.edu>
- [9] C. M. MCDermott, M. N. Hardy, N. G. Bazan, and J. C. Magee, Sleep deprivation-induced alterations in excitatory synaptic transmission in the CA1 region of the rat hippocampus, *J. Physiol.* 570(3) (2006) 553–565.
- [10] S. Majumdar S and B. N. Mallick, Increased levels of tyrosine hydroxylase and glutamic acid decarboxylase in locus coeruleus neurons after rapid eye movement sleep deprivation in rats, *Neurosci. Lett.* 338(3) (2003) 193–196.
- [11] H. Murck, T. Struttmann, M. Czisch, T. Wetter, A. Steiger, and D. P. Auer, Increase in amino acids in the pons after sleep deprivation: a pilot study using proton magnetic resonance spectroscopy, *Neuropsychobiology* 45(3) (2002) 120–123.
- [12] L. Bettendorff, M. Sallanon-Moulin, M. Touret, P. Wins, I. Margineanu, and E. Schofeniels, Paradoxical sleep deprivation increases the content of glutamate and glutamine in rat cerebral cortex, *Sleep* 19(1) (1996) 65–71.
- [13] P. Maquet, D. Dive, E. Salmon, B. Sadzot, G. Franco, R. Poirrier and G. Franck, Cerebral glucose utilization during sleep-wake cycle in man determined by positron emission tomography and [18F]-2-fluoro-2 deoxy-D-glucose method, *Brain Research* 513 (1990) 136–143.
- [14] E. T. Rolls, Absolute refractory period of neurons involved in MFB self-stimulation, *Physiology & Behavior* 7(3) (1971) 311–315.

# A unifying model to measure consensus solutions in a society

J. C. R. Alcantud\*, R. de Andrés Calle\*<sup>†</sup>, and J. M. Cascón\*

(\*) Universidad de Salamanca

Facultad de Economía y Empresa, 37007 Salamanca, Spain

(<sup>†</sup>) Presad Research Group

November 30, 2011

## 1 Introduction

A classical group decision making problem arises when a group of voters or experts have to make a decision on a set of alternatives, candidates, etc. The experts' opinions about the alternatives are usually characterized by their ideas, principles, knowledge, etc. which is a source for conflict when it comes to making a collective decision. There are many aspects which govern the final decision in group decision making problems. Two such aspects are the experts opinions (expression domain, criteria to evaluate, expertss knowledge on the alternatives or candidates, etc.) and the methodology followed to compute the final decision (aggregation operators, voting systems, etc.).

Due to this fact, it may be interesting not only to know the result of aggregating the individuals' opinions but also to measure how much agreement with respect to the individual opinions this final decision has conveyed. This point is important in order to be aware of which decision making methodologies or voting systems could capture the coherence within the group better.

In this work we contribute to the formal analysis of the measurement of consensus in a society by taking advantage of a general method to measure

---

\*e-mail: {jcr,rocioac,casbar}@usal.es

consensus which unifies previous methodologies (cf., Alcantud et al. [1]). That model for addressing the problem can be specialized via two ways: the “voting rule” that is selected, and the measure of agreement between profiles and orderings. We here focus our analysis on two relevant and specific subcategories: the Borda and the Copeland rules under a Kemeny-type distance, whose computational construction and performance for small societies is analyzed.

## 2 Algorithmic construction of RCM-B and RCM-C

Our generic proposal (see [1] for details) includes the referenced consensus measures (RCMs) given by the tie-breaking Borda and Copeland rules as particular instances. Let  $\partial_{\mathcal{K}}$  denote the Kemeny’s measure between a profile of complete preorders and a preorder. Then RCM-B is defined as  $\mathbf{M}_B = (\mathcal{C}_B, \partial_{\mathcal{K}})$ , where  $\mathcal{C}_B$  assigns with a profile its consensus preorder by the Borda rule, and RCM-C is defined as  $\mathbf{M}_C = (\mathcal{C}_C, \partial_{\mathcal{K}})$ , where  $\mathcal{C}_C$  assigns with a profile its consensus preorder by the Copeland rule.

Below we present three algorithms that permit to compute both RCMs.

**Kemeny’s measure:**  $\mathcal{K}(R_1, R_2)$  is the probability that the binary ordering between two randomly chosen alternatives is the same for the two voters.

Given two weak orders  $R_1$  and  $R_2$  on  $k$  options, their respective  $k \times k$  preference matrices  $\mathbf{P}_1$  and  $\mathbf{P}_2$  are:

$$\mathbf{P}_s(i, j) = \begin{cases} 1 & \text{if agent } s \text{ strictly prefers } i \text{ to } j \\ 0 & \text{otherwise} \end{cases} \quad (s = 1, 2)$$

```

d = 0
For i = 1 to k
  For j = i + 1 to k
    If (
      P1(i, j) + P2(i, j) == 2
      OR P1(j, i) + P2(j, i) == 2
      OR P1(i, j) + P2(i, j) + P1(j, i) + P2(j, i) == 0 )
      Then
        d = d + 1;
      EndIf
    EndFor
  EndFor
EndFor

$$\mathcal{K}(R_1, R_2) = \frac{2d}{k(k-1)}$$


```

**RCM-B: Referenced consensus measure Borda**

**Input:** A profile  $\mathcal{R} = (R_1, \dots, R_N)$  of  $N$  weak orders on  $k$  options

- (1) For each preorder compute its  $k \times k$  preference matrix  $P_s(i, j)$
- (2) Calculate the aggregate preference matrix:  $A(i, j) = \sum_{s=1}^N P_s(i, j)$
- (3) Compute rule score  $A_B = A - A^t$ ,  $R_{\mathcal{R}}(i) = \sum_{j=1}^N A_B(i, j)$
- (4) Compute rule preference matrix yielding  $R_{\mathcal{R}}$

$$P_{\mathcal{R}}(i, j) = \begin{cases} 1 & \text{if } R_{\mathcal{R}}(i) > R_{\mathcal{R}}(j) \\ 0 & \text{otherwise} \end{cases}$$

- (5) Compute ‘individual’ measures  $\mathcal{K}(R_s, R_{\mathcal{R}})$  for each  $s = 1, \dots, N$

**Output:** Measure  $\mathcal{K}(\mathcal{R}, R_{\mathcal{R}}) = \frac{\sum_{s=1}^N \mathcal{K}(R_s, R_{\mathcal{R}})}{N}$

The RCM-C algorithm is the same as the RCM-B above, except in that step (3) is replaced by:

- (3) Compute rule score  $A_C = \text{sign}(A - A^t)$ ,  $R_{\mathcal{R}}(i) = \sum_{j=1}^N A_C(i, j)$ .

### 3 Data for comparing consensus via Borda and Copeland

In this section we carry out a computational exploration of the behaviour of our proposals for  $k = 3$  and  $k = 4$  alternatives and small societies (the case  $k = 2$  admits a simple and complete analysis: see [1] for details). Besides we restrict our study to the case where the voters linearly order the alternatives, both for expository and computational reasons. Of course, even though all voters have linear orders the models can produce ties in the consensus preorder. Table 1 shows the respective performance when  $k = 3$  and  $k = 4$ . We provide the number of cases where the Borda and Copeland methods convey the same consensus, resp., Borda gives a higher consensus than Copeland. The table conveys relative values (in larger types) and absolute values. We observe that RCM-C performs better than RCM-B when the number of experts is odd, and that the situation is the opposite for even-numbered groups (cf., Fig. 1).

Table 1: Comparing RCMs for 3 and 4 options

		$k = 3$							
		Experts							
		$n = 3$	$n = 4$	$n = 5$	$n = 6$	$n = 7$	$n = 8$	$n = 9$	$n = 10$
		56	126	252	462	792	1,287	2,002	3,003
B=C	79%	44	90%	74%	82%	73%	78%	73%	75%
			114	186	378	582	993	1,468	2,253
B>C	0%	0	10%	2%	13%	4%	14%	5%	15%
			12	6	60	30	186	90	450
B<C	21%	12	0%	24%	5%	23%	8%	22%	10%
			0	60	24	180	108	444	300

		$k = 4$							
		Experts							
		$n = 3$	$n = 4$	$n = 5$	$n = 6$	$n = 7$	$n = 8$	$n = 9$	$n = 10$
		2,600	17,550	98,280	475,020	2,035,800	7,888,725	28,048,800	92,561,040
B=C	57%	44	72%	46%	59%	43%	52%	42%	49%
			114	186	378	582	993	1,468	2,253
B>C	6%	0	18%	14%	24%	18%	27%	20%	29%
			12	6	60	30	186	90	450
B<C	37%	12	10%	40%	17%	39%	20%	38%	22%
			0	60	24	180	108	444	300

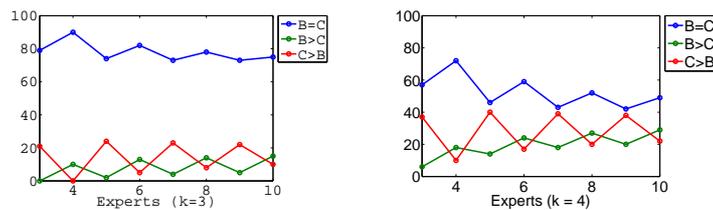


Figure 1: Comparison between RCM-B and RCM-C for  $k = 3$  and  $k = 4$ .

## 4 Concluding remarks

In this paper we analyse the measurement of consensus from a descriptive point of view. We have given two particular specifications of a unifying model, and their relative performance for small societies has been explored.

## References

- [1] J. C. R. Alcantud, R. de Andrés Calle and J. M. Cascón. Measurement of consensus with a reference. MPRA Paper No. 32155, <http://mpra.ub.uni-muenchen.de/32155/>, 2011.

# Contrast of Two Methods of Reconstruction of CT Images Using High Performance Computing

L. Flores<sup>\* \*</sup>, V. Vidal<sup>\*</sup>, P. Mayo<sup>†</sup>, F. Rodenas<sup>‡</sup>, and G. Verdú<sup>\*\*</sup>

(<sup>\*</sup>) Departamento de Sistemas Informáticos y Computación,

Universidad Politécnica de Valencia, Camino de Vera s/n, 46022, Valencia, Spain,

(<sup>†</sup>) Servicios Tecnológicos, Grupo Dominguis,

Sorolla Center, local 10 Avda. de las Cortes Valencianas, 46015 Valencia, Spain,

(<sup>‡</sup>) Departamento de Matemática Aplicada,

Universidad Politécnica de Valencia, Camino de Vera a/n, 46022, Valencia, Spain,

(<sup>\*\*</sup>) Departamento de Ingeniería Química y Nuclear,

Universidad Politécnica de Valencia, Camino de Vera s/n, 46022, Valencia, Spain.

November 30, 2011

## 1 Introduction

In medicine, the diagnosis based on CT images is fundamental for the detection of abnormal tissues by different attenuation on X-ray energy, which frequently is not clearly distinguished for radiologists. In CT imaging, a set of projections taken with a scanner is used to reconstruct the internal structure of an object.

Presently, the reconstruction process in clinical scanners is based on analytical algorithms which use the inverse Fourier transform. However, the algebraic methods represent a dominant option due to two reasons. Firstly,

---

\*e-mail: liuflo@posgrado.upv.es

the analytical methods require complete data collection which is not always possible. Secondly, they do not provide the optimal reconstruction in noisy conditions in the image [1]. Algebraic methods allow reconstructing images with higher contrast and precision in noisy conditions from a small number of projections than the methods based on the Fourier transform [2]. In CT, it is common to find incomplete set of no equally spaced projections. In these cases, algebraic reconstruction methods provide images with better quality. One of the lines of research in imaging techniques concerns portable tomography devices which employ carbon nanotubes because of their capacity to produce X-rays [4]. This type of scanners might be used to realize urgent examination at any place. They do not provide equally spaced data, so, the algebraic reconstruction is more suitable for these devices.

Let's suppose that matrix  $X$  is a digital version of the image  $f(x, y)$ . That is,  $f(x, y)$  can be approximated by the elements of the matrix  $X$ , which values represent intensities of the image. We assume that the image fits some square region and is approximated by  $n \times n$  matrix  $X$ . A single projection  $P_k$  taken at angle  $\phi_r$  can be expressed by the following formula:

$$\sum_{i=1}^n \sum_{j=1}^n W_{ij}(k, \phi_r) X_{ij} = P_{k, \phi_r}. \quad (1)$$

In matrix form the equation (1) is given as follows:

$$Ax = P, \quad (2)$$

where the system matrix  $A$  is formed by the elements  $W_{ij}(k, \phi_r)$ , simulates TAC functioning and may not be square. For a given angle, we assume that the number of projections ranges from 1 to  $m$ . If there are  $k$  different angles, then in (2)  $P$  is the column matrix with  $mxk$  elements,  $x$  is the column matrix with  $n^2$  elements, and  $A$  is the  $mk \times n^2$  rectangular matrix.

Fundamentally, the algebraic methods of image reconstruction from projections are schemes for solving a linear system of equations (2). The dimensions of  $A$  grow proportionally to the resolution of the image to be reconstructed and the number of projections, increasing therefore the computational cost. Today the technology provides the possibility to parallelize calculations and allows more efficient management of the resources of a system. In this work we employ PETSc library [5] for parallel solving of the system (2).

## 2 Methodology

Matrix  $A$  simulates the TAC functioning and many properties of the reconstructed image depend on the approximations when calculating the matrix. In this work we use the Siddon algorithm to calculate elements of the matrix in a rectangular grid. For experimental purposes we worked with 200 fan-beam real projections acquired from the hospital Clinico Universitario in Valencia. We reconstructed images with both analytical  $FBP$  and algebraic  $LSQR$  methods with the aim to analyze the possibility of the reconstruction from a fewer number of projections and to perform the quality comparison of the reconstructed images. As a measure of the quality we used Mean Square Error ( $MSE$ ) and Peak Signal-to-Noise Ratio ( $PSNR$ ) functions.

## 3 Results and Discussion

The results have been measured on a multiprocessor system with 8 processors Intel(R) Xeon(R) CPU 5365 @3GHz with 4 cores each one of them and 4GB RAM. For an image of 256x256 pixels the reconstruction time (in seconds) is summarized in Table 1. It can be seen that the paralelization speeds up a great deal the whole reconstruction process.

Table 1: The time of the reconstruction (in seconds)

Number of projections	Number of processors			
	Np=4	Np=8	Np=16	Np=32
256x200	3000	1681	882	528
256x100	779	427	132	67
256x50	193	104	55	29

Some of the reconstructed images from equally spaced projections are presented in Figure 1. Figure 2 illustrates the capacity of the iterative algorithm to reconstruct images from incomplete and unequally spaced data. This quality might be useful in portable tomography devices that produce incomplete set of projections.

The obtained results show the capacity of the algebraic methods to reconstruct images of a good quality and with low computational cost without the necessity of the complete and equally spaced experimental data. We expect

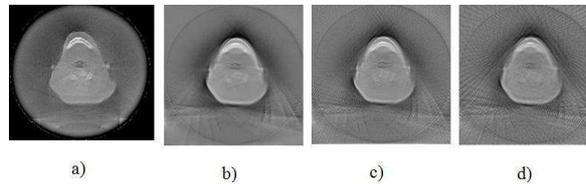


Figure 1: Reconstructed images: a) original image; b),c),d) - iterative reconstruction from 200, 100, 50 angles at the iteration 12 when the given tolerance is achieved.

more significant results in undergoing work of 3D image reconstruction where a lot of computation is involved.

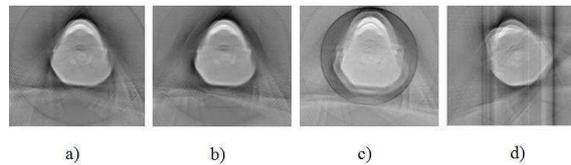


Figure 2: Reconstruction from incomplete data: -iterative (a) and analytical (c) reconstruction from the set (256x170) with removed angles; -iterative (b) and analytical (d) reconstruction from the set (226x200) with removed sensors.

## References

- [1] G. Wang, H. Yu, and B. De Man. An outlook on X-ray CT research and development. *Medical Physics*, 35 (3): 1051–1064, 2008.
- [2] R. S. Deans. *The Radon Transform and Some of Its Applications*. Dover Publications, INC. Mineola, New York, 2007.
- [3] R. G. Wells, M. A. King, P. H. Simkin, P. F. Judy, A. B. Brill, H. C. Giord, R. Licho, P. H. Pretorius, P. B. Schneider, and D. W. Seldin. Comparing Filtered backprojection and ordered-subsets expectation maximization for small-lesion detection and localization in 67Ga SPECT. *Nucl. Med.*, 41: 1391–1399, 2000.
- [4] N. Sinha and J. T. W. Yeow. Carbon nanotubes for biomedical applications. *IEEE Trans. Nano.*, 4(2): 180–196, 2005.
- [5] <http://acts.nersc.gov/petsc/index.html>. Last access Sep. 2011

# A finite element approach for the acoustic modelling of perforated dissipative mufflers with non-homogeneous properties.

A. G. Antebas, F. D. Denia\*, A. M. Pedrosa, F. J. Fuenmayor

(\*) Centro de Investigación de Tecnología de Vehículos,

Universitat Politècnica de València,

Camino de Vera s/n, 46022 Valencia, España.

November 30, 2011

## 1 Introduction

In this work, a finite element approach is presented for modelling sound propagation in perforated dissipative mufflers with non-homogeneous properties (see Figure 1). The spatial variations of the acoustic properties can arise, for example, from uneven filling processes during manufacture [1, 2] and degradation associated with the flow of soot particles within the absorbent material [3]. First, the finite element method is applied to the wave equation for a propagation medium with variable properties [4] (outer chamber with absorbent material) and a homogeneous medium (central passage). The usual calculation of the finite element matrices has been modified to include these heterogeneous acoustic properties. For the case of a dissipative muffler, the characterization of the absorbent material is carried out by means of its equivalent complex density and speed of sound [5]. To account for the spatial variations of these properties, a coordinate-dependent function is proposed for the filling density of the absorbent material. The coupling between the outer chamber and the central passage is achieved by using the acoustic impedance of the perforated central pipe, that relates the acoustic pressure jump and the normal velocity through the perforations [6, 7]. The acoustic impedance of the perforated central duct includes the influence of the absorbent material [8] and therefore a spatial variation of the impedance is also taken into account. A detailed study is then presented to assess the influence of the heterogeneous properties and the perforated duct porosity on the acoustic attenuation performance of the muffler.

---

\*Corresponding author. Tel.: +34 963879620; fax: +34 963877629; E-mail: fdenia@mcm.upv.es

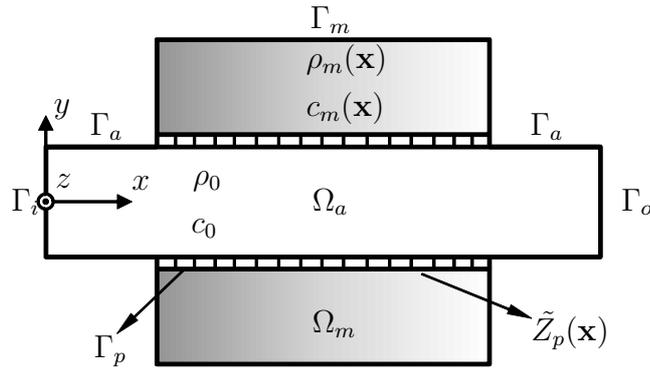


Figure 1: Perforated dissipative silencer with heterogeneous absorbent material.

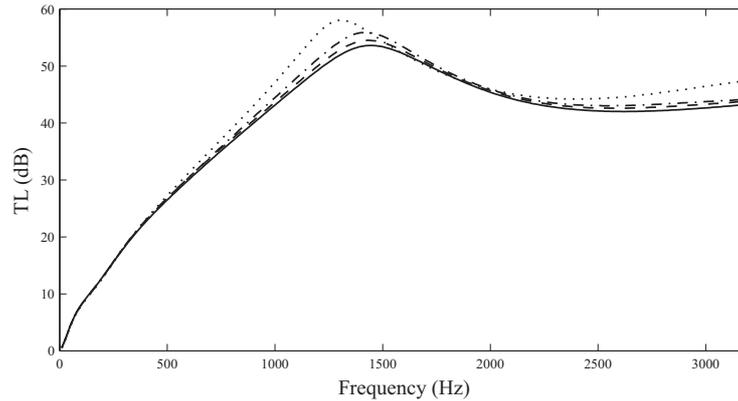
## 2 Results and discussion

### 2.1 Validation

Figure 2 shows the comparison of the transmission loss (TL) [6] predictions obtained by the implemented procedure as well as those from the segmentation method [2, 9]. The muffler geometry is axisymmetric, and the dimensions of the studied configuration are: central passage radius  $R_1 = 0.0245$  m, outer chamber radius  $R_2 = 0.0822$  m, length of the inlet and outlet ducts  $L_i = L_o = 0.1$  m and length of the absorbent material  $L_m = 0.3$  m. The perforated duct is characterized by a porosity  $\sigma = 20\%$ , thickness  $t_p = 0.001$  m and hole diameter  $d_h = 0.0035$  m. In the proposed approach for the heterogeneous absorbent materials, the filling density is assumed to vary according to the function  $\rho_c(x) = -653.333x + 359.333$ , which yields a heterogeneous steady airflow resistivity  $R(x)$ , and hence coordinate-dependent acoustic properties such as the equivalent complex density  $\rho_m(x)$  and the speed of sound  $c_m(x)$ . This assumption provides a mean filling density of value  $\rho_c = 196$  kg/m<sup>3</sup>. For the segmentation method calculations three cases have been computed with 1, 2 and 3 segments, considering a homogeneous absorbent material in each region. Therefore, the case with  $N = 1$  corresponds to a perforated dissipative muffler with homogeneous absorbent material. The lengths of segments and the associated filling densities for the three cases are shown in Table 1. In the segmentation method predictions and the proposed approach results as well, the fibrous material total mass is constant with a value of 1.137 kg. As can be seen in Figure 2, all the curves exhibit a reasonable agreement at low frequencies, where the attenuation of the muffler is mainly dictated by reactive phenomena [6]. In the mid and high frequency range, considerable differences between the implemented procedure results and the segmentation method predictions with 1 segment (homogeneous absorbent material) can be observed (at 1310 Hz the difference is about 6 dB). By increasing the number of segments, the results provided by the segmentation approach exhibit a convergence to the proposed method and the discrepancies are gradually lower. For  $N = 3$ , some slight differences are still detected between the results obtained by the proposed approach and the segmentation procedure.

$N$	Length (m)			Filling density ( $\text{kg/m}^3$ )		
	$L_1$	$L_2$	$L_3$	$\rho_{c1}$	$\rho_{c2}$	$\rho_{c3}$
1	0.30	–	–	196	–	–
2	0.15	0.15	–	245	147	–
3	0.10	0.10	0.10	261.3	196	130.7

Table 1: Lengths of the segments and the associated filling density.

Figure 2: TL of perforated dissipative silencer:  $\cdots$ , segmentation method,  $N = 1$ ;  $-\cdots-$ , segmentation method,  $N = 2$ ;  $---$ , segmentation method,  $N = 3$ ;  $—$ , implemented procedure.

## 2.2 Effect of filling density

The results have been obtained by applying the proposed approach and the segmentation method with  $N = 1$  as well (homogeneous absorbent material case). The values  $\rho_c = 90 \text{ kg/m}^3$ ,  $\rho_c = 166.7 \text{ kg/m}^3$  and  $\rho_c = 196 \text{ kg/m}^3$  are considered for the geometry analyzed in the previous figure. In the segmentation method calculations these values represent the constant filling densities within the muffler, while in the implemented procedure they denote the mean filling densities. In this latter case, the axial distributions are given by  $\rho_c(x) = -300x + 165$ ,  $\rho_c(x) = -555.733x + 305.653$  and  $\rho_c(x) = -653.333x + 359.333$ , respectively. The perforated duct is characterized by  $\sigma = 20\%$ ,  $t_p = 0.001 \text{ m}$  and  $d_h = 0.0035 \text{ m}$ . For all the considered filling densities, the transmission loss curves associated with the proposed approach are clearly different from those obtained by the segmentation method. As the filling density increases, the two approaches show higher discrepancies. Also, with higher filling densities, the transmission loss predictions obtained by the two approaches are increased at mid to high frequencies, and the resonant peaks are shifted to lower frequencies. At very low frequencies, an opposite trend appears, that is, higher filling densities tend to reduce the acoustic attenuation.

### 2.3 Effect of porosity

The influence of the perforated duct porosity has been analyzed, where the values  $\sigma = 5\%$ ,  $\sigma = 10\%$  and  $\sigma = 20\%$  have been considered while keeping  $t_p = 0.001$  m and  $d_h = 0.0035$  m. As in the latter case, the transmission loss curves have been obtained by applying the proposed approach as well as the segmentation method with  $N = 1$  in order to assess the impact of heterogeneities within the outer dissipative chamber. The absorbent material used for the calculations has a filling density  $\rho_c$  of  $196 \text{ kg/m}^3$ . This value is constant in the segmentation method and provides the mean density in the implemented procedure. In this latter case (heterogeneous absorbent material) the axial variation of the filling density is defined by  $\rho_c(x) = -653.333x + 359.333$ . For all the considered porosities, the two approaches exhibit a reasonable agreement at low frequencies, while at mid to high frequencies significant discrepancies appear between them. By increasing the porosity, the acoustic performance of the dissipative silencers is considerably improved at high frequencies since more sound attenuation is generated within the absorbent material. In addition, the resonant peaks are shifted to high frequencies due to the decrease of the mass reactance. At low frequencies, higher porosities lead to lower transmission loss.

## 3 Conclusions

A finite element approach has been developed to predict the acoustic behaviour of perforated dissipative mufflers with non-homogeneous properties. The heterogeneities of the absorbent material have been modelled by introducing a spatial variation of its filling density, that leads, through the steady airflow resistivity, to coordinate-dependent acoustic properties such as the equivalent complex density and the speed of sound. Therefore, the usual calculation of the finite element matrices has been modified to include these equivalent heterogeneous acoustic properties. In addition, the coupling between the central passage and the dissipative outer chamber has been carried out considering a perforated surface. Since the acoustic impedance of the perforations in the presence of a backing porous medium strongly depends on the acoustic properties of the fibrous material, additional features have been implemented in comparison with the models of earlier studies. Specifically, the perforated acoustic impedance models normally used in the bibliography have been modified to introduce the spatial influence of the heterogeneous properties of the fibrous material.

To validate the numerical approach, the results provided by the proposed technique have been compared with a segmentation method that considers a sequence of homogeneous regions within the outer chamber, showing a good agreement for increasing number of segments. Although the calculations exhibit some discrepancies, particularly when a reduced number of regions is considered, a suitable convergence to the proposed approach has been found for the selected configuration under analysis as the number of segments increases. Finally, a study has been presented to assess the influence of the heterogeneity, the filling density and the porosity of the perforated duct on the sound attenuation of the

dissipative mufflers.

## 4 Acknowledgments

The authors gratefully acknowledge the financial support of Ministerio de Ciencia e Innovación and the European Regional Development Fund by means of the projects DPI2007-62635 and DPI2010-15412.

## 5 References

### References

- [1] A. Selamet, M. B. Xu, I. J. Lee, and N. T. Huff. Dissipative expansion chambers with two concentric layers of fibrous materials *International Journal of Vehicle Noise and Vibration*, 1(3-4):341–357, 2005.
- [2] A. Selamet, M. B. Xu, I. J. Lee, and N. T. Huff. Effects of voids on the acoustics of perforated dissipative silencers *International Journal of Vehicle Noise and Vibration*, 2(4):357–372, 2006.
- [3] S. Allam, and M. Åbom. Sound propagation in an array of narrow porous channels with application to diesel particulate filters *Journal of Sound and Vibration*, 291(3-5):882–901, 2006.
- [4] J. E. Murphy, and S. A. Chin-Bing. A finite element model for ocean acoustic propagation and scattering *Journal of the Acoustical Society of America*, 86(4):1478–1483, 1989.
- [5] J. F. Allard, Propagation of Sound in Porous Media: Modelling Sound Absorbing Materials. Elsevier Science Publishers LTD, 1993.
- [6] M. L. Munjal, Acoustics of Ducts and Mufflers. Nueva York, Editorial, 1987.
- [7] R. Kirby, and F. D. Denia. Analytic mode matching for a circular dissipative silencer containing mean flow and a perforated pipe *Journal of the Acoustical Society of America*, 122(6):3471–3482, 2007.
- [8] R. Kirby, and A. Cummings. The impedance of perforated plates subjected to grazing gas flow and backed by porous media *Journal of Sound and Vibration*, 217(1):619–636, 1998.
- [9] F. D. Denia, A. G. Antebas, J. Martínez-Casas, and F. J. Fuenmayor. Transmission loss calculations for dissipative mufflers with temperature gradients *Twentieth International Congress on Acoustics*, Sydney, 2010.

# Numerical solution of a nonlinear equation in mathematical finance

Abraham J. Arenas<sup>†</sup> \*

Gilberto González-Parra<sup>‡</sup> and Blas Melendez Caraballo<sup>†</sup>

(<sup>†</sup>)Departamento de Matemáticas y Estadística, Universidad de Córdoba, Montería, Colombia,

(<sup>‡</sup>) Grupo de Matemática Multidisciplinar, Universidad de los Andes, Venezuela.

November 30, 2011

## 1 Introduction

In the recent several years stock option was one of the most popular financial derivatives. Numerical analysis has been applied since Merton publish a paper expanding the mathematical understanding of the options pricing model and coined the term Black-Scholes options pricing model [1]. The Black-Scholes equation governs the price of the option over time. The key idea behind the derivation was to perfectly hedge the option by buying and selling the underlying asset in just the right way and decrease the risk.

However, the Black-Scholes linear model is not very realistic, since the Black-Scholes model had been derived under very restrictive assumptions, such as frictionless, liquid and complete market, i.e. for idealized financial markets [2]. Applying Ito's calculus to the Black-Scholes model a partial differential equation is obtained. It can be approximated and integrated numerically by various methods. The basic idea is to approximate the partial differential equation by a system of equations with a finite number of unknowns, which may be solved numerically to obtain a discrete solution.

In the mathematical literature, few results can be found on the numerical solution of nonlinear Black-Scholes equations. We propose an explicit

---

\*e-mail: aarenas@correo.unicordoba.edu.co

nonstandard finite difference scheme method to solve numerically a nonlinear Black-Scholes equation modeling illiquid markets. In particular we use a method that use an exact difference scheme on the linear term and the spatial derivative is approximated using a nonstandard finite difference scheme [3]. This methodology has been applied in many sciences areas of sciences including biology and epidemic models [3, 4].

The nonlinear Black-Scholes equation modeling illiquid markets that is studied in this work was presented in [5]. These authors show the existence and uniqueness of this generalized Black-Scholes model under certain conditions of the payoff function.

## 2 Nonlinear equation in mathematical finance

In this paper we are interested in the following generalized Black-Scholes pricing PDE:

$$\frac{\partial v}{\partial t} + \frac{\sigma^2 S^2}{2(1 - \lambda(S, t)S \frac{\partial^2 v}{\partial S^2})^2} \frac{\partial^2 v}{\partial S^2} + rS \frac{\partial v}{\partial S} - rv = 0, \quad (S, t) \in \Omega_0,$$

$$v(S, T) = f(S), \quad 0 < S < +\infty,$$

where, the domain  $\Omega_0 = ]0, +\infty[ \times ]0, T[$ ;  $S$  is the underlying asset,  $v$  is the price of the option,  $T$  is the maturity,  $\frac{\partial v}{\partial S}$  is the Greek Delta of the option,  $\frac{\partial^2 v}{\partial S^2}$  is the Gamma of the option, and the payoff function  $f(S)$  of the European contingent claim is a continuous piecewise linear function. Now, we consider the change of variable  $\tau = T - t$ ,  $u(S, \tau) = v(S, t)$ , and we put  $0 < \underline{S} < \bar{S} < b$ , where  $\underline{S}$  and  $\bar{S}$  represent, respectively, the lower and upper limits of the stock price within which there is a price impact. Thus, the discretization domain is  $\Omega = ]0, b[ \times ]0, T[$ , see [6]. Therefore, the above equation can be written as

$$\frac{\partial u}{\partial \tau} - \frac{\sigma^2 S^2}{2(1 - \lambda(S, T - \tau)S \frac{\partial^2 u}{\partial S^2})^2} \frac{\partial^2 u}{\partial S^2} - rS \frac{\partial u}{\partial S} + ru = 0, \quad (S, \tau) \in \Omega, \quad (1)$$

$$u(S, 0) = f(S), \quad 0 < S < +\infty,$$

where the parameter  $r > 0$  is the interest rate and the reference volatility is  $\sigma > 0$ . Now,  $\lambda(S, \tau)$  is the function that is consistent with the price impact, and given by

$$\lambda(S, T - \tau) = \begin{cases} \frac{\gamma}{\bar{S}}(1 - e^{-\beta\tau}), & \text{if } \underline{S} \leq S \leq \bar{S}, \\ 0, & \text{otherwise,} \end{cases}$$

where  $\gamma > 0$  measures the price impact per traded share.

### 3 Proposed numerical scheme

First we construct an "exact" numerical scheme of the convection-reaction part of (1) using the ideas of the "exact" time-stepping given in [7] and the techniques of the nonstandard finite difference methodology [3]. Thus, we consider the linear equation

$$\begin{aligned} \frac{\partial u}{\partial \tau} - rS \frac{\partial u}{\partial S} &= -ru, & (S, \tau) \in (0, +\infty) \times (0, T], \\ u(S, 0) &= f(S), & 0 < S < +\infty. \end{aligned} \tag{2}$$

The equation (2) can be easily solved using the *characteristics method*, and the solution is the following

$$u(S, \tau) = f(Se^{r\tau})e^{-r\tau}. \tag{3}$$

If we evaluate the solution (3) at time  $\tau + \Delta\tau$ , such that  $\tau + \Delta\tau \in [0, T]$ , it obtains the following expression

$$u(S, \tau + \Delta\tau) = u(\hat{S}, \tau)(e^{-r\Delta\tau} - 1) + u(\hat{S}, \tau),$$

where  $\hat{S} = Se^{r\Delta\tau}$ . The "exact" difference scheme for the equation (2) is given by

$$\frac{u(S, \tau + \Delta\tau) - u(\hat{S}, \tau)}{\varphi(\Delta\tau)} = -ru(\hat{S}, \tau) \tag{4}$$

where the denominator function  $\varphi(\Delta\tau)$  is given by  $\varphi(\Delta\tau) = \frac{1-e^{-r\Delta\tau}}{r}$ . The left-hand side of (4) can be viewed as a nonstandard difference approximation of the *characteristic* derivative  $\frac{Du}{D\tau} = \frac{\partial u}{\partial \tau} - rS \frac{\partial u}{\partial S}$ , while the right-hand side represents a nonlocal modeling of the linear reaction term.

Second, we discretize the domain  $\Omega$  in the following form. We chose positive integers  $N, L$ , such that the spatial step size is given by  $h = b/N$ , while the time step size is  $k = T/L$ , and put  $S_j = jh$ ,  $h = \Delta S$ ,  $\tau_n = nk$ ,  $k = \Delta\tau$ ,  $0 \leq j \leq N$ ,  $0 \leq n \leq L$ . We seek approximations of the solution at these

mesh points, these approximate values will be denoted by  $U_j^n \approx u(S_j, \tau_n)$ . We suppose that  $u(S, \tau) \in C^{4,2}(\Omega)$ , and we define the second derivatives as

$$\frac{\partial^2 u(S_j, \tau_n)}{\partial S^2} = \frac{u(S_{j+1}, \tau_n) - 2u(S_j, \tau_n) + u(S_{j-1}, \tau_n)}{\phi(h)} + O(h^2),$$

where,  $\phi(h) = h^2 + O(h^4)$ , and  $\varphi(k) = k + O(k^2)$ , see [3]. For the sake of simplicity, we define the difference operator

$$\delta_j^n = \delta(U_j^n) := \frac{U_{j-1}^n - 2U_j^n + U_{j+1}^n}{\phi(h)}, \quad 1 \leq j \leq N - 1, \quad (5)$$

where we take the denominator function  $\phi(h) = (e^{\frac{\sqrt{2}}{3}h} - 2 + e^{-\frac{\sqrt{2}}{3}h})/\frac{2}{9}$ , see [3, 8]. Moreover, for  $h > 0$ ,  $\phi(h) > 0$ . The characterization of these functions is given by: (i)  $f(h) = h^2 + O(h^4)$  and (ii) the denominator function is a real, positive function of  $h$  over a interval  $0 < h < h_c$ , where  $h_c$  is a critical value. For some cases,  $h_c$  can be unbounded.

Using (4) and (5), we can construct an explicit nonstandard numerical scheme for the problem (1) as follow

$$U_j^{n+1} = (1 - r\varphi(k))U^n(\hat{S}_j) + \frac{\varphi(k)}{2}\beta_j^n\delta_j^n, \quad 0 \leq j \leq N - 1, \quad 0 \leq n \leq L - 1 \quad (6)$$

$$\beta_j^n = \beta_j^n(U) = \frac{\sigma^2 S_j^2}{[1 - v_j^n]^2} \geq 0, \quad 0 \leq n \leq L - 1, \quad 0 \leq j \leq N, \quad (7)$$

$$v_j^n = v_j^n(U) = \begin{cases} \gamma(1 - e^{-\beta nk})\delta_j^n, & \text{if } \underline{S} \leq S_j \leq \bar{S}, \\ 0, & \text{otherwise,} \end{cases} \quad (8)$$

where  $U^n(\hat{S}_j) \approx u(S_j e^{rk}, \tau_n)$ . Now,  $U^n(\hat{S}_j)$  is calculated using

$$U^n(\hat{S}_j) = U_j^n + \frac{\hat{S}_j - S_j}{h}(U_{j+1}^n - U_j^n), \quad 1 \leq n \leq L - 1, \quad 1 \leq j \leq N, \\ U^n(\hat{S}_0) = U_0^n.$$

Thus, we obtain the following numerical scheme

$$U_j^{n+1} = \left( rj\varphi(k) + \frac{\varphi(k)}{2\phi(h)}\beta_j^n \right) U_{j+1}^n + \left( e^{-rk} - rj\varphi(k) - \frac{\varphi(k)}{\phi(h)}\beta_j^n \right) U_j^n \quad (9) \\ + \frac{\varphi(k)}{2\phi(h)}\beta_j^n U_{j-1}^n, \quad 1 \leq j \leq N - 1, \quad 0 \leq n \leq L - 1,$$

for the interior points and in the boundary points

$$U_0^{n+1} = e^{-(n+1)kr} f(0), \quad U_N^{n+1} = e^{-rk} U_N^n + N(1 - e^{-rk})(U_N^n - U_{N-1}^n).$$

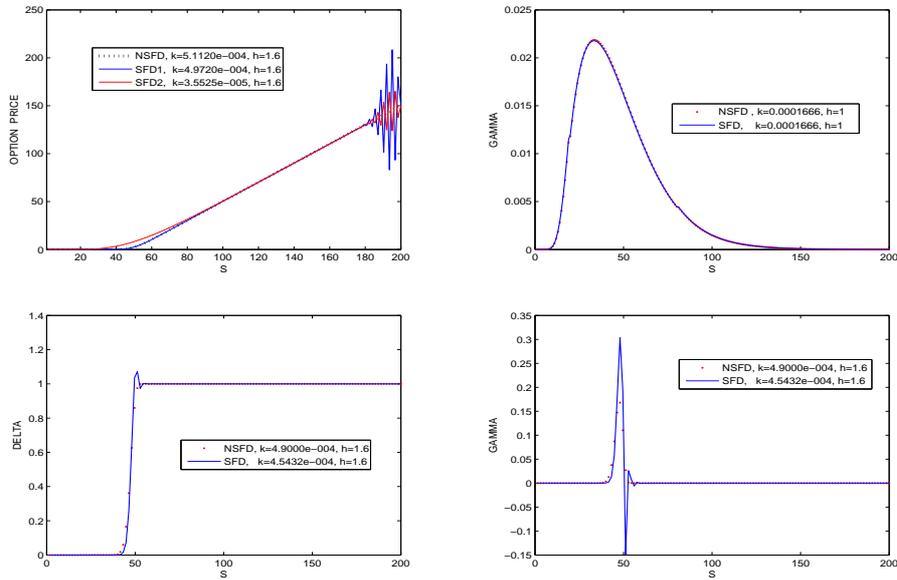


Figure 1: Consider  $f(S) = \max\{S - E, 0\}$ , with *Strike price*  $E = 50$ ,  $\underline{S} = 20$ ,  $\bar{S} = 80$ ,  $\beta = 100$ ,  $\gamma = 1$ ,  $T = 1$ . The spatial parameters are  $b = 200$ ,  $N = 125$ . Upper figures  $\sigma = 0.4$ ,  $r = 0.06$ . Lower figures  $\sigma = 0.04$  and  $r = 0.06$ .

### 4 Numerical results for the nonlinear equation in mathematical finance

In Fig. 1 we present the numerical results using the proposed nonstandard finite difference scheme (6)-(8). The first numerical result is obtained for the option price and is compared with the numerical schemes given in [2] and [6]. We can see that our scheme is able to produce an accurate solution with a large time step ( $h = 1.6$ ). Next result (from right to left) shows that our solution for the Gamma function agrees with the one proposed in [6] for a small time step. The last two figures show that our proposed scheme produces smooth and positive solutions for the Delta and Gamma function with higher values of  $k$  than the other schemes.

## 5 Conclusions

In this paper we proposed an nonstandard finite difference methodology to solve numerically a generalized nonlinear Black-Scholes equation modeling illiquid markets. In particular, the proposed method uses an exact difference scheme in the linear convection-reaction part and the spatial derivative is approximated using a nonstandard finite difference scheme. We show here that the numerical solutions of the proposed scheme, preserves the positivity.

## References

- [1] F. Black and M. Scholes. The pricing of options and corporate liabilities. *J. Pol. Econ.*, 81:637–659, 1973.
- [2] R. Company, L. Jódar, and J.R. Pintos. Numerical analysis and computing for option pricing models in illiquid markets. *Math. Comput. Model.*, 52(7-8):1066 – 1073, 2011.
- [3] R.E. Mickens. *Nonstandard Finite Difference Models of Differential Equations*. World Scientific, 1994.
- [4] A.J. Arenas, G. González-Parra, and B.M. Chen-Charpentier. A nonstandard numerical scheme of predictor-corrector type for epidemic models. *Comput. Math. Appl.*, 59(12):3740–3749, 2010.
- [5] H. Liu and J. Yong. Option pricing withan illiquid underlying asset market. *J. Econ. Dynamics Control*, 29:2125–2156, 2005.
- [6] R. Company, L. Jódar, and J.R. Pintos. A consistent stable numerical scheme for a nonlinear option pricing model in illiquid markets. *Math. Comput. Simul.*, In press, 2010.
- [7] H. Kojouharov and B.M. Chen. Nonstandard methods for the convective transport equation with nonlinear reactions. *Numer. Meth. Partial Diff. Eqns.*, 14:467–485, 1998.
- [8] S. Jerez. A nonstandard difference-integral method for the viscous burgers equation. *Appl. Math. Comput.*, 200:378–386, 2008.

# Water supply system component evaluation from GPR radargrams using a multi-agent approach <sup>\*</sup>

D. Ayala–Cabrera <sup>‡ †</sup>, J. Izquierdo<sup>‡</sup>, I. Montalvo<sup>‡</sup>  
and R. Pérez–García<sup>‡</sup>

(<sup>‡</sup>) FluIng-IMM, Universitat Politècnica de València,  
C. de Vera s/n, Edif 5C, 46022 Valencia, Spain

November 30, 2011

## 1 Introduction

Ground penetrating radar (GPR) has been extensively used as a non-destructive methodology to analyze components and anomalies in Water Supply Systems (WSS). The components most frequently analyzed are pipes and, especially, metallic pipes. Information about components, undergone changes, and anomalies is completely necessary for productive control and management of WSS. The multi-agent paradigm is used in this paper to assess components of WSS from GPR radargrams. The aim of this work is to provide non-highly qualified technicians with non-destructive, easy and quick procedures of interpretation of GPR survey files. These procedures will enable them to gain insight into the sometimes unknown layouts of the systems, and to unveil various concealed characteristics of the components of WSS. Following the same line of research on GPR image processing started in a

---

<sup>\*</sup>This work has been supported by project IDAWAS, DPI2009-11591, of the Dirección General de Investigación of the Ministerio de Ciencia e Innovación of Spain, and ACOMP/2011/188 of the Conselleria de Educació of the Generalitat Valenciana.

<sup>†</sup>contact e-mail: daaycab@upv.es

previous work by the authors [1], this paper takes the matter further by presenting a new multi-agent algorithm.

## 2 Proposed method

The agent racing will provide an interpretation and a grouping method for data from GPR radargrams. The proposed algorithm has been developed in MatLab and is based on Game Theory. The input of the agent racing algorithm is the radargram, which is a matrix of size  $m \times n$ . The dimension,  $m$ , is the volume of signal data each trace records, which depends on the characteristics of the equipment used. The proposed game is a *payoff function* specified for each player. So, the game is a function  $\pi : \prod_{s \in P} \Sigma^s \rightarrow \mathbb{R}^n$ , [2],

where  $P$  is the set of agents (players). It is a finite set, which we label  $\{1, 2, \dots, n\}$ . Each agent  $s$  in  $P$  has a finite number of strategies making up a strategy profile set,  $\Sigma$ . The  $n$  traces generated by the GPR survey are used as pseudo-parallel tracks for the  $n$  agents to compete. During the race, each agent  $s$  in  $P$  builds its vector of strategies  $k_s$ , whose  $i$ -th coordinate is the strategy taken by the agent at time  $i$ . To build these successive strategies the agent examines its associated column, its track, which we call  $b$ , in the prospection matrix. The agents' competition evolves in time from  $i = 1$  till  $i = m$ . In the competition each agent has four properties. These properties of agents are explained next.

1. *Interpretation.* For each time during the race, an agent takes one value of the trace ( $b_i$ ); then, this value is compared with two more signal values, the before-value  $b_{i-1}$  and the next-value  $b_{i+1}$ ; as a result, a binary value is generated.
2. *Decision to move.* An agent's decision to move is based on the binary value variation. According to this variation, a property called stamina varies positively (variable *StaIni*) or negatively (variable *StaEnd*). When the total stamina is zero, that is to say *StaIni* equals *StaEnd*, the agent receives its payoff for the effort performed. This is accomplished by the variable *AgeMov*. As explained in *the race phases property*, this is applied during the 'official' race, just after the *warming-up*.

3. *Movement time.* Each effort developed by an agent happens between a start time and end time. These values, associated to the agent movement (*AgeMov*) are stored in two agent personal vectors, namely, *StaTiIni* and *StaTiEnd*, respectively. Also, every agent movement (*AgeMov*), has one movement time associated *MovTi* that we define as the average time between the stamina's time start (*StaTiIni*), and the stamina's time end (*StaTiEnd*). A component of *MovTi* is defined every time the difference between these stamina values is 0.
4. *The race phases.* The race comprises two phases: a) *warming-up*, and b) *racing*. The phases are characterized by two times: a warming-up time ( $t_w$ ), and a racing time ( $t_r$ ), totaling a time  $t = t_w + t_r$ , where the  $t_w$  time corresponds to the time for the agent to overcome the end wave amplitude value (*AmplEnd*) in some percent of the average wave amplitude value for before values for the current time (*AmplProm*).

### 3 Experimental study

The case-study corresponds to the GPR image taken from one plastic pipe commonly used in WSS. The GPR image was obtained by burying the pipe in dry soil in the test tank. The following task consists on post-processing the captured GPR radargram with the proposed method. The movement times (*MovTi*) for each agent is rendered graphically and as result the Figure 1, b is generated. The input for the race (radargram) and the schematic test configuration are shown in Figures 1, a and c, respectively. Images shown in Figure 1, b present a lower amount of points than their corresponding image in Figure 1,a, thus enabling more easily interpretation. It can also be observed when comparing Figures 1, b and c, that obtained points with more similar configurations also produce more similar images. As a result, we claim that the application of multi-agent proposed method improve the insight into the subsoil properties.

### 4 Conclusions

In this paper we propose a tool for WSS component evaluation from GPR radargrams using a multi-agent approach. In the raw captured radargram without post-processing, we can see how the weakly reflective plastic pipe

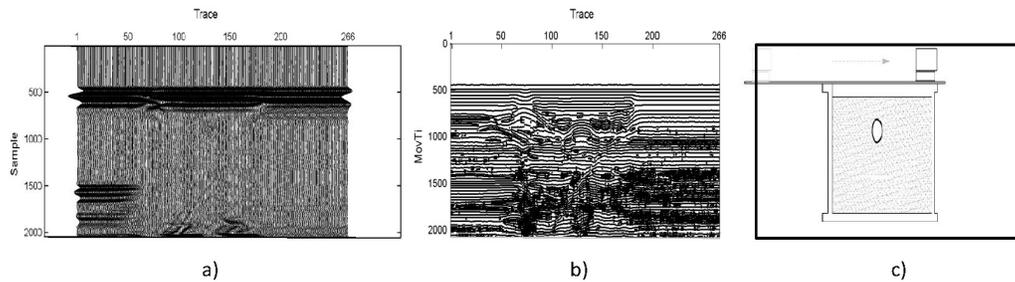


Figure 1: a) radargram, b) final image, and c) schematic configuration for test.

materials are difficult to identify. The transformation of the raw data based on the proposed multi-agent method improves the visualization of plastic pipe images by producing a better representation of the signal characteristics. It should be noted that this is a simple process that does not depend on specialist skills and is repeatable. The proposed multi-agent method is efficient on computational resources and accurate in their results. The amount of information dealt with has been reduced, while preserving its reliability. Also, the proposed method offers the possibility of more detailed analysis in terms of time with the movements of agents, which generates the possibility of better interpretations and this could serve as a basis for intelligent training systems. This will help give WSS managers more accurate vision of the systems they operate and, as a result, better service for users.

## References

- [1] Ayala-Cabrera, D., Herrera, M., Montalvo, I., Pérez-García, R., 2011. Towards the visualization of water supply system components with GPR images. *Mathematical and Computer Modelling*, Vol. 54, No. 7-8, pp. 1818-1822.
- [2] Shoham, J. and Leyton-Brown, K., 2009. *Multiagent systems: algorithmic, game-theoretic, and logical foundations*. Cambridge University, *in press*.

# Hybrid mode matching method for the efficient analysis of rods in waveguided structures

C. Bachiller\* ; H. Esteban<sup>†</sup>, J.V. Morro<sup>†</sup>, and V. Boria<sup>‡</sup>

(\*) Departamento de Comunicaciones

Universidad Politécnica de Valencia

Camino de Vera S/N, 46022 Valencia, Spain.

November 30, 2011

## 1 Introduction

Filtering is one of the main processes in any communication system. The signals arriving to the receptor must be filtered in frequency before being demodulated and processed. The different frequencies and applications define the technology used to manufacture the filters. For space communication applications the traditional filters are built with waveguided structures. The structure is divided into cavities separated by obstacles (posts, irises, etc.) which resonate at the central frequency of the filter. The number of cavities defines the order of the filter and the size and position of the obstacles determine its bandwidth.

There are many reasons that lead to develop new topologies of high frequency filters for space applications. The traditional all metal waveguide resonant cavities are heavy and bulky, in many cases they are a limiting factor in the payload of a satellite. When including dielectric rods inside the resonant cavities, the size of the whole structure is dramatically reduced. Ceramic dielectrics with high dielectric constant and temperature stability are commercially available. They allow to implement small size, good temperature stability, and high-Q dielectric resonators [1]. If the height of the

---

\*e-mail: mabacmar@dcom.upv.es

dielectric posts is the same as the height of the waveguide, that is, if the posts are all inductive or H-plane, better out of band rejection and Multi-pactor response can be obtained [2].

Nevertheless the accurate modeling of the circular dielectric resonators is far more complex than the analysis of square metal structures, since circular and rectangular geometries must be analyzed together. The authors have developed a technique for the analysis of these passive devices which consists on dividing the device in simpler building blocks: empty waveguides, steps and sections of waveguide loaded with the dielectric or metallic posts. The elements of the generalized scattering matrix are the parameters that characterize the frequency response of the structure. The matrix of each block is obtained by the most suitable analysis method, and then all the matrices are connected by a new and efficient iterative technique that provides the generalized scattering matrix of the whole filter [3].

The generalized scattering matrix of the empty sections of the waveguide is well known from the literature, and the waveguide steps can be analyzed by means of several modal techniques, such as the well know mode matching method [4] or the integral equation technique [5].

However, the analysis of waveguide sections loaded with circular obstacles is far more complex. In order to characterize these structures, circular and rectangular geometries must be considered at the same time. This complexity in the geometry makes it difficult to use purely analytic techniques, while the numerical methods are highly time-consuming. This is a serious drawback when using the simulator in a design process, since it typically demands a huge number of simulations before it finds a suitable design that fulfills the specifications. For this reason the accurate and efficient analysis of H-plane circular rods inside rectangular waveguides has received considerable attention for more than 60 years.

In this paper the authors propose a method for this analysis of the structure by decomposing the electromagnetic fields in two different regions: near the cylindrical post as cylindrical modes and far away as guided modes. The cylindrical field waves are series expansions of Bessel and Hankel functions and the field guided waves are series expansions of sinusoid functions. Then a circular boundary is used to match cylindrical and guided waves. The method increases the accuracy and stability of the analysis, no matter the number of modes or the nature of the H-plane problem under study. This is accomplished with a new mode matching procedure which uses the fast Fourier transform to solve the matching between cylindrical and guided modes. An-

alytical expressions are used to characterize the circular posts in the inner region, thus resulting in a highly efficient technique for the analysis of single or multiple posts inside a rectangular waveguide.

The results of the work were assessed by the automated design of a filter with dielectric rods by a CAD tool, the manufacture of a prototype and the measurement of its frequency response.

## **2 Results and Conclusions**

In this section a traditional all metallic filter and a filter with centered circular dielectric rods have been design to present the same electrical response. Then both filters have been manufactured and measured.

The filters were designed to have a Chebychev response in the X-band (centered at 11 GHz) with a bandwidth of 300 MHz and return losses greater than 20 dB.

### **2.1 Metal Resonant Cavities Filter.**

This is a classical filter, which has been designed and analyzed in the literature on multiple occasions. In this case, this classical filter was designed and analyzed in order to allow equal comparison with the new filter with dielectric rods also studied in this work. The filter was fabricated by the company Pinach S.L. (Alborache, Valencia, Spain) using a precision milling system and electroforming in the inner corners of the cavities. This system provides a nominal accuracy of 30 microns. The fabricated filter is shown in Figure 1. A dimensional control of the filter was commissioned to AIMME Metalworking Technology Institute. This dimensional control showed that there was a systematic drift so that all the dimensions of the manufactured filter were around 80 microns smaller than the nominal average value.

Figure 2 depicts the measured frequency response of the manufactured filter together with the simulated response of the designed filter. It can be observed that in the measured response there is an increase of the losses, due mainly to losses in the connectors used to perform the measurements, a frequency shift of about 40 MHz and a decrease in the bandwidth. This is due to the fact that the real dimensions of the resonant cavities are around 80 microns smaller than the nominal values. This makes that the central frequency increases in frequency and the bandwidth decreases slightly. The simulated

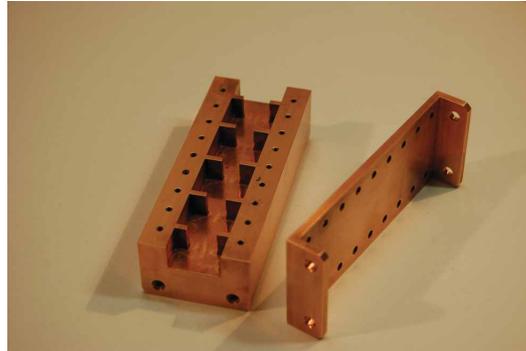


Figure 1: Manufactured filter.

response of the same filter but corrected with the dimensions provided by the dimensional control is also shown in Figure 2. It can be observed that the response of the measurements and the simulated response of the corrected filter are in good agreement, thus proving that the analysis method provides very accurate results.

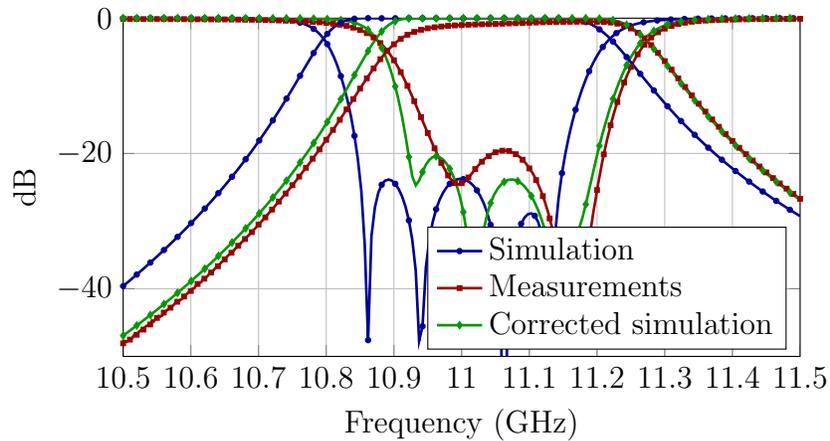


Figure 2: All metallic filter. Measurements, simulated responses (reflexion and transmission) of the designed filter, and simulated response of the filter with the dimensions provided by the dimensional control.

## 2.2 Cavities Filter with Dielectric Resonators

In this section the analysis technique is used to design the same band-pass filter of the previous section. But this time the topology of the filter is an H plane filter with coupled cavities and circular dielectric posts placed in the center of each cavity. This new filter was also manufactured by Pinach S.L. Once the metal structure was made the dielectric cylinders were fixed to it using an epoxy resin. The manufactured filter is shown in Figure 3.

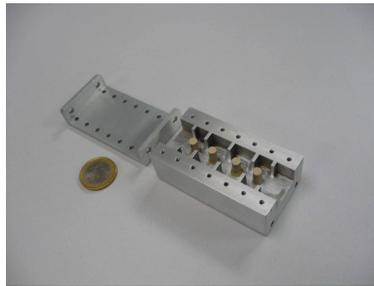


Figure 3: Manufactured filter with dielectrics.

A dimensional control was also commissioned to AIMME for the filter with dielectric rods. This time the dimensional control showed that there was again a systematic drift of around 80 microns in the inner dimensions of the metal structure, and it also showed that the diameters of the cylinders had maximum variations of  $\pm 20$  microns, but the drift was random.

When the measured and simulated results were compared, there was no agreement, even if the simulation was made with the proper dimensions provided by the dimensional control. So the permittivity of the cylinders was accurately calculated, since the manufacturer stated that the tolerances for the permittivity and loss tangent were wide and they varied with the frequency. In order to measure the permittivity and loss tangent of the dielectric rods, a section of WR90 waveguide with a small recess to place a dielectric cylinder was built. In this structure the dielectric cylinder was placed in the center and the frequency response of the structure was measured. Then the analysis tool was used to calculate the frequency response of the structure and an optimization algorithm, the Nelder-Mead simplex method, was used to iteratively change the values of permittivity and loss tangent in the simulations until the simulated response was in good agreement with the measured response. This agreement between measurement and simulations was obtained

with a permittivity of 18.85 and a loss tangent of 0.002.

Once the permittivity is fixed to 18.85 and the loss tangent to 0.002, a new simulation is performed and compared to the measurements in figure 4. It can be observed that there is good agreement in both responses in the central frequency and in the bandwidth. The difference in the return losses may be due to additional losses introduced by the connectors.

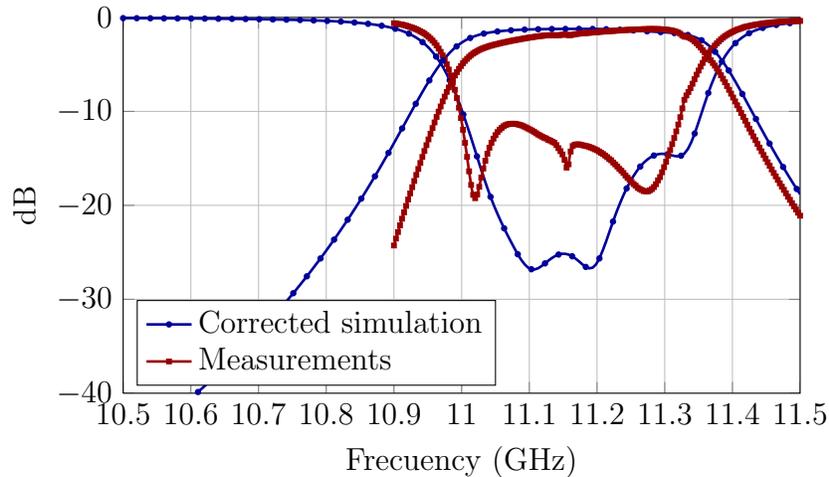


Figure 4: Reflexion and transmission of the filter simulated with  $\epsilon_r = 18,85$ ,  $\tan\delta = 2 \cdot 10^{-3}$ , compared to the measurements.

## 2.3 Conclusions

It is possible the efficient analysis of a complex waveguide structure if it is segmented adequately. In this work a waveguide topology of H-plane waveguide filters with dielectric cylinders, whose analysis and design is complex, but that has advantages over traditional topologies, has been successfully designed, manufactured and measured. For the analysis of these filters a hybrid technique developed by the authors has been used. This technique expands the field near the cylinders using open space modes (Bessel and Hankel functions), and guided modes in the accessing ports. Then a mode matching is performed between open space and guided modes and the generalized scattering matrix is obtained. All the integral needed to solve the problem can be efficiently computed using the fast Fourier transform thanks to the use of a circular boundary between open space and guided modes.

## References

- [1] P. Arcioni, V. E. Boria, M. Bozzi, G. Conciauro, and B. Gimeno, “Analysis of H-plane waveguide components with dielectric obstacles by the BI-RME method,” in *Proc. of the 32nd Eu. Microwave Conf.*, Milano, 2002, pp. 1–3.
- [2] C. Bachiller, H. Esteban, F. Diaz, V. Boria, and J. Morro, “Comparative study of multipactor breakdown in waveguide H-plane filters loaded with dielectric resonators,” in *Int. Workshop on Multipactor, Corona and Passive Intermodulation*, Valencia, Sept. 2008, pp. 1–8.
- [3] C. Bachiller, H. Esteban, V. E. Boria, A. Belenguer, and J. V. Morro, “Efficient technique for the cascade connection of multiple two port scattering matrices,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 55, no. 9, pp. 1880–1886, September 2007.
- [4] R. Harrington, *Field Computation by Moment Methods*. New York: The Mac Millan Company, 1968.
- [5] G. Conciauro, M. Bressan, and C. Zuffada, “Waveguide modes via an integral equation leading to a linear matrix eigenvalue problem,” *IEEE Trans. Microwave Theory Tech.*, vol. 32, no. 11, pp. 1495–1504, Nov. 1984.

# Mathematical modelling of the combined optimization of a pumped-storage hydro-plant and a wind park

L. Bayón\*, J.M. Grau, M.M. Ruiz and P.M. Suárez

Department of Mathematics. University of Oviedo.

E.P.I., Campus of Viesques, Gijón, 33203, Spain

## 1 Introduction

The new Spanish regulations (BOE [1]) allows wind farms to go to the market to sell the energy generated. In the case of over- or undersupply, other producers must reduce or increase their production to resolve the so-called deviation, thereby incurring financial losses. Faced with this situation, wind farms have several options. In this paper we consider one promising method: the combined optimization of a pumped-storage hydro-plant and a wind park and we present a tool to design the optimal configuration.

## 2 Mathematical optimization

When pumped-storage plants are considered, the fixed-head hydro-plant model for the active power generated  $P$  is defined piecewise as:

$$P(z') := \begin{cases} A \cdot z' & \text{if } z' \geq 0 \\ \eta \cdot A \cdot z' & \text{if } z' < 0 \end{cases} \quad (1)$$

where  $z'$  is the rate of water discharge, and  $\eta$  is the efficiency ([2]). Let  $b$  be the volume of water that must be discharged over the optimization interval

---

\*e-mail: bayon@uniovi.es

$[0, T]$ . So, the following boundary conditions will have to be fulfilled:

$$z(0) = 0, \quad z(T) = b \quad (2)$$

Besides the previous statement, we consider  $z'(t)$  to be bounded by technical constraints

$$q_{\min} \leq z'(t) \leq q_{\max}, \quad \forall t \in [0, T] \quad (3)$$

In this section we focus on the new short-term problem faced by a generation company in a deregulated electricity market when preparing its offers for the day-ahead market. In our problem, the objective function is given by hydraulic profit over the optimization interval,  $[0, T]$ . Profit is obtained by multiplying the hydraulic production of the pumped-storage hydro-plant by the clearing price,  $\pi(t)$ , at each hour,  $t$ . Taking our objective functional  $F(z)$  in continuous time form, a standard Lagrange-type Optimal Control problem can be mathematically formulated as follows:

$$\begin{aligned} \max_{(u,z)} \int_0^T L(t, z(t), u(t)) dt &= \max_{(u,z)} \int_0^T \pi(t) P(u) dt \\ z' &= u; \quad z(0) = 0, z(T) = b; \quad u_{\min} \leq u(t) \leq u_{\max} \end{aligned} \quad (4)$$

### 3 Optimization algorithm

For the Optimal Control problem (4), the resulting Hamiltonian,  $H$ , is linear in the control variable,  $u$ , and results in an optimal singular/bang-bang control policy. In general, the application of Pontryagin's Maximum Principle [3] is not well suited for computing singular control problems as it fails to yield a unique value for the control. In our problem, however, an added complication arises: the Hamiltonian is defined piecewisely and the derivative of  $H$  with respect to  $u$  presents discontinuity at  $u = 0$ , which it is the border between the power generation zone (positive values of  $u$ ) and the pumping zone (negative values of  $u$ ). When non-differentiable objective functions arise in optimization problems, the generalized (or Clarke's) gradient ([4]) must be considered.

On the basis of the above theoretical results, in [5] we determined the optimal solution: the bang-singular-bang segments and the boundary on which the solution is situated:

$$u^*(t) = \begin{cases} u_{\max} & \text{if } A \cdot \pi(t) > -\lambda_0 \\ u_{\text{sing}} = 0 & \text{if } -\lambda_0 \in [A \cdot \pi(t), \eta \cdot A \cdot \pi(t)] \\ u_{\min} & \text{if } \eta \cdot A \cdot \pi(t) < -\lambda_0 \end{cases} \quad (5)$$

The algorithm that leads to the optimal solution comprises the following steps: (i) First, for a given  $\lambda$ , we have to determine the switching times:  $t_1, t_2, \dots$ . These instants are calculated solving the equations

$$A \cdot \pi(t) = -\lambda; \quad \eta \cdot A \cdot \pi(t) = -\lambda \tag{6}$$

(ii) Second, the optimal value,  $\lambda_0$ , must be determined in order for:

$$z_\lambda(T) = \sum_{i=1}^{N_u} \delta_i^u \cdot q_{\max} + \sum_{i=1}^{N_l} \delta_i^l \cdot q_{\min} = b \tag{7}$$

$\delta_i^u$  and  $\delta_i^l$  being the duration of the  $i$ -th bang-bang segment in the upper and lower bound respectively,  $N_u$  and  $N_l$  the number of such segments, and  $z_\lambda(T)$  the final volume obtained for each  $\lambda$ . (iii) To calculate an approximate value of  $\lambda_0$ , we propose a classic iterative method (like, for example, bisection or the secant method).

## 4 Combined optimization

With the aid of the algorithm presented in the previous section, we are now in a position to analyze the combined optimization of a pumped-storage hydro-plant and a wind farm. In this section we shall analyze whether it is in the interest of wind farms to go to market. To do so, we address two configurations (see Figure 1):

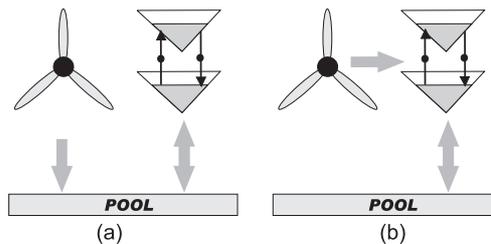


Figure 1. Two configurations.

(a) The wind farm and the pumped-storage hydro-plant work independently, each selling the energy it produces on the market.

(b) The wind farm does not sell energy on the market, but uses the generated power to pump water to the upper reservoir of the pump-plant.

There are a number of factors that influence the final result, like for example: the efficiency of the hydro-plant  $\eta$ , the volume of water available  $b$ , deviation penalties, or wind power production. The first two factors have already been analysed in [5]. We shall now analyse the last two factors, which correspond to the wind farm. In order for our study to reflect a broad range of possibilities, we propose two scenarios: (i) Low wind power production in peak hours; and (ii) High wind power production in peak hours.

## 5 Results and conclusions

A program was written using the Mathematica package to apply the results obtained in this paper to an example of a hydraulic system made up of one fixed-head pumped-storage hydro-plant and a wind farm. The problem studied in this paper analyses the convenience, or not, of the wind farms going to market. In conclusion, both configurations (a) and (b) seem interesting options. But, the system is highly sensitive to numerous factors and each company must carefully assess particular situations that may result in variations in the optimum configuration. Our algorithm allows the optimal solution to be obtained easily and the obtained results provide real-time information to determine which configuration is preferable in each specific real situation of the electricity market.

## References

- [1] Real Decreto 661/2007, [www.boe.es](http://www.boe.es)
- [2] M.E. El-Hawary and G.S. Christensen, *Optimal Economic Operation of Electrical Power Systems*, Academic Press, New York, 1979.
- [3] R. Vinter, *Optimal Control, Systems & Control: Foundations & Applications*. Birkhäuser Boston, Inc., Boston, MA, 2000.
- [4] F.H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley-Interscience, New York, 1983.
- [5] L. Bayon, J.M. Grau, M.M Ruiz and P.M. Suarez, Optimization of a pumped-storage fixed-head hydro-plant: the bang-singular-bang solution, *Math. Prob. Eng.*, Vol. (2011): 11 pages, 2011.

# Design of B-*spline* multidimensional deformable models in the frequency domain

R. Berenguer-Vidal\*, R. Verdú-Monedero and J. Morales-Sánchez

Dpto. Tecnologías de la Información y las Comunicaciones.

Universidad Politécnica Cartagena

Plaza del Hospital, 1, 30202, Cartagena (Murcia), SPAIN

November 30, 2010

## 1 Introduction

Deformable models are useful tools in image processing, used mainly to analyze the shape and movement of objects [1]. The model is obtained from the minimization of an energy functional by means of the multidimensional Euler-Lagrange equation.

Practical implementations result in a second order system, where the parameters of the mathematical model are able to emulate physical characteristics of objects: elasticity, rigidity, mass and damping. In the process, the model is influenced by external and internal forces, varying its shape adaptively to the data under analysis.

In the original formulation [2], unidimensional deformable models or active contours are defined in the spatial domain, which are difficult to generalize to a higher dimensional models. This abstract presents a frequency-based formulation based on [3], using B-*spline* as shape function in the finite element method for the spatial discretization. This process allows us to obtain a computationally efficient iterative system in matrix form, suitable for models of any dimension. These models can be used for fast segmentation and tracking of three-dimensional objects with non-rigid motion or deformation.

\*e-mail: rberenguer@pdi.ucam.edu

## 2 Formulation

A dynamic deformable model is defined as a parametric hypersurface in  $\mathbb{R}^d$ ,  $\mathbf{v}(\mathbf{s}, t)$ , where  $\mathbf{s} \equiv [s_1, \dots, s_e]$  with  $e \leq d - 1$ , is the vector of the parametric variables of the space domain. The model is governed by an energy functional  $\mathcal{E}(\mathbf{v}) = \mathcal{S}(\mathbf{v}) + \mathcal{P}(\mathbf{v})$  [1]. The first term is the internal deformation energy,  $\mathcal{S}(\mathbf{v}) = \frac{1}{2} \sum_{l=1}^d (\int_{\Omega} \alpha(\mathbf{s}) \|\nabla v_l(\mathbf{s})\|^2 + \beta(\mathbf{s}) |\Delta v_l(\mathbf{s})|^2 ds)$ , where  $\alpha$  and  $\beta$  control the elasticity and the rigidity in any coordinate  $\mathbf{s}$  of the model. The term  $\mathcal{P}(\mathbf{v})$  contains the rest of energies applied to the model.

According to the variational calculus, the model  $\mathbf{v}(\mathbf{s})$  that minimizes  $\mathcal{E}(\mathbf{v})$  must satisfy the Euler-Lagrange (E-L) equations [4], which produce a set of  $d$  decoupled partial differential equations,  $\mu(\mathbf{s})\partial_{tt}\mathbf{v}(\mathbf{s}, t) + \gamma(\mathbf{s})\partial_t\mathbf{v}(\mathbf{s}, t) - \nabla \cdot (\alpha(\mathbf{s})\nabla\mathbf{v}(\mathbf{s}, t)) + \Delta(\beta(\mathbf{s})\Delta\mathbf{v}(\mathbf{s}, t)) = \mathbf{q}(\mathbf{v}(\mathbf{s}, t))$  where  $\mathbf{q}$ ,  $\mu$  and  $\gamma$  represent the external forces, and the mass and the damping density of the model respectively.

By applying a spatial discretization by means of finite elements  $\mathbf{v} = \mathbf{f} \otimes \mathbf{u}$ , and using B-splines as shape function  $\mathbf{f}$  [5], the Garlekin's method allow us to transform the E-L equations of motion in a set of  $d$  second-order partial differential equations (PDE),  $\mathbf{M}d_{tt}\mathbf{u}_i(t) + \mathbf{C}d_t\mathbf{u}_i(t) + \mathbf{K}\mathbf{u}_i(t) = \mathbf{q}_i(t)$ , where  $\mathbf{u}_i$  are the nodes of the model for the dimension  $i$  reshaped to a column vector,  $\mathbf{M}$ ,  $\mathbf{C}$  and  $\mathbf{K}$  are the matrices of the model and  $\mathbf{q}_i$  represents the external forces.

For a practical implementation, the time is discretized  $\mathbf{u}_{\xi} = \mathbf{u}(\xi\Delta t)$ , and the time derivatives of  $\mathbf{u}$  are replaced by their discrete approximations. Thus, the system can be rewritten as,  $\eta^{-1}(\eta\mathbf{f} + \mathbf{k}) \otimes \mathbf{u}_{\xi} = a_1 \mathbf{f} \otimes \mathbf{u}_{\xi-1} + a_2 \mathbf{f} \otimes \mathbf{u}_{\xi-2} + \eta^{-1}\mathbf{q}_{\xi-1}$ , where  $\otimes$  indicates  $e$ -dimensional circular convolution.  $\eta$ ,  $\gamma$ ,  $a_1$  and  $a_2$  are constants obtained from the model parameters.

The discrete spatial domain  $\bar{\mathbf{n}}$  is translated into the frequency domain  $\bar{\omega}$  using the  $e$ -dimensional discrete Fourier Transform ( $eDFT$ ). This allows us to isolate the nodes of the model,  $\hat{\mathbf{u}}_{\xi} = \hat{\mathbf{h}}(a_1\hat{\mathbf{u}}_{\xi-1} + a_2\hat{\mathbf{u}}_{\xi-2} + (\eta\hat{\mathbf{f}})^{-1}\hat{\mathbf{q}}_{\xi-1})$ , where  $\hat{\mathbf{u}}$ ,  $\hat{\mathbf{f}}$  and  $\hat{\mathbf{q}}$  are the  $eDFT$ 's of their respective spatial sequences. This equation provides an efficient iterative system to adapt the deformable model to the data. Internal forces are imposed by filter  $\hat{\mathbf{h}}$  and external forces are applied by means of  $\hat{\mathbf{q}}$ . The  $e$ -dimensional low-pass filter  $\hat{\mathbf{h}}$  can be calculated as  $\hat{\mathbf{h}} = \eta\hat{\mathbf{f}}/(\eta\hat{\mathbf{f}} + \hat{\mathbf{k}})$ , where  $\hat{\mathbf{k}}$  is the  $eDFT$  of  $\mathbf{k}$ . The one and two-dimensional filters are defined in [6].

The spectral characteristics of filter  $\hat{\mathbf{h}}$  depend on the parameters of the model  $\alpha$ ,  $\beta$  and  $\eta$ , which control the roughness of the model. The parameter

$\gamma$  determines the dynamic response of the model. So, increasing  $\alpha$  augments the elastic behavior, the rigidity of the model is directly proportional to the value of  $\beta$  and the parameters  $\gamma$  and  $\eta$  control the inertia of the model.

### 3 Results and discussion

A two-dimensional model, is tested on two experiments involving three-dimensional medical data sets using the proposed formulation. That is,  $e = 2$  and  $d = 3$ . The first data set  $\mathcal{D}_1$  is a static three-dimensional computed tomography (CT) image of a human skull with  $256 \times 256 \times 27$  voxels. In contrast,  $\mathcal{D}_2(\tau)$  is a dynamic CT (4D) of human thorax, i.e., a time series of 3D CT images of size  $512 \times 512 \times 60$ , where  $\tau$  represents the time,  $1 \leq \tau \leq 20$ .

The model is used to segment objects in the data in both experiments. For  $\mathcal{D}_1$  the target of the model is the skull and for  $\mathcal{D}_2$  it is the external surface of the heart. A model with  $N_1 = 64$  and  $N_2 = 32$  nodes is used for both data sets. Given that the model is open for some parametric variables, extensions must be added in such variables [7]. The gradient of the data is used as external forces in the iterative process. The model is initialized in the shape of a cylindrical mesh around the objects. External forces drive the model to fit the surface of the object. The internal forces regularize the model, smoothing the discontinuities of the data. At the same time the elasticity behavior increases the speed of adaptation, stretching the nodes to the surface of the object. The model is adapted completely to the data in about 50 iterations. Figures 1a and 1b show the model fitted to  $\mathcal{D}_1$  and  $\mathcal{D}_2(\tau)|_{\tau=1}$  respectively.

### 4 Conclusions

In this abstract a formulation of  $e$ -dimensional parametric deformable models defined in the frequency domain. This approach results from application of the Euler-Lagrange equation to the multidimensional energy functional has been proposed. By discretizing the spatial domain by means of B-spline based finite elements, a system in matrix form is obtained. With this, a compact iterative system is achieved by translating the system to the Fourier domain.

As shown in the results, the model can be used effectively in applications where parametric models are suitable, that is structures with non-changing topologies and data with a limited level of concavities. Future work includes

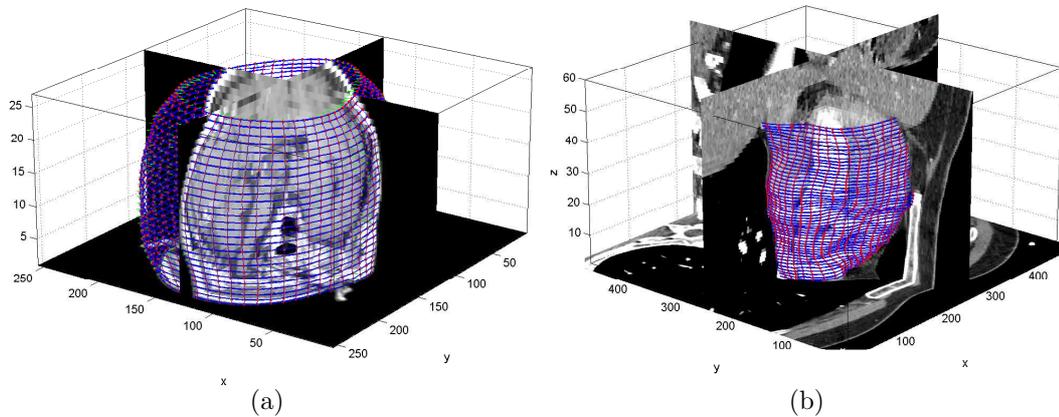


Figure 1: (a) 3D CT of a human skull, (b) 4D CT of a human thorax.

the extension of the model to other shape functions, analysis of the convergence and stability of the system, and definition of arrays of the system for models with three parametric variables,  $e = 3$ .

## References

- [1] M. Kass, A. Witkin, D. Terzopoulos, Snakes: Active contour models, *Int. J. of Computer Vision* 1 (4) (1988).
- [2] J. Liang, et al., United snakes, *Medical Image Analysis* 10 (2) (2006).
- [3] R. Verdú, J. Morales, L. Weruaga, Efficient frequency design of deformable models, *Proc. of ACIVS* (2004) 223–229.
- [4] J. Larrey-Ruiz, R. Verdú-Monedero, J. Morales-Sánchez, A fourier domain framework for variational image registration., *JMIV* 32 (1) (2008).
- [5] M. Unser, Splines: a perfect fit for signal and image processing, *Signal Processing Magazine, IEEE* 16 (6) (1999).
- [6] R. Berenguer-Vidal, et al., Characterization of three-dimensional data with multidimensional deformable models based on B-splines in the fourier domain, *Proc. of IEEE ICASSP* (2011).
- [7] L. Weruaga, R. Verdú, J. Morales, Frequency domain formulation of active parametric deformable models, *IEEE Trans. PAMI.* 26 (12) (2004).

# Adaptive modelling for automotive applications

C. Guardiola, B. Pla, D. Blanco-Rodriguez, and P. Cabrera \*

CMT Motores Térmicos, Universidad Politécnica de Valencia, Camino de Vera s/n  
E-46022 Valencia, Spain

October 10, 2011

## 1 Introduction

Control systems in engineering require reliable information about the states that mainly affect these systems. In the automotive field, engine and after-treatment performance and diagnostics, including On Board Diagnostics system (OBD), and maps calibration depend directly on this. Sensors are an alternative but can be expensive and their responses present delayed and variable dynamic performance.

Look-up tables model systems by means of a piecewise function with  $n$  inputs creating an array of  $n$  dimensions; and constitutes a good option for complex situations where it is difficult to derive an analytical expression or the processes are not well-known, such as current engine systems. However, on one hand the calibration effort is really high including big quantity of tests and human resources, and on the other hand, engine behaviour is affected by manufacturing discrepancies (unit-to-unit dispersion), ambient conditions and drift (variation of the engine behaviour over the working life), making that the used of fixed relationships is not an optimised way to treat the problem.

The design of adaptive concepts for updating these tables is a good solution. Several authors have proposed alternatives: Kalman filter [1] for

---

\*e-mail: {carguaga, benplamo, dablarod, pedcablo}@mot.upv.es

observing and updating table elements [2] with good results but high computational cost (all table elements are treated in each iteration including the ones that are locally unobservables); proportional updating concepts [3] with low computational cost but without taking into account noise in measurement and not having an optimal solution; and many others, such as [4].

Current work presents a novel learning method affecting only observable nodes and updating these by using membership functions based on a Kalman filter. This approach keeps noise treatment of Kalman filter and derives a converged solution of the Kalman problem in a proportional way, ensuring a low computational cost allowing this to implement in a commercial electronic control unit (ECU).

## 2 Problem requirements and conditions

The problem can be stated in discrete form (sampled system are represented using discrete formulation) as having a certain information about one system in a given instant  $k$   $z_k$ , (sensor, model or observation reference), and given a certain look-up table  $T_k$   $n \times D$  with  $n$  inputs  $u_{n,k}$ , designing a learning algorithm  $F$  (see (2)) that can adapt  $T_k$  elements to approximate  $z_k$  with the best accuracy, i.e. with the minimum error. For the evaluation of the method, an actual reference matrix  $T_r$  is created although this does not have validity for real applications, i.e. it is created by means of synthetic signals for simulating, and will be variable in a real context (if fast measurement systems are used,  $T_r$  can be measured). An estimation error  $e_k$  is defined in (1), where  $\hat{x}_k$  is calculated by means of linear interpolation  $q$  involving the  $2n$  observable elements.

$$\begin{aligned}\hat{x}_k &= q(T_{k-1}, u_{1,k}, \dots, u_{n,k}) \\ e_k &= z_k - \hat{x}_k\end{aligned}\tag{1}$$

This error is weighted for updating parameters of  $T_k$  in (2) coping with system variations.

$$T_k = F(T_{k-1}, u_{1,k}, \dots, u_{n,k}, e_k)\tag{2}$$

Given this scenery, the learning method must ensure the convergence of the method as fast as possible (i.e. any of the error metrics that could be defined must tend to zero in the infinite), the robustness of the estimation against some noise or perturbances in  $z_k$  and a low computational burden.

For showing the benefits of the method, a quadratic error  $eq_k$  is defined in (3) where  $\theta_{p,k}$  are  $T_k$  elements and  $\theta_{r,p,k}$ ,  $T_r$  elements in instant  $k$ , being  $eq_u$  a the maximum expected one and  $p$  the position of each element or node in the given array:

$$eq_k = 100 \sum_{p=1}^{P_{end}} (\theta_{p,k} - \theta_{r,p,k})^2 / eq_u \quad (3)$$

Note that the information about the system  $z_k$  is just the output of  $T_r$  only when noise and other perturbances  $P$  are zero:

$$z_k = f(T_r, u_{1,k}, \dots, u_{n,k}) + P \quad (4)$$

### 3 Adaption principle

The algorithm  $F$  defines the learning principle for updating the table as a function of  $z_k$  and  $e_k$ . This algorithm will take the basis of parameter identification by means of observing table elements as states [2] using a Kalman filter in a proportional way as in [3]. For this, membership functions are derived from a Kalman filter that finally weights  $e_k$  between all the involved elements in the interpolation for every iteration  $k$ . This algorithm remains weighting properties of Kalman filter and avoids numerical and computational problems associated with the elements that are unobservables in  $k$  in the standard Kalman filter.

The key aspect is the derivation of converged membership functions that are calculated taking into account white noises  $\sigma_\theta^2$  and  $\sigma_z^2$  added to the observed table elements ( $\theta_{p,k}$ ) and the output ( $z_k$ ) respectively. Then, it can be demonstrated that membership functions  $k_p$  are function only of noise trade-off  $\sigma_z^2/\sigma_\theta^2$  and inputs [5]. In every iteration and for the nD case, a maximum of  $2n$  membership functions are active while the rest of them will be zero. These functions are precalculated and stored off-line using Kalman gain properties of convergence after certain number of iterations in the given case of constant noise.

$$k_p = f(u_{1,k}, \dots, u_{n,k}, \sigma_z^2/\sigma_\theta^2) \quad (5)$$

These functions are symmetrical and an example for three different noise values and a 1D case is shown in 1 where  $\eta_1$  is the normalised distance

between  $u_1$  (only 1 input due to 1D case) and the previous closest table node where  $\eta_1 \in [0, 1[$ .

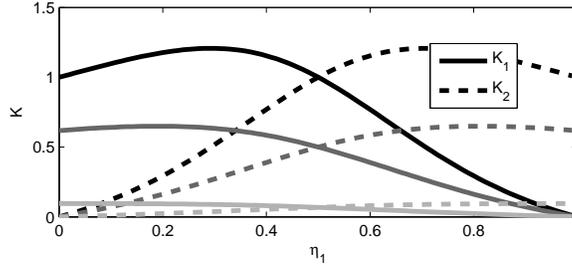


Figure 1: Membership functions:  $\sigma_z^2/\sigma_\theta^2 = 0$  (black),  $\sigma_z^2/\sigma_\theta^2 = 0.01$  (medium grey) and  $\sigma_z^2/\sigma_\theta^2 = 0.1$  (light grey).

These functions define the correction for each element of the table in each iteration  $k$ :

$$\theta_{p,k} = \theta_{p,k-1} + k_p(u_{1,k}, \dots, u_{n,k})e_k \quad (6)$$

Figure 1 shows two interesting properties: when  $u_1$  is closed to one table node ( $\eta_1$  is around 0 or 1), correction for the opposite node tends to zero; and when  $\sigma_z^2/\sigma_\theta^2$  decreases, reliability on  $z_k$  increases and  $k_p$  is bigger.

## 4 Simulation results

An analytical 1D function deriving one approximated 1D actual table  $T_r$  discretised with 6 nodes is used for checking the method and shown in left plot of figure 2. Initialising the observed matrix  $T_0$  with the same nodes and all elements zero, the objective is checked the adaption of  $T_k$  seeing the  $eq_k$  evolution, where objective is that  $T_k$  tends to  $T_r$ . Nodes take values between 1000 and 3500 with equal space between them of 500 (although irrelevant to this paper, these values are based on engine speed in rpm).

First simulation is supposing that  $z_k$  is reliable, and for this an stochastic distribution of 1000 inputs  $u$  is prepared and results are shown in right plot of 2, showing clearly how method converges in a very low number of iterations, being faster when  $\sigma_z^2/\sigma_\theta^2$  is lower, logical conclusion as far as  $P$  is supposed as zero.

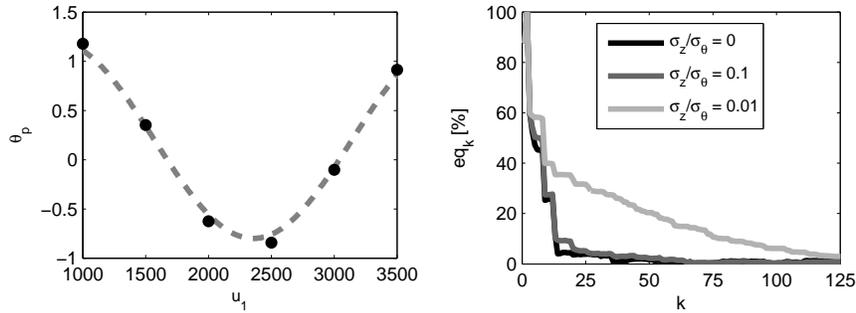


Figure 2: Left: Grey colour represents defined function while black dots the fitted table  $T_{1000}$ . Right: Evolution of  $eq_k$  in the first 125 iterations.

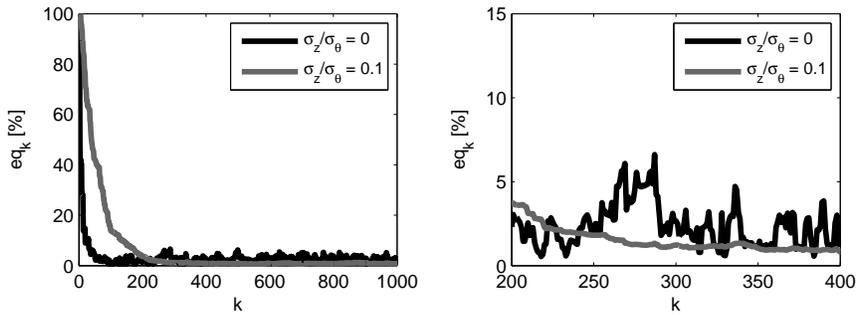


Figure 3: Evolution of  $eq_k$  in 1D case applying a P distribution with an amplitude of 0.2 in the measurement.

A second simulation with same inputs is made including some perturbation  $P$  modelled as an uniform distribution with zero mean and amplitude of 0.2 and shown in figure 3. Although case with  $\sigma_z = 0$  is faster in the first iterations until  $k = 200$  approximately, then grey line with higher  $\sigma_z$  exhibits a better performance, with a lower noise transmission.

The simulation shows that the method is capable of updating the table in a limited number of iterations and for different cases including some perturbation checking the convergence and robustness. Final calibration of the method depends on the reliability of  $z_k$ , knowing that lower values of  $\sigma_z^2/\sigma_\theta^2$  lead in general to faster corrections but with higher noise transmission, compromising the stability and robustness of the updating.

## 5 Conclusion

The current work develops a novel learning algorithm for look-up tables by means of using membership functions pre-calculated off-line on the basis on a Kalman filter with constant noise distribution. The method shows good agreement for different cases, including ones with some perturbances in the information source used for the updating.

The on-line adaption of look-up tables is useful for estimating engine variables valid for control or diagnosis purposes, without forgetting that pre-calibration work and both engine and sensor variations during performance can be coped. The simple structure of the proposed method allows it to be implemented in commercial ECUs.

## References

- [1] RE Kalman. A New Approach to Linear Filtering and Prediction Problems *Journal of Basic Engineering*, 82:35–45, 1960.
- [2] E Höckerdal, E Frisk and L Eriksson. EKF-based adaptation of look-up tables with an air mass-flow sensor application *Control Engineering Practice*, 19:442–453, 2011.
- [3] G Wu. A Table Update Method for Adaptive Knock Control *SAE paper*, 2006-01-0607, 2006.
- [4] JC Peyton Jones and KR Muske. Identification and adaptation of linear look-up table parameters using an efficient recursive least-squares technique *ISA Transactions*, 48:476–483, 2009.
- [5] C Guardiola, B Pla, D Blanco-Rodriguez, A Mazer and O Hayat. Bias compensation of  $\lambda$  estimator for internal combustion engines *International Journal of Engine Research*, submitted 2011.

# Geometric integrators for solving optimal control problems appearing for quadrotor UAV

P. Bader\*, S. Blanes, J. C. Cortés and E. Ponsoda

Instituto de Matemática Multidisciplinar.

Building 8G, Universitat Politècnica de València.

46022 Valencia, Spain.

November 30, 2011

## 1 Introduction

The optimal control of unmanned air vehicles (UAV) has attracted a great attention in recent years [5]. Helicopters are classified as vertical take-off and landing (VTOL) aircraft and are among the most complex flying objects because their flight dynamics is nonlinear and their variables are strongly coupled.

In this work, we address the optimal control of a VTOL quadrotor based on the model presented in [4, 5, 10], that is, a vehicle with four rigid propellers, whose rotational speeds are independent, placed around a main body. Linear techniques to control the system have been frequently used. However, more realistic problems have time-varying parameters [11], require a time dependent state reference [7] or involve nonlinear equations [6, 10].

Linear quadratic (LQ) optimal controllers are widely used, in particular for the control of small aircrafts [10], where they have shown to produce better results than other standard methods, like proportional integral derivative methods (PID) [4]. The techniques presented here, however, are valid for the

---

\*e-mail: phiba@imm.upv.es

general optimal LQ control problem

$$\min_{u \in L^2} \int_0^{t_f} (X^T(t)Q(t)X(t) + u^T(t)R(t)u(t)) dt, \quad (1)$$

subject to

$$X'(t) = A(t)X(t) + B(t)u(t), \quad X(0) = X_0, \quad (2)$$

where  $X'(t)$  is the time-derivative of the state vector  $X(t) \in \mathbb{R}^n$ ,  $u(t) \in \mathbb{R}^m$  is the control,  $R(t) \in \mathbb{R}^{m \times m}$  is symmetric and non-negative,  $Q(t) \in \mathbb{R}^{n \times n}$  is symmetric and positive definite,  $A \in \mathbb{R}^{n \times n}$ ,  $B \in \mathbb{R}^{n \times m}$ , and  $M^T$  denotes the transpose of a matrix  $M$ .

The optimal control problem (1)-(2) is solved by the linear feedback controller  $u = -R^{-1}B^T P X$ , with  $P(t)$  verifying the matrix Riccati differential equation (RDE)

$$P'(t) = -P(t)A(t) - A^T(t)P(t) + P(t)B(t)R^{-1}(t)B^T(t)P(t) - Q(t), \quad (3)$$

under the final condition  $P(t_f) = 0$ . It can be shown that the solution  $P(t)$  is a symmetric and nonnegative matrix [1]. To compute the optimal control  $u(t)$ , we solve for  $P(t)$  and plugging the control law into (2) yields a linear equation for the state vector

$$X'(t) = (A(t) - B(t)R^{-1}(t)B^T(t)P(t)) X(t), \quad X(0) = X_0. \quad (4)$$

## 2 Nonlinear dynamics

The dynamics of a quadrotor, as well as most UAV, shows nonlinear dynamics [5] and thus, changing in (1),  $Q(t)$  and  $R(t)$  by  $Q(t, X(t))$  and  $R(t, X(t))$  respectively, the corresponding state equation can be written in a nonunique way as

$$X'(t) = A(t, X)X(t) + B(t, X, u)u(t), \quad X(0) = X_0. \quad (5)$$

Usually,  $B$  is assumed to be independent of  $u$ . There are various strategies to control the nonlinear motion of the vehicle.

The *quasilinearization* method takes the exact equation and uses an approximate control law as if the problem were linear, whereas iterative linearization procedures approximate the equation and use the optimal control

law for the approximated equation. The formulation (5) is the basic ingredient for the State Dependent Riccati Equation (SDRE) control technique [6]. Then, formal similarity to the linear problem (1)-(2), leads to

$$X' = (A(t, X) - B(t, X)R(t, X)^{-1}B(t, X)^T P(t, X)) X, \quad X(0) = X_0. \quad (6)$$

The usual approach is to start from  $X(0) = X_0$ , and then to advance step by step in time solving an algebraic Riccati equation ((3) with  $P' = 0$ ) and applying some numerical methods. Alternatively, we can linearize (6) by iterating

$$X^{m+1} = (A(t, X^n) - B(t, X^n)R(t, X^n)^{-1}B(t, X^n)^T P(t, X^n)) X^{n+1}, \quad (7)$$

where starting from a guess solution,  $X^0(t)$ , one obtains iteratively, by solving linear systems of non autonomous ordinary differential equations, a sequence of solutions,  $X^1(t), X^2(t), \dots, X^n(t)$  to be stopped once consecutive solutions differ by less than a given tolerance. This method is called *waveform relaxation*.

Finally, similarly to [9], we can use a third method, called *Taylor-type linearization*, based in a Taylor-expand of the vector field in (5) around an approximate solution  $X^n(t)$  and use optimal LQ controls for the approximated equation.

### 3 Lie group methods

The matrix RDE (3) and the initial value problem (IVP) (4) appear frequently in linear optimal control problems and have been extensively studied in the literature, see [1]. In this work, Lie group methods (see [3] and references therein) are proposed in order to solve these problems. They are geometric integrators which have shown in many cases a high performance for the numerical integration of linear differential equations because they preserve some of the qualitative properties of the exact solution. Regulator problems and their solutions have an important qualitative structure, see [2], which is not preserved by most classical numerical methods, and this can seriously affect the numerical solution.

For all presented methods, after substitution of a control law, we need to solve an IVP,  $X'(t) = D(t, P(t))X(t) + C(t)$ ,  $X(0) = X_0$ . Here,  $P(t)$  is a matrix function satisfying (3) with final condition  $P(t_f) = 0$ . If  $C \neq 0$ , the

linear non homogeneous equation can be formulated as a homogeneous one in the following way

$$\begin{bmatrix} X(t) \\ 1 \end{bmatrix}' = \begin{bmatrix} D(t, P(t)) & C(t) \\ 0_n^T & 0 \end{bmatrix} \begin{bmatrix} X(t) \\ 1 \end{bmatrix}, \quad \begin{bmatrix} X(0) \\ 1 \end{bmatrix} = \begin{bmatrix} X_0 \\ 1 \end{bmatrix},$$

where  $0_n \in \mathbb{R}^n$ . The RDE can also be written as a linear problem

$$\begin{bmatrix} V(t) \\ W(t) \end{bmatrix}' = \begin{bmatrix} -A(t)^T & -Q(t) \\ -B(t)R^{-1}(t)B^T(t) & A(t) \end{bmatrix} \begin{bmatrix} V(t) \\ W(t) \end{bmatrix}, \quad \begin{bmatrix} V(0) \\ W(0) \end{bmatrix} = \begin{bmatrix} 0 \\ I \end{bmatrix}, \quad (8)$$

where the solution  $P(t)$  of problem (3) is given by

$$P(t) = V(t)W(t)^{-1}; \quad P(t), V(t), W(t) \in \mathbb{R}^{n \times n}, \quad (9)$$

in the region where  $W(t)$  is invertible (see, for instance, [8] and references therein). Hence, both the IVP for  $X(t)$  and the RDE for  $P(t)$  reduce to a matrix linear homogeneous equation to which we focus our attention.

Magnus integrators, a special class of exponential integrators as well as Lie group integrators, when used to numerically solve, e.g. the eq. (8), can be interpreted as exactly solving a slightly perturbed matrix RDE with a similar structure (i.e. replacing the matrices  $Q, R$  by perturbed ones  $\tilde{Q}, \tilde{R}$  close to  $Q, R$  which are also symmetric, non negative, etc.) and then, the numerical solution for  $P(t)$  will also be a symmetric and nonnegative matrix.

Let us present an explicit symmetric second order Magnus integrator to solve a general linear equation

$$y'(t) = S(t) y(t), \quad y(t_0) = y_0; \quad (10)$$

with  $y(t) \in \mathbb{R}^p$ . Let us denote by  $\Phi(t, t_0) \in \mathbb{R}^{p \times p}$  the fundamental solution, such that  $y(t) = \Phi(t, t_0)y(t_0)$ . Then,  $\exp\left(\int_t^{t+h} S(t)dt\right) = \Phi(t+h, t) + \mathcal{O}(h^3)$ , corresponds to the first order approximation (second order in the time step,  $h$ ) for most exponential methods like e.g. the Magnus, Fer or Wilcox expansions, see [3] details. Here, it suffices to approximate the integral by a second order symmetric rule. From the computational point of view, we found it useful for the numerical algorithm to consider the trapezoidal rule.

Firstly, let us consider the RDE (3) that corresponds to (10) with the data (8). Let us consider an equidistant time grid  $t_n = t_0 + nh$ ,  $0 \leq n \leq N$ ,

with constant time step  $h = (t_f - t_0)/N$ . By (9), applying the trapezoidal rule in order to approximate the integral, and taking into account that this equation has to be solved backward, we obtain

$$\begin{bmatrix} V_n \\ W_n \end{bmatrix} = \exp\left(-\frac{h}{2}[S(t_n) + S(t_{n+1})]\right) \begin{bmatrix} V_{n+1} \\ W_{n+1} \end{bmatrix} \Rightarrow P_n = V_n W_n^{-1}, \quad (11)$$

the matrix function values  $A(t_n)$ ,  $B(t_n)$ ,  $Q(t_n)$ ,  $R(t_n)$  at the same grid points as the time symmetric second order approximation  $P(t_n)$ . A superscript  $P^k$  indicates the value in the  $k$ -th iteration. The final step is the integration (forward in time) of the dynamics depending on the linearization, (6), (7) or Taylor linearization, the last of which are inhomogeneous. Note that for the Magnus method, e.g. (11), one has to compute the action of an exponential on a matrix or a vector. In our problem, the matrices  $A$  and  $B$  are sparse and  $Q, R$  are diagonal. Thus, we can take diagonal Padé matrix approximations to preserve the Lie group structure. These rational approximations can be easily computed using a fixed point iteration algorithm.

## 4 Numerical simulations and conclusions

An analysis of the dynamics of the quadrotor shows that the control of the attitude can be separated from the translation of the UAV [10] and we focus our attention on the stabilization of the attitude, neglecting the gyroscopic effect. The state vector is given by  $X(t) = \left(\phi(t), \dot{\phi}(t), \theta(t), \dot{\theta}(t), \psi(t), \dot{\psi}(t)\right)^T \in \mathbb{R}^6$ , where  $\phi(t)$ ,  $\theta(t)$  and  $\psi(t)$  are the rolling, pitching and yawing angles respectively, and the input vector  $u \in \mathbb{R}^3$  is formed by linear combinations of the thrust of each propeller.

The system designer can choose the weight matrices to tune the behavior of the control according to the requirements,  $R(t)$  is used to suppress certain movements and  $Q(t)$  limits the use of the control inputs. Usually, these matrices are chosen constant, nonnegative definite, and often even diagonal, see [5, 7]. For the numerical experiments, we consider the problem (5) with the following values taken from [10]

$$\begin{aligned} a_{1,2} = a_{3,4} = a_{5,6} = 1, & \quad a_{2,4} = \lambda\alpha_1 I_1 \dot{\psi}, & \quad a_{2,6} = \lambda(1 - \alpha_1) I_1 \dot{\theta} \\ a_{4,2} = \lambda\alpha_2 I_2 \dot{\psi}, & \quad a_{4,6} = \lambda(1 - \alpha_2) I_2 \dot{\phi}, & \quad a_{6,2} = \lambda\alpha_3 I_3 \dot{\theta}, \\ a_{6,4} = \lambda(1 - \alpha_3) I_3 \dot{\phi}, & \quad b_{2,1} = L/I_x, \quad b_{4,2} = L/I_y, \quad b_{6,3} = 1/I_z \end{aligned} \quad (12)$$

where  $\alpha_i$  reflects the non-uniqueness in the SDRE formulation and is set to 1 in the following,  $L$  is the length of the arms connecting the propellers with the center,  $I_1 = (I_y - I_z)/I_x$ ,  $I_2 = (I_z - I_x)/I_y$ ,  $I_3 = (I_x - I_y)/I_z$ . Here,  $m_{i,j}$  denotes the element located at  $i$ -th row and  $j$ -th column of the matrix  $M$ . Other entries of  $A(t) \in \mathbb{R}^{6 \times 6}$  and  $B(t) \in \mathbb{R}^{6 \times 3}$  not indicated in (12) are null elements. The numerical values are extracted from [4] and are given in the SI units,  $I_x = 0.0075$ ,  $I_y = 0.0075$ ,  $I_z = 0.0130$ ,  $L = 0.23$ . The weight matrices are fixed at

$$Q = 0.01 \cdot \text{diag}\{1, 0.1, 1, 0.1, 1, 0.1\} \in \mathbb{R}^{6 \times 6}, \quad R = \text{diag}\{1, 0.1, 1\} \in \mathbb{R}^{3 \times 3}.$$

We set the time frame to  $t_f = 10$  seconds, with a stepsize of  $h = 0.125s$  and initial state  $X_0 = (70^\circ, 10, 70^\circ, 20, -130^\circ, -1)^T$ , that corresponds to a disadvantageous orientation and high rotational velocities that is sought to be stabilized at  $0 \in \mathbb{R}^6$ .

We have implemented a variety of methods to test against the Magnus integrators presented in section 3. As initial condition, we have taken  $X^0(t) = (1 - t/t_f)X_0$  and iterated until  $\|X^n - X^{n-1}\|_2 < 10^{-3}$ . Some experimental results are given in Table 1, where we can see that the Magnus method approximates better the optimal control.

	Type	$X(t)$	$P(t)$	$V(t)$	Cost	It.
S1)	SDRE	Euler	are	N/A	0.1114	
S2)		Impl. Euler (IE)	are	N/A	0.1021	
		Optimal	$\Rightarrow$		0.0977	
W1)	WAVE	Euler	Euler	N/A	0.1071	3
W2)		IE	IE	N/A	0.1036	3
W3)		Magnus	Magnus	N/A	0.0926	3
		Optimal	$\Rightarrow$	N/A	0.0888	
T1)	TAYLOR	Euler	Euler	Euler	N/A	Inf
T2)		IE	IE	IE	0.0789	12
T3)		Magnus	Magnus	Magnus	0.0707	12
		Optimal	$\Rightarrow$		0.0707	

Table 1: Comparison of numerical methods, Type indicates the linearization procedure (section 2), and It. denotes the number of iterations necessary until convergence. The cost is the value of the integral (1) evaluated at the grid points.

Figure 1 shows the dynamics of the quadrotor subject to the obtained controls for the schemes S2, W3, T3 given by table 1. We can appreciate how the Magnus methods maximizes the use of the controls to reach an overall minimum of the cost functional.

From the numerical experiments we conclude that Lie group methods, like Magnus integrators, are very useful tools for solving optimal control problems of UAV. The results shown for a quadrotor can be easily extend to other helicopters. In addition, for more involved trajectories the structure of the equations will play a more important role and Lie group methods can provide efficient numerical algorithms.

## Acknowledgements

The authors acknowledge the support of the FPU, fellowship AP2009-1892, and the Ministerio de Ciencia e Innovación (Spain) under the projects MTM2010-18246-C03 and MTM2009-08587.

## References

- [1] H. Abou-Kandil, G. Freiling, V. Ionescu, V. & G. Jank, Matrix Riccati equations in control and systems theory. Basel, Birkhäuser Verlag, 2003.
- [2] B. D. O. Anderson & J. B. Moore, *Optimal control. Linear quadratic methods*. New York, Dover Publications, 2007.
- [3] S. Blanes, F. Casas, J. A. Oteo, & J. Ros, The Magnus expansion and some of its applications. *Physics Reports*, 470: pp. 151–238, 2009.
- [4] S. Bouabdallah, A. Noth & R. Siegwart. PID vs LQ control techniques applied to an indoor micro quadrotor, *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS 2004)*, 3: 2451–2456, 2004.
- [5] P. Castillo, R. Lozano & A. E. Dzul, Modelling and control of mini-flying machines. (Adv. Industrial Control Series). London, Springer, 2005.
- [6] T. Çimen. State-dependent Riccati equation (SDRE) control: A survey, *Proc. of the 17th IFAC World Congress(IFAC'08)*, pp. 3761–3775, 2008.

- [7] I. D. Cowling, J. F. Whidborne, & A. K. Cooke. Optimal trajectory planning and LQR control for a quadrotor UAV, *Proc. UKACC Int. Conf. Control 2006 (ICC2006)*. Glasgow, 2006.
- [8] L. Jódar & E. Ponsoda, Non-autonomous Riccati-type matrix differential equations: existence interval, construction of continuous numerical solutions and error bounds. *IMA J. Num. Anal.*, 15: 61-74, 1995.
- [9] E. Ponsoda, S. Blanes, & P. Bader. New efficient numerical methods to describe the heat transfer in a solid medium, *Math. Comput. Mod.* 54: 1858-1862, 2011
- [10] H. Voos. Nonlinear state-dependent Riccati equation control of a quadrotor UAV, *Proc. Int. Conf. Control Appl. pp. 2547-2552*. Munich, 2006.
- [11] R. Zhang, Q. Quan, & K. Y. Cai. Attitude control of a quadrotor aircraft subject to a class of time-varying disturbance. *IET Control Theory Appl.*, 5: 1140-1146, 2011.

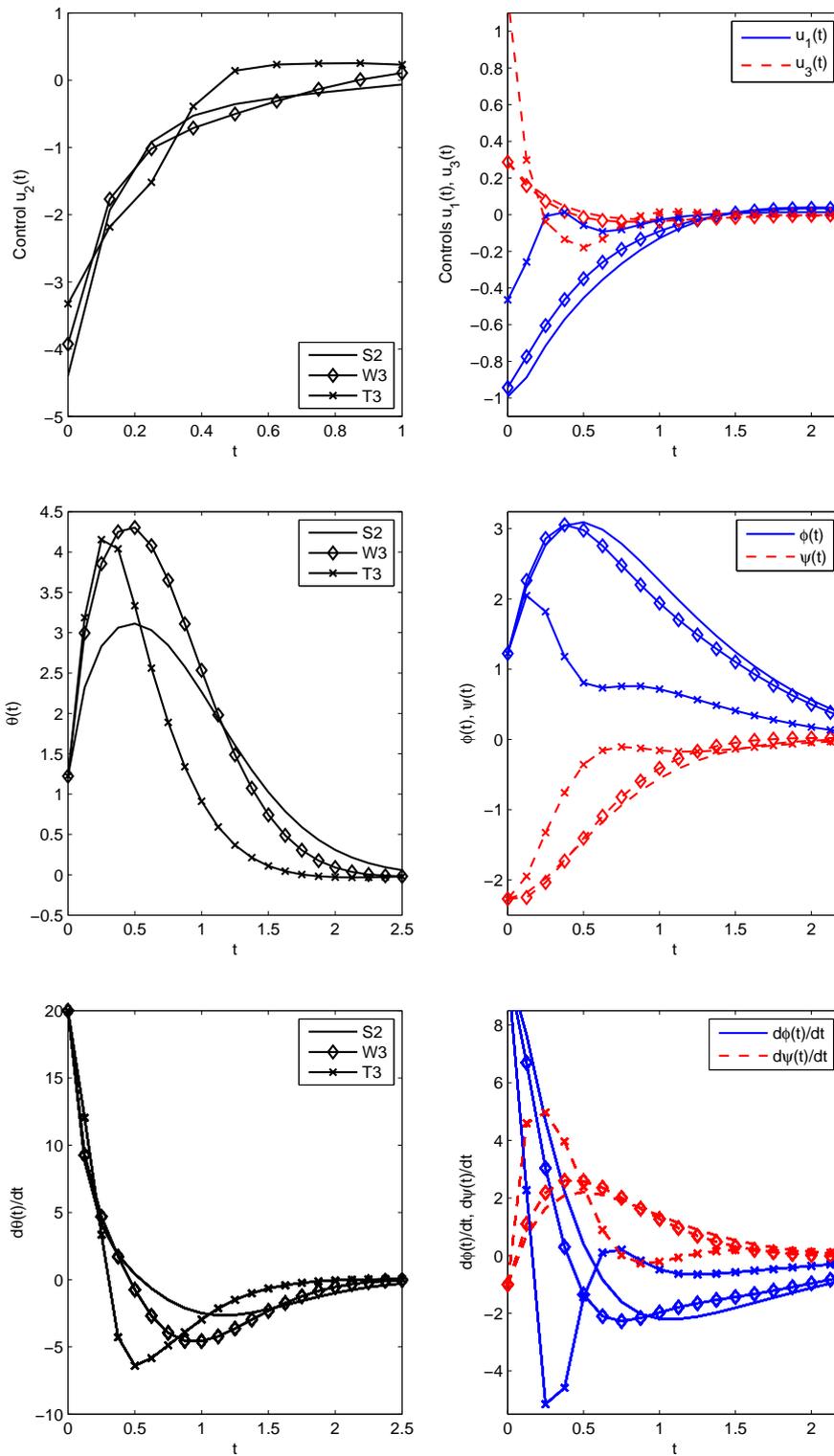


Figure 1: Evolution of the control vector (top row), orientation of the quadrotor (middle) and angular velocities (bottom). The left column shows the control that has been less penalized  $u_2$  and the coordinates  $\theta(t)$  and  $\theta'(t)$ . All curves are given for all methods S2 (line), W3 (diamond) and T3 (cross).

# **Title: “A standardization model by equivalent patient to control outpatient pharmaceutical expenditure”**

**Maria Caballer-Tarazona<sup>a</sup>, Laia Buigues-Pastor<sup>b</sup>, Inmaculada Saurí-Ferrer<sup>b</sup>, Ruth Usó-Talamantes<sup>b</sup>, Jose Luís Trillo-Mata<sup>c</sup>.**

**a Corresponding autor.** Maria Caballer-Tarazona. Departamento de Economía Aplicada. Facultad de Economía. University of Valencia. Avinguda dels Tarongers s/n. c.p:46022. Valencia. Spain. Phone: +34 961625388

**b Oficina de Farmacoeconomía. Dirección General de Farmacia y Productos Sanitarios. Valencian Agency of Health. C/Micer Mascó 31. c.p:46010. Valencia. Spain. Phone: +34 96 386 92 18**

**c Área de Farmacia y Productos Sanitarios Managing Director. Dirección General de Farmacia y Productos Sanitarios. Valencian Agency of Health. Comunitat Valenciana. Spain Phone: +34 963862828.**

## **INTRODUCTION**

The outpatient pharmaceutical expenditure in the Comunitat Valenciana represents about 20,7% of the total health expenditure. With the obvious premise that public resources are limited and must be optimized, it is necessary to design rational tools which ease policymakers' decisions regarded to the control of pharmaceutical expenditure.

Some of these tools are represented by indicators aimed to quantify pharmaceutical expenditures. Indicators are very useful for assess and monitor health performance by public administration managers. As mentioned before, in the Comunitat Valenciana and in developed countries in general, pharmaceutical expenditure represents a considerable proportion of the total public budget on health. Given this, it is necessary to quantify and compare pharmaceutical expenditure in order to stimulate more efficient prescribing practices. We cannot control those dimension that we cannot quantify and, therefore, it cannot be help in policy makers decisions.

Since 2010, in the Comunitat Valenciana it has been used an indicator aimed to evaluate the performance of each health department regarding pharmaceutical expenditure. This indicator showed the theoretical outpatient pharmaceutical consumption level that patients must had, considering their co-payment status (differentiating between outpatients with free pharmacy and 40% of co-payment). This was achieved across the standardization of the population, based on the co-payment variable.

Considering that pharmaceutical expenditure depends on socio-demographic characteristics and population morbidity but also on the prescribing profile of physicians (in terms of quantity and price of the prescription), an indicator which aims to properly show the performance regarding to pharmaceutical expenditure must include individual characteristics of assigned patients. Therefore, in this paper we propose a model of population standardization in which it is considered the co-payment status but also other socio-demographic variables such as gender, age and international coverage in order to weigh patients in a more adjusted way.

Under this context, the aim of this paper is to design a tool capable to assess and monitor pharmaceutical expenditure performance in each level of organization (either microeconomic or macroeconomic) in order to establish rational criteria for the control of drug expenditure, to stimulate better prescribing practices and, therefore, encouraging the efficient prescription of drugs.

## MATERIALS AND METHODS

For this study three types of data sources have been used: COFs (Spanish initials for Official Colleges of Pharmacy) dispensation database, SIP (Spanish initials for Population Information System) database and CRC (Catalogue of Corporate Resources) database.

## METHODS

The explanation of methods applied in the analysis is divided in two parts: description of the improved indicator and methods concerning the population standardization.

### *The indicator*

As mentioned before, since 2010 the regional health services from Comunitat Valenciana are using an indicator aimed to evaluate the performance of each health department regarding pharmaceutical expenditure. This indicator has been called *Outpatient Expenditure per standardized patient* and shows the outpatient pharmaceutical consumption level per patient, considering their co-payment status, both for a macro organization level (department) as for a micro organization level (health areas, health centres and physicians individually). Thus, the population standardization is based on the co-payment variable (as shown in Expression 1) which allows us to obtain two outpatient categories: with 0% of co-payment and with 40% of co-payment.

$$Ep = \frac{\sum_{i=1}^2 X_i}{\sum_{i=1}^2 Pop_i^{std}} = \frac{X_1}{Pop_1^{std}} + \frac{X_2}{Pop_2^{std}} \quad (1)$$

Where:

$i$	Co-payment status.
$Ep$	Pharmaceutical expenditure per standardized outpatient (previous model)
$X_i$	Total pharmaceutical expenditure among outpatients with co-payment status $i$
$Pop_i^{std}$	Total standardized population with co-payment status $i$

The improvement of the indicator that we propose above requires the inclusion of additional socio-demographic variables in order to obtain more accurate results. Our proposal is a model of population standardization in which it is considered the co-payment status but also variables such as gender, age (20 stratum of 5 years each) and international coverage (with or without international coverage). We have considered as people with international coverage, those registered in the population list that also have a European health card, those registered and also adhered to international agreements or those born abroad without economic resources but also registered in the population list. People without international coverage are those only registered in the population list.

Therefore, the new formula showed in the Expression 2 considers 160 different outpatient categories (2 co-payment status x 2 gender status x 20 age stratum x 2 international coverage status).

$$En = \frac{\sum_{i=1}^{160} X_j}{\sum_{i=1}^{160} Pop_j^{std}} = \frac{X_1}{Pop_1^{std}} + \frac{X_2}{Pop_2^{std}} + \frac{X_3}{Pop_3^{std}} + \dots + \frac{X_{160}}{Pop_{160}^{std}} \quad (2)$$

Where:

$j$	Outpatient category
$En$	Pharmaceutical expenditure per standardized outpatient with category $j$ (new model)
$X_j$	Total pharmaceutical expenditure among outpatients with category $j$
$Pop_j^{std}$	Total standardized population with category $j$

### The population standardization

The standardization coefficient for each outpatient category is obtained as the Expression 3:

$$C_j = \frac{Y_j}{S} = \frac{X_j / Pop_j}{\sum_{i=1}^{160} X_j / \sum_{i=1}^{160} Pop_j} \quad (3)$$

Where:

$j$	Outpatient category
$C_j$	Standardization coefficient among outpatients with category $j$
$Y_j$	Expenditure per outpatient for category $j$
$S$	Average expenditure per outpatient (standard)
$X_j$	Total pharmaceutical expenditure among outpatients with category $j$
$Pop_j$	Total population with category $j$

The standardization coefficient shows how much the pharmaceutical expenditure for each category is over or under the average expenditure per outpatient (the standard). The rate represents the weight that each outpatient category has, it is: the number of standard outpatients that each category represents.

## CONCLUSIONS

Taking into account differences between health departments' population structure is necessary in order to be able to compare their pharmaceutical consumption trends. In this paper we have shown how important is the consideration of socio-demographic variables. Our results demonstrate that there is a clear evidence of the limitations that the old standardization model (based on the co-payment status as a unique variable) implied for any analysis. The new indicator (which standardizes taking into account the co-payment status but also gender, age and international coverage status) gives more accurate results which allow us to offer a better pharmaceutical expenditure analysis.

The proposed tool makes possible the analysis of outpatient drug expenditure patterns for the whole valencian region, but also the comparison of departments and the evaluation of their efficiency and their scope of the pharmaceutical expenditure aim. The indicator can also be used in a micro level for comparing physicians, health centres or health zones from any department, which makes possible the establishment of mechanisms to stimulate physicians to reduce prescription costs.

Even when the new standardization proposal improves the old model, it is important to notice that variables such as gender or age are not significant when they are applied with more powerful variables like morbidity.

In this line, risk adjusting methodologies like Clinical Risk Groups (CRGs) or Adjusted Clinical Groups (ACGs) that assign each individual to a single mutually exclusive risk group based on measure the different cases according to the diagnostic try to cluster the cases depending on their homogeneous level of resource consumption and at the same time have clinical significance.

However, the indicator developed in this paper can be useful when morbidity data is not available, as it is the case of the majority regions in Spain.

# Differences between adult health and early adult / late adolescent health: an empirical analysis

J. David Cabedo\* , Aurora Esteve†, and José Miguel Tirado‡

(\*) Departamento de Finanzas y Contabilidad  
Universidad Jaume I, Castellon (Spain)

(†) U.P de Enfermería,  
Universidad Jaume I, Castellon (Spain)

(‡) Departamento de Finanzas y Contabilidad  
Universidad Jaume I, Castellon (Spain)

November 30, 2011

## 1 Introduction

We use the terms adolescents and young people to refer to the population group that shares the gradual transition from childhood to adulthood; the process of constructing identities and becoming part of broader social life occurs in this period. A set of studies has appeared in recent years that analyse the health problems of adolescents and young people.[1] proposes a specific category of development called early adulthood for the period between 18 and 25 years. [3] analyse health trends in adolescents and young people from 10 to 24 years. [4] analyse the way adolescents and young people between the ages of 15 and 24 perceive their health. The study by [5] analyses the relationship between the type and severity of mental disorders and the search for help in young people aged between 16 to 24 years in comparison to people

---

\*e-mail: cabedo@cofin.uji.es

aged from 25 to 44 and 45 to 85. In a systematic review of the literature, [2] determine the relationship between the social, cognitive, behavioural, emotional and moral competences of young people up to the age of 20 and their sexual and reproductive health. [6] study the relationship between family structure, involvement of parents who do not live with the family and eating habits in a sample of young people between the ages of 11 and 21. In this line of studies that jointly analyse the problems of adolescent and young people's health, the aim of this work is to determine whether particularly significant health problems in adolescence show any differences between late adolescence/youth and adulthood.

## **2 Data and methodology**

Many organisations are concerned with adolescent health (at world, national, regional and local level). These bodies and the literature identify a set of health problems related to risk behaviours which affect adolescents in particular and endanger not only their present but also their future health. These problems can be grouped into the following categories: sexual and reproductive health, substance use (tobacco, alcohol and drugs), mental health, diet, nutrition and physical activity, violence and accidents.

For our study, we use the data from the Spanish national survey into adult health for the year 2006 undertaken by the Spanish Ministry of Health. The survey is not specifically directed at detecting the above-mentioned health problems, although it contains a series of questions which are able to provide such information. In particular, the survey provides information on respondents' behaviour in relation to the following aspects: road safety, addictive behaviours, mental health, eating habits and physical exercise. For other items, such as sexual and reproductive health the survey does not contain relevant data. As indicated above, this work aims to determine to what extent the health problems identified as particularly relevant in adolescence persist in the transition to adulthood. We have therefore contrasted differences in these health problems between two population groups: late adolescents / young people (individuals aged between 17 and 25 years referred to henceforth simply as adolescents) and adults (persons over 25 and under 51). We have decided to limit adult age to 51 so that the comparison group for adolescents contains information on adult health problems without including other problems corresponding to later ages in human development.

The differences in categorical variables between both populations have been analysed using Pearson's Chi-squared test to contrast independence between two attributes. In cases where there are two categories the above test is applied with Yates' correction. For the numerical variables we test for equality of measurements (Student t-test) and equality of medians (Wilcoxon range test) of the sample (survey data). In the case of the means, before the above equality test, we used the F-test to test for equality of variances between the two groups (adolescents and adults). Student's t-test is formulated depending on whether the equality of variances is accepted or rejected.

### **3 Empirical analysis**

#### **3.1 Road traffic accidents**

Within the section on road traffic accidents helmet and seat belt use has been analysed for travel in and outside the city and as driver or passenger. In all cases over 90% of adolescents and adults use seat belts. However, statistically significant differences have been found between both groups when they are passengers. In this case, the percentage of individuals who use a seat belt is lower in the case of adolescents. In particular, when travelling as passengers in the city 10% of adolescents do not usually wear a seat belt in contrast to 6.2% of adults. In the case of helmet use the situation is similar: there are differences between both groups when travelling in the city: 11.6% of adolescents do not always use a helmet, in contrast to 6.1% of adults.

#### **3.2 Addictive behaviours**

There have been detected statistically significant differences in the number of smokers, where in proportional terms, there are more smokers in the group of adults. There are also differences between both groups in terms of the number of cigarettes that habitual smokers usually consume: adults smoke more than adolescents which confirms that acquisition of the smoking habit in adolescence continues into adulthood and that addiction generates higher consumption in the long term. On the other hand, we have detected that consumption of types of alcohol is related to age. The main differences centre on the consumption of wine or cava (higher in the adult population) and the consumption of brandy and other liqueurs (higher in the group of ado-

lescents). We have also detected statistically significant differences in the amount of each type of alcohol consumed by the groups analysed. Furthermore, it is striking that in all cases adolescents consume higher amounts than adults. Finally we have analysed the days when each type of alcohol is consumed. Adolescents concentrate consumption mainly at the weekend and each time they consume alcohol, they consume more than adults resulting in very different consumption patterns between the two groups. Adolescents drink greater amounts in a shorter space of time.

### **3.3 Mental Health**

In relation to mental health, first of all we analysed the items related to how the individuals in the groups feel about themselves and those related to satisfaction and stress at work. In the few items where there are statistically significant differences the results for adolescents can be classified as more favourable. Another aspect analysed in mental health is how individuals feel in relation to their environment. For most of the items analysed, no differences between groups have been detected or if there are any differences, the problem is greater for adults than adolescents. Only for two items statistically significant differences have been detected and the results show that the problem is greater for adolescents than for adults. In particular, adults speak of domestic problems in the home or take important decisions jointly with the other members of the household more frequently than adolescents.

### **3.4 Diet and physical exercise**

The different consumption habits of adults and adolescents include in particular the finding that adolescent consumption of fruit, vegetables and pulses is lower and their consumption of prepared meats, sweets and sugary soft drinks is higher. On the other hand, our results show that more adults follow diets. However, when the reasons are analysed it is striking that most adolescents who follow a diet do so to lose weight. Finally, the only variable in the national health survey concerning physical health does not show less healthy behaviour in adolescents.

## 4 Conclusions

As regards risk behaviours related to road traffic accidents, it has been detected that when in the city or as a passenger, helmet and seat belt use is more frequent in the adult population. In relation to addictive behaviours, our findings confirm that the acquisition of the smoking habit in adolescence continues through to adulthood and that addiction generates greater consumption in the long term. As regards alcohol, there have been found different consumption patterns in the groups analysed: adolescents drink greater amounts in a shorter space of time. This adolescent risk behaviour detected in various studies, persists in the transition to adulthood and therefore indicates that in the future, it could become a problem in adulthood. With regard to eating habits, the results of the statistical tests show in general that consumption of healthy foods is lower in adolescents than in adults. In addition, analysis of the motives for following a diet shows that, in comparison to adults, the reason mostly given by adolescents concerns body image; the desire to become thin. In short, this work highlights that most unhealthy habits are detected in adolescence and persist in the period of transition towards adulthood. It is therefore necessary to reinforce educational action directed at adolescents to reduce risk behaviour which may have a negative impact on their health and quality of life when they reach adulthood.

## References

- [1] Arnett, J.J. Competence as a predictor of sexual and reproductive health outcomes for youth: a systematic review *Journal of Adolescent Health*, 55(5): 469–480, 2000.
- [2] House, L.D. Bates, J. Markham, C.M. and Lesesne, C. A theory of development from the late teens through the twenties *American Psychologist*, 46: S7-S22, 2010.
- [3] Mulye, T.P.; Park, M.J. Nelson C.D.; Adams, S.H.; Irwin and C.E.; Brindis, C.D. Trends in adolescent and young adult health in the United States. *Journal of Adolescent Health*, 45: 8-24, 2009.
- [4] Ott, M.A. Rosenberger, J.G. McBride, K.R. and Woodcox, S.G. How do adolescents view health? Implications for state health policy *Journal of Adolescent Health*, 48: 398-403, 2011.

- [5] Reavley, N.J., Cvetkovski, S., Jorm, A.F. and Lubman, D.I. Help-seeking for substance use, anxiety and affective disorders among young people: results from the 2007 Australian National Survey of Mental Health and Wellbeing *Australian and New Zealand Journal of Psychiatry*,44: 729-735, 2010.
- [6] Stewart, S.D. Menning, C.L Family structure, non resident father involvement, and adolescent eating patterns *Journal of Adolescent Health*,45: 193-201, 2009.

# Measuring performance of social and non-profit Microfinance Institutions (MFIS)

C. Bartual Sanfeliu<sup>(†)</sup>,  
R. Cervelló Royo<sup>(†)</sup>\*, I. Moya Clemente<sup>(†)</sup>

(†) Departamento de Economía y Ciencias Sociales,  
Facultad de Administración y Dirección de Empresas,  
Universidad Politécnica de Valencia,  
Camino de Vera s/n, 46022, Valencia,

November 30, 2011

## 1 Introduction

Most studies dealing with business performance are focused on the traditional trading bank sector: Private Capital banks [1, 2, 3] among others, and Savings banks [4]. Institutions blamed for being responsible of the credit restriction to the less-privileged classes. In order to meet this need, new financial intermediaries have arisen, the Microfinance Institutions (MFIs), which provide small loans (microcredits) to poor people who have promising and feasible investment ideas that can lead to profitable ventures. However, these special financial institutions are also interested in financial matters like profitability, returns and efficiency. The evolution and expansion of the Microfinance Industry have led to consider all these aspects; so a set of standardized performance indicators have been introduced. Since then, there is a lot of literature and information which deals with aspects like sustainability/profitability, asset/liability management and/or portfolio quality [6],

---

\*e-mail: rocerro@esp.upv.es

whereas there is little literature on efficiency/productivity of these institutions [7, 8].

The main aim of the multicriteria business global performance is to combine all these institution's performance criteria into a coherent whole in order to synthesize the information contained in a series of single-criterion business performance. The definition of the weights of the variables used in the multicriteria performance ranking represents the most difficult task, trying to minimize as far as possible the subjectivity of the person who decides the weights. To this end, the present study proposes the multicriteria goal programming technique (Goal Programming) [9].

## 2 Methodology and data sources

### 2.1 Methodology

Starting out from the idea of [10] and by applying different versions of the goal programming model, a collective approach is considered (giving greater weight to criteria that show similar performance over the more conflictive criteria), as well as an individual approach (greater weight to the more conflictive criteria). In the first case, the absolute difference between the multicriteria value and the standardized single-criterion value (norm  $L_1$ ) is minimum. In the second, the greatest difference recorded between the multicriteria value and the standardized single-criterion values (norm  $L_\infty$ ) is minimum.

The goal programming model in norm  $L_1$  appears in (1):

$$\begin{aligned}
& \min \sum_{i=1}^n \sum_{j=1}^c (n_{ij} + p_{ij}), \\
& \text{s.a.} \\
& \sum_{j=1}^c w_j v_{ij} + n_{ij} - p_{ij} = v_{ij}, \quad i = 1, \dots, n, \quad j = 1, \dots, c, \\
& \sum_{j=1}^c w_j = 1, \\
& \sum_{j=1}^c w_j v_{ij} = V_i, \quad i = 1, \dots, n, \\
& \sum_{i=1}^n (n_{ij} + p_{ij}) = D_j, \quad j = 1, \dots, c, \\
& \sum_{j=1}^c D_j = Z,
\end{aligned} \tag{1}$$

where:

- $w_j$ : weight to be estimated for the  $j$ -th criterion.
- $n_{ij}(p_{ij})$ : negative (positive) deviation variable. Quantifies the difference by excess (defect) between the value of the  $i$ -th MFI in the  $j$ -th criterion and the multicriteria value obtained by applying the weights  $w_j$ . This is,  $n_{ij} - p_{ij} = v_{ij} - \sum_{j=1}^c w_j v_{ij}$ , with  $n_{ij}, p_{ij} \geq 0$ . The objective function of (1) ensures that only one of the deviation variables can have a value greater than zero:  $n_{ij} \times p_{ij} = 0$ .
- $D_j$ : degree of disagreement between the  $j$ -th criterion and the multicriteria value.
- $Z$ : magnitude of global disagreement.

The model (1) has a total of  $n \times c$  goal constraints. This means that for each criterion  $j$  ( $j = 1, \dots, c$ ) the model implements  $n$  constraints, one for each alternative  $i$  ( $i = 1, \dots, n$ ) and must determine the weight associated with the criterion  $j$ ,  $w_j$ . This is obtained by minimizing the difference in absolute terms between the single-criterion performance of each alternative in the criterion  $j$ ,  $v_{ij}$ , and the multicriteria performance  $V_i$  with  $V_i = \sum_{j=1}^c w_j v_{ij}$ . This value is the ultimate objective of the methodology, since on assigning a single value to each alternative as the total of all single-criterion performances, the ranking of the alternatives is immediately obtained.

The value of the target function provides the degree of non-achievement of the set of goals. Weightings are restricted to sum 1. The last constraints are used to compute the MFIs' multicriteria performance ( $V_i$ ), the degree of disagreement of each single-criterion measurement in relation to the multicriteria value ( $D_j$ ) and the degree of global disagreement ( $Z$ ). In the literature, the model that minimizes the sum of absolute deviations is known as the weighted goal programming model (WGP).

The  $L_\infty$  norm is implemented by the MINMAX goal programming model (2), in which  $D$  represents the maximum deviation between the multicriteria value and the single-criterion values. The rest of the variables keep the same meaning as in (1).

$$\begin{aligned}
 & \min D, \\
 & \text{s.a.} \\
 & \sum_{j=1}^c w_j v_{ij} + n_{ij} - p_{ij} = v_{ij}, \quad i = 1, \dots, n, \quad j = 1, \dots, c, \\
 & \sum_{i=1}^n (n_{ij} + p_{ij}) \leq D, \quad j = 1, \dots, c, \\
 & \sum_{j=1}^c w_j = 1, \\
 & \sum_{j=1}^c w_j v_{ij} = V_i, \quad i = 1, \dots, n, \\
 & \sum_{i=1}^n (n_{ij} + p_{ij}) = D_j, \quad j = 1, \dots, c, \\
 & \sum_{j=1}^c D_j = Z,
 \end{aligned} \tag{2}$$

Criteria weights ( $v_{ij}$ ) were normalized from the original variables ( $u_{ij}$ ), so that  $v_{ij} = (u_{ij} - u_{*j}) / (u_j^* - u_{*j})$ , with  $u_j^* = \max_i u_j^*$  and  $u_{*j} = \min_i u_{ij}$ . Normalization is needed when the original variables are given in different measures (percentage, monetary units, etc.). The solutions provided by models (1) and (2) represent extreme cases in which conflicting strategies are opposed to each other: favoring global consensus (WGP) or favoring the criteria that generate performance rankings with a higher degree of idiosyncrasy (MIN-MAX GP).

An interesting option for a compromise between (1) and (3) is to employ

an extended goal programming model (3) in which the  $\lambda$  parameter provides more balanced solutions. This widens the range of possibilities when deciding which multicriteria value is the most suitable and representative of the individual criteria. Note that if the same solution is obtained as in model (1), while if the solution coincides with model (2).

$$\begin{aligned}
 & \min \lambda \sum_{i=1}^n \sum_{j=1}^c (n_{ij} + p_{ij}) + (1 - \lambda)D, \\
 & \text{s.a.} \\
 & \sum_{j=1}^c w_j v_{ij} + n_{ij} - p_{ij} = v_{ij}, \quad i = 1, \dots, n, \quad j = 1, \dots, c, \\
 & \sum_{i=1}^n (n_{ij} + p_{ij}) \leq D, \quad j = 1, \dots, c, \\
 & \sum_{j=1}^c w_j = 1, \\
 & \sum_{j=1}^c w_j v_{ij} = V_i, \quad i = 1, \dots, n, \\
 & \sum_{i=1}^n (n_{ij} + p_{ij}) = D_j, \quad j = 1, \dots, c, \\
 & \sum_{j=1}^c D_j = Z,
 \end{aligned} \tag{3}$$

## 2.2 Data sources

To select the variables used in the study, those used in the works cited in the first section were considered, as were all the different business performance categories (Institutional Characteristics, Financing Structure, Overall Financial Performance, Expenses, Efficiency and Productivity, Risk and Liquidity).

Considering the special features of the MFIs, the Corielli and Marcelino [11] algorithm was used in order to get the most important or most representative factors for each one of the mentioned categories and as is usual in studies on performance, certain areas were assigned several variables that measure different aspects. All the criteria were directly combined in the performance ranking, assuming the higher the value of any of the criteria, the higher the perception of the performance; except for the following aspects: Financial Expense/Assets (FEA), Personnel Expense/Assets (PEA), Cost per Borrower (CPB) and Personnel Allocation Ratio (PA), which had

a negative effect. The database for the present study was compiled from the MIX Market reports published by Microfinance Information Exchange.

### 3 Results, Discussion and Conclusions

The 6 categories and the 12 criterion obtained are used as a starting point, with the original variables being normalized. On solving (3) for different values of  $\lambda \in [0, 1]$  we obtain 1) the weighting or relative importance of each individual criterion in the overall performance ranking and 2) the multicriteria value which ranks the MFIs according to performance.

For each  $\lambda$  value, we check the weight of each criterion, the deviations between the multicriteria performance and each of the criteria ( $D_j$ ,  $j = 1, \dots, 12$ ), the maximum deviation  $D$  between them, and the global deviation  $Z$  as the sum of all  $D_j$ . As the value of the  $\lambda$  parameter diminishes, the number of variables also diminishes. At the extreme  $\lambda = 0$  only five variables appear with a non-null coefficient: P (Institutional Characteristics), PTA (Financing Structure), ROE (Overall Financial Performance), PA (Efficiency and Productivity) and PEA (Expenses). This means that these are the five variables that differ most from the rest as to the amount of information they contribute to the performance. If we evaluate the  $D_j$  values for each of the variables we see that the biggest of these corresponds to ( $D_9$ ) for most values of parameter  $\lambda$ . This means PA is the variable most in disagreement with the rest of the single-criterion performance measurements. We also considered the weight of each of the categories contained in the analysis, obtained as the sum of the individual weights of each criterion. We check that as the value of  $\lambda$  diminishes, Risk and Liquidity gives part of its weight to Overall Financial Performance and Financing Structure, while Institutional Characteristics remains around 25% or even more for the entire range of  $\lambda$  values analyzed. Although the weight of each criterion, or the set calculated for the category, offers an idea of the relative importance of each measurement in calculating multicriteria performance, a Spearman correlation analysis must be carried out to analyze the correlation between each of the single-criterion measurements and final performance. The variables with the highest correlation coefficient are ROE and PTA. Thus, it can be concluded that ROE and PTA clearly have the highest correlation with multicriteria performance and will no doubt be considered key factors in any decisions that have to be taken in the near future by the MFIs.

## References

- [1] Bonin J, Hasan I and Wachtel P. Bank performance, efficiency and ownership in transition countries *J. Banking Finan.*,29(1):31–53, 2005.
- [2] Pastor J.M., Credit risk and efficiency in the European banking system: A three-stage analysis *Appl. Finan. Econ.*,12:895–911, 2002.
- [3] Tortosa-Ausina E., Exploring efficiency differences over time in the Spanish banking industry *Eur. J. Oper. Res.*,13:643–644, 2002.
- [4] Garcia F, Guijarro F and Moya I. Ranking Spanish savings Banks: A multicriteria approach *Math. Comput. Model.*,52:1058–1065, 2010.
- [5] Ahlin C, Lin J and Maio M. Where does Microfinance flourish? Microfinance institution performance in Macroeconomic context *J. Dev. Econ.*,95(2):105–120, 2011.
- [6] Jansson T, Stauffenberg V, Kenyon N and Barluenga-Badiola MC. Performance Indicators for Microfinance Institutions: Technical Guide. Washington DC, MicroRate and Inter-American Development Bank Ed., 2003.
- [7] Gutiérrez-Nieto B, Serrano-Cinca C and Mar-Molinero C. Microfinance institutions and efficiency *Omega-Int. J. Manage. Sci.*,35:131–142, 2007.
- [8] Garmaise MJ, and Natividad G. Information, the cost of credit and operational efficiency: an empirical study of Microfinance *Rev. Financ. Stud.*,23(6):2560–2590, 2010.
- [9] Charnes A, Cooper WW and Ferguson RO. Optimal estimation of executive compensation by linear programming *Management Science*,1:138–150, 1955.
- [10] Linares P and Romero C. Aggregation of preferences in an environmental economics context: a goal-programming approach *Omega-Int. J. Manage. Sci.*,30(2):89–95, 2002.
- [11] Corielli F and Marcellino M. Factor based index tracking *J. Banking Finan.*,30:2215–2233, 2006.

# Virus Propagation with Randomness

B. Chen-Charpentier<sup>\*</sup> and D. Stanescu<sup>†</sup>

(<sup>\*</sup>) Department of Mathematics,

University of Texas at Arlington, Arlington, TX 76019-0408, USA,

(<sup>†</sup>) Department of Mathematics,

Department of Mathematics, University of Wyoming, Laramie, WY 82071-3036, USA.

November 30, 2011

## 1 Introduction

Here we consider two models: a simple model that includes only uninfected cells, infected cells and free viruses, and a model that also considers CTL dynamics. Cytotoxic T lymphocytes (CTLs) form a branch of the immune system essential in the fight against many viral infections [6].

As is the case in many epidemic models, the transmission or infection coefficients and other parameters in virus models are difficult to determine and have variability due to errors in measurements, differences in population size

Monte Carlo methods [3] are a traditional way to deal with randomness because of their simplicity; however, they usually require many realizations for reasonable accuracy. Other possible choices are the method of moments [5] and polynomial chaos. In this paper we use a non-intrusive version of the method of polynomial chaos (NIPC) to deal with the randomness in the model equations.

---

<sup>\*</sup>e-mail:bmchen@uta.edu

## 2 The Models and Results

Two virus population models are considered for the purpose of this work. The first model was proposed by Nowak and Bangham [4] in their study of the human T cell leukemia virus (HTLV-1) and the human immunodeficiency virus (HIV-1). It is given by a system of three ordinary differential equations.

The virus population has a much faster growth rate than than the populations of uninfected and infected cells, so a common assumption is that the virus population is almost steady, and we get only two equations.

The second model considered here [4], [2] incorporates CTL dynamics. To fight the infection CTL cells kill infected cells and thus try to stop the spread of the virus. Making the same assumption as above that the free virus population grows at much faster rates than the other populations we again can eliminate the equation for the virus.

In the general case one may be interested in a quantity of interest which is usually a functional of the solution of a model that describes certain phenomena under investigation. Let this quantity be denoted by  $X$ , and assume that the model  $X = M(\Lambda)$  used to compute it depends on parameters  $\Lambda$ , both  $X$  and  $\Lambda$  considered to be finite-dimensional. Thus, the quantity of interest also becomes a random variable  $X = M(\Lambda(\omega)) = X(\omega)$ . If the parameter distributions are known, one way to obtain the distribution of the output  $X$  is to use the Monte-Carlo (MC) method [3]. This entails a large numbers of runs of the model to obtain a certain accuracy.

The NIPC method is a reasonable alternative to the MC method. Like the latter, it only requires running the model (together with solving a relatively small-size linear system of equations) to obtain a probabilistic description of the output. The number of model runs may be, however, much smaller than that required for the MC method for the same level of accuracy. It makes use of the Cameron-Martin theorem [1] that guarantees convergence of the chaos expansion. For computational purposes this chaos expansion is truncated

after a finite number of terms and rearranged as  $X(\omega) = \sum_{i=0}^P X_i \Phi_i(\vec{\xi})$  where

$\Phi_i(\vec{\xi})$  are given basis functions. The same expansion can be performed for the parameters of the model. Thus, one value for the random parameter of the model  $\Lambda$  is mapped into one value for the quantity of interest  $X$  as

$X(\omega) = M(\Lambda(\vec{\xi}(\omega)))$ , where  $\Lambda(\vec{\xi}(\omega)) = \sum_{i=0}^P \Lambda_i \Phi_i(\vec{\xi})$  and the coefficients  $\Lambda_i$  are

known.

To obtain the NIPC approximation a sample for the parameter set is first obtained by first sampling the vector  $\vec{\xi}$  and then obtaining  $\Lambda$  from its known expansion. Let such a sample be denoted by  $\vec{\xi}_i = \vec{\xi}(\omega_i)$ , where  $i = 1, 2, \dots, S$  and  $S$  is the total number of samples. Then one equation is obtained for the coefficients of the output variable by running the model and requesting that  $X(\omega) = M(\Lambda(\vec{\xi}_i))$ . This leads to the following system of equations

$$\begin{bmatrix} 1 & \Phi_1(\vec{\xi}_1) & \dots & \Phi_P(\vec{\xi}_1) \\ 1 & \Phi_1(\vec{\xi}_2) & \dots & \Phi_P(\vec{\xi}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \Phi_1(\vec{\xi}_S) & \dots & \Phi_P(\vec{\xi}_S) \end{bmatrix} \begin{bmatrix} X_0 \\ X_1 \\ \vdots \\ \vdots \\ X_P \end{bmatrix} = \begin{bmatrix} M(\Lambda(\vec{\xi}_1)) \\ M(\Lambda(\vec{\xi}_2)) \\ \vdots \\ \vdots \\ M(\Lambda(\vec{\xi}_S)) \end{bmatrix} \quad (1)$$

for the coefficients of the output variable. Note that the number of equations  $S$  is not necessarily related to the number of coefficients  $P + 1$ . In actual simulations we usually construct an over-determined system and solve it using least-squares.

The first set of numerical experiments is performed on the first model. A base case with deterministic values of the parameters is considered.

Randomness is then introduced by asking that some or all of the coefficients be uniformly distributed random variables with mean their deterministic value and spread over an interval of length 0.4 times their mean value. A first case we examined was two random parameters, while the other parameters have their deterministic value. Good accuracy can be noticed in the results for the higher polynomial orders as soon as  $S > P + 1$ .

For a three-dimensional chaos case we also allow a third parameter to be random. One of the advantages of NIPC versus usual polynomial chaos using a Galerkin projection is that in the latter higher-dimensional chaos may lead to lengthy inner products that are difficult to evaluate, whereas in NIPC the chaos dimension makes no difference other than solving a bigger matrix; the setup process of this matrix is the same.

Next we consider the CTL model with the same parameter values as before, which leads to a CTL response. We focus on the ability of NIPC to obtain accurate statistics for larger chaos dimensions. To this purpose we let all the parameters in the model be random and compute the error in the expected value and variance, for which a seven-dimensional chaos is necessary.

The multivariate polynomials in the chaos basis are constructed via tensor products of monomials with order up to  $p = 3$ .

### 3 Conclusions

Parameters in virus models have a large variability; even for the same type of virus the parameters involve have variations, in particular the number of viral particles produced from a single infected cell. The NIPC method is an useful for studying the effect this variability has on the outcome of the mathematical models. Like MC, it does not require the development of any special code except the one for the deterministic model itself, but requires much less sampling, and consequently much less computer time than the latter, for comparable accuracy. Another advantage of this method is that it produces a series expansion of the output process. This is particularly valuable if the output process becomes in turn an input for some other model: realizations of this output process are then readily at hand.

### References

- [1] R. Cameron, and W. Martin, The orthogonal development of nonlinear functionals in series of Fourier-Hermite functionals, *Ann. Math.*, 48: 385–392, 1947.
- [2] R. J. De Boer and A. S. Perelson, Target cell limited and immune control models of HIV infection: a comparison, *J Theor. Biol.*, 190: 201–14, 1998.
- [3] G. S. Fishman, Monte Carlo: Concepts, Algorithms, and Applications, New York, Springer Verlag, 1995.
- [4] M.A. Novak and C.R.M. Bangham, Population Dynamics of Immune Responses to Persistent Viruses, *Science*, 272: 74–79, 1996.
- [5] T. Soong, Random Differential Equations in Science and Engineering, New York, Academic Press, 1973.
- [6] D. Wodarz, Killer Cell Dynamics: Mathematical and Computational Approaches to Immunology, New York, Springer, 2007.

# A dynamical model to study the academic underachievement in Spain: A epidemiological approach.

L. Acedo, J.C. Cortés\*, E. Defez,  
E. Ponsoda, M.J. Rodríguez-Álvarez,  
A. Sánchez-Sánchez, M. Tung and R.J. Villanueva

Instituto Universitario de Matemática Multidisciplinar,  
Building 8G, Access C, 2<sup>nd</sup> floor,  
Universitat Politècnica de València (SPAIN),

November 30, 2011

## 1 Introduction

The analysis of school performance has an important role in achieving quality education and increasing its efficiency. The number of students that abandon their studies not only in the stage of *Bachillerato* but also in other educational levels, and therefore they are not going to be qualified workers is greater than the number of qualified jobs expected over the following years. This will lead to an increase in unemployment among young students, which is also a primer concern of the European Union [1]. In this paper we are interested in studying the academic underachievement in *Bachillerato*. To address this problem, in addition to consider autonomous behavior of students who take decisions about their own priorities, goals, etc., we assume that academic underachievement is a socially transmitted behavior [2, 3]. We treat school failure as a problem that is transmitted through social contact. These social

---

\*e-mail: jccorttes@imm.upv.es

contacts have an influence on the probability of transmission of bad study habits. The main idea behind this approach is that these inappropriate habits may spread from one student to another, more probably between students of the same academic level. These facts lead us to propose an epidemiological-type mathematical model to study the evolution (transmission dynamics) of the academic underachievement in *Bachillerato*. According to [4], there exists a significative difference of academic performance depending on genre. In general, women obtain better academic results than men. This motivates the fact that our model considers genre into its formulation.

## 2 Modelling the academic underachievement

In this paper, we have considered the available data corresponding to the first and second stage of *Bachillerato*, in both, public and private high schools all over Spain during academic years 1999/2000 to 2007/2008. We point out that we study this period because after the academic year 2007/2008 only partial data corresponding to specific regions of Spain are available. For each of these academic years, Table 1 collects the percentage of women/men that pass ( $W_i/M_i$ ) and do not pass ( $\overline{W}_i/\overline{M}_i$ ) for each level  $i = 1, 2$  (that corresponds to the first and second stage of *Bachillerato*, respectively) over the total Spanish *Bachillerato* students. After some algebraic manipulations, these data have been obtained from the official database [5, 6] and they are referred to the month of September, when each academic year ends officially.

Table 1: The available data corresponding to the first and second stage of *Bachillerato*, in both, public and private high schools all over Spain during academic years 1999/2000 to 2007/2008.

Academic year	First stage of Bachillerato (Women/Men)		Second stage of Bachillerato (Women/Men)	
	% Pass ( $W_1/M_1$ )	% Do not Pass ( $\overline{W}_1/\overline{M}_1$ )	% Pass ( $W_2/M_2$ )	% Do not Pass ( $\overline{W}_2/\overline{M}_2$ )
1999-00	19.68/15.24	9.75/9.33	16.21/11.64	9.52/8.63
2000-01	22.65/17.54	9.91/10.12	14.07/10.04	8.24/7.43
2001-02	19.23/14.23	8.61/9.10	17.86/13.06	9.32/8.59
2002-03	18.87/14.19	8.36/8.51	19.14/13.97	8.76/8.20
2003-04	19.93/15.06	7.74/7.88	19.19/13.80	8.44/7.96
2004-05	20.11/15.14	7.65/7.94	18.90/13.92	8.39/7.95
2005-06	20.07/15.39	7.64/7.93	19.14/13.97	8.08/7.78
2006-07	20.06/15.34	7.67/7.87	19.14/14.29	7.98/7.65
2007-08	20.25/15.82	7.57/7.66	19.37/14.61	7.60/7.12

We build our mathematical model based on an epidemiological-type model by considering that academic underachievement is a process that takes place when a female (W) or a male (M) student that initially belongs to the promotable group of a specific level,  $W_i$  or  $M_i$ ,  $i = 1, 2$ , leaves her/his good academic habits due to the negative influence (*contagion*) from other students of the *same* educational level who belong to the group of students with academic underachievement,  $\overline{W}_i$  and  $\overline{M}_i$ . We shall consider that, for each specific academic level under study, the bad academic habits just spread between students of the same course, independently of their genre. Thus, the transitions described can be modeled as follows:

- For a specific *Bachillerato* academic level  $i = 1, 2$ , a student in  $W_i$  (respec.  $M_i$ ) transits to  $\overline{W}_i$  (respec.  $\overline{M}_i$ ) because students in  $\overline{W}_i$  and  $\overline{M}_i$  transmit their negative academic habits at rates  $\beta_i^W$  (respec.  $\beta_i^M$ ). Therefore, this is a nonlinear term modeled by  $\beta_i^W W_i(\overline{W}_i + \overline{M}_i)$  (respec.  $\beta_i^M M_i(\overline{W}_i + \overline{M}_i)$ ). Note that modeling above assumes implicitly population mixing for each academic level under consideration. This is a usual assumption in type-continuous epidemiological models [7].
- For a specific *Bachillerato* academic level  $i = 1, 2$ , students can also acquire bad academic habits because they autonomously decide to strive less and to give up appropriate study habits due to lack of self-motivation, personal problems, etc. This type of transition is modeled by means of a linear term  $\alpha_i^W W_i$  (respec.  $\alpha_i^M M_i$ ).
- Data collected in Table 1 refer to the end of each academic course, i.e., in September when, according to educational regulation in force in Spain, every student in  $W_1$  and  $M_1$  will pass to  $W_2$  and  $M_2$ , respectively. Taking into account that time  $t$  is measured in years and we identify each academic year in the period 1999/2000 to 2007/2008 with  $0, 1, \dots, 8$ , respectively, parameter  $\delta$  indicated in Figure 1 is defined as follows:

$$\delta = \begin{cases} 1 & \text{if } \frac{9}{12} + j \leq t \leq \frac{10}{12} + j, \quad j = 0, 1, 2, \dots, 8, \\ 0 & \text{otherwise.} \end{cases}$$

- For a specific *Bachillerato* academic level  $i = 1, 2$ , a student in  $\overline{W}_i$  (respec.  $\overline{M}_i$ ) transits to  $W_i$  (respec.  $M_i$ ), when she/he gives up her/his

bad academic habit. An individual in  $\overline{W}_i$  (respec.  $\overline{M}_i$ ) transits to  $W_i$  (respec.  $M_i$ ) at rate  $\gamma_i^W$  (respec.  $\gamma_i^M$ ) proportionally to the size of  $\overline{W}_i$  (respec.  $\overline{M}_i$ ). Analogously to  $\alpha_i^W$  and  $\alpha_i^M$ , parameters  $\gamma_i^W$  and  $\gamma_i^M$  also contain those autonomous decisions adopted by students belonging to  $\overline{W}_i$  and  $\overline{M}_i$ , respectively.

Then, the transitions between these different subpopulations are described by the following coupled nonlinear system of differential equations where the unknowns are  $W_i = W_i(t)$ ,  $M_i = M_i(t)$ ,  $\overline{W}_i = \overline{W}_i(t)$  and  $\overline{M}_i = \overline{M}_i(t)$  ( $t$  denotes time in years),

$$\begin{aligned}
 W_1'(t) &= -\delta W_1(t) - \alpha_1^W W_1(t) - \beta_1^W W_1(t)[\overline{W}_1(t) + \overline{M}_1(t)] + \gamma_1^W \overline{W}_1(t), \\
 \overline{W}_1'(t) &= \alpha_1^W W_1(t) + \beta_1^W W_1(t)[\overline{W}_1(t) + \overline{M}_1(t)] - \gamma_1^W \overline{W}_1(t), \\
 W_2'(t) &= \delta W_1(t) - \alpha_2^W W_2(t) - \beta_2^W W_2(t)[\overline{W}_2(t) + \overline{M}_2(t)] + \gamma_2^W \overline{W}_2(t), \\
 \overline{W}_2'(t) &= \alpha_2^W W_2(t) + \beta_2^W W_2(t)[\overline{W}_2(t) + \overline{M}_2(t)] - \gamma_2^W \overline{W}_2(t), \\
 M_1'(t) &= -\delta M_1(t) - \alpha_1^M M_1(t) - \beta_1^M M_1(t)[\overline{W}_1(t) + \overline{M}_1(t)] + \gamma_1^M \overline{M}_1(t), \\
 \overline{M}_1'(t) &= \alpha_1^M M_1(t) + \beta_1^M M_1(t)[\overline{W}_1(t) + \overline{M}_1(t)] - \gamma_1^M \overline{M}_1(t), \\
 M_2'(t) &= \delta M_1(t) - \alpha_2^M M_2(t) - \beta_2^M M_2(t)[\overline{W}_2(t) + \overline{M}_2(t)] + \gamma_2^M \overline{M}_2(t), \\
 \overline{M}_2'(t) &= \alpha_2^M M_2(t) + \beta_2^M M_2(t)[\overline{W}_2(t) + \overline{M}_2(t)] - \gamma_2^M \overline{M}_2(t).
 \end{aligned}
 \tag{1}$$

The parameters of the model are:

- $\alpha_i^g$ , denotes the rate at which a student of Spanish *Bachillerato* academic level  $i$  and genre  $g$  who belongs to the promoting group passes to have bad academic habits by an autonomous decision. In accordance with personal trait patterns and academic performance of adolescents [8], it is assumed that:

$$\alpha_1^W < \alpha_1^M, \quad \alpha_2^W < \alpha_2^M; \quad \alpha_1^W > \alpha_2^W, \quad \alpha_1^M > \alpha_2^M. \tag{2}$$

- $\beta_i^g$ , denotes the transmission rate at which a student of Spanish *Bachillerato* academic level  $i$  and genre  $g$  adopts bad academic habits due to the negative influence from students that do not pass and belong to the same academic level  $i$  including both genre.
- $\gamma_i^g$ , denotes the rate at which a student of Spanish *Bachillerato* academic level  $i$  and genre  $g$  who has bad academic habits, by an autonomous decision, decides to change her/his bad academic habits and she/he ends up getting into the passing group.

The flow diagram, associated to the model, is depicted in Figure 1.

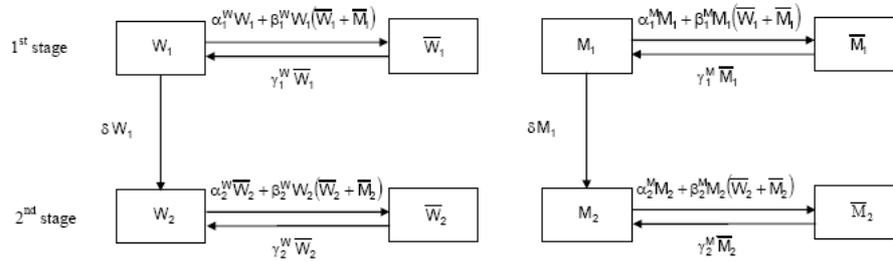


Figure 1: Flow diagram of the mathematical model for *Bachillerato* academic underachievement in Spain.

### 3 Fitting the model and forecasting

In this section we will estimate the parameters of (1) by fitting the model in the mean square sense to the data collected in Table 1. We have carried out computations with *Mathematica 7.0* [9]. As initial conditions of the system of differential equations (1), we take data of the academic year 1999/2000 (corresponding to  $t = 0$ ), so  $W_1(0) = 19.68$ ,  $M_1(0) = 15.24$ ,  $\overline{W}_1(0) = 9.75$ ,  $\overline{M}_1(0) = 9.33$ ,  $W_2(0) = 16.21$ ,  $M_2(0) = 11.64$ ,  $\overline{W}_2(0) = 9.52$  and  $\overline{M}_2(0) = 8.63$ . The estimated parameters are collected in Table 2. The value of the function in the global minimum (i.e., the least square error) that we have obtained is 0.0095.

Table 2: Estimation of the model parameters

Parameter	Value	Parameter	Value	Parameter	Value
$\alpha_1^W$	0.0180156	$\beta_1^W$	0.538867	$\gamma_1^W$	0.285557
$\alpha_1^M$	0.0180175	$\beta_1^M$	2.59115	$\gamma_1^M$	0.853703
$\alpha_2^W$	0.000119175	$\beta_2^W$	0.668998	$\gamma_2^W$	0.270119
$\alpha_2^M$	0.00307025	$\beta_2^M$	1.38138	$\gamma_2^M$	0.405319

Now, once the model is stated and the parameters estimated, we are able to give predictions of each subpopulation over the next few years computing the solutions of the model,  $W_1(t)$ ,  $\overline{W}_1(t)$ ,  $M_1(t)$ ,  $\overline{M}_1(t)$ ,  $W_2(t)$ ,  $\overline{W}_2(t)$ ,  $M_2(t)$  and  $\overline{M}_2(t)$  for values of time  $t$  in the future. The solution to model

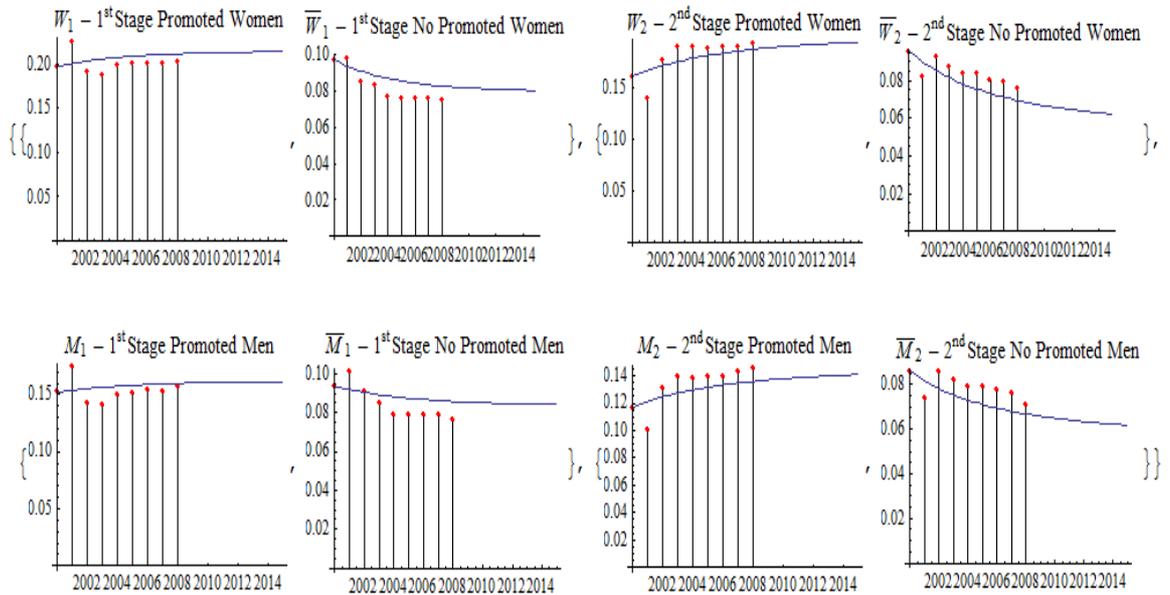


Figure 2: Fitting and prediction of the academic performance of *Bachillerato* Spanish students over the academic years 1999/2000 to 2014/2015

(1) is plotted in Figure 2. According to the model predictions, the percentage of students that pass (do not pass) in the second stage of *Bachillerato* seems to increase (decrease) slightly as time goes. While in the first stage of *Bachillerato* the situation is different. In this case, the model forecasts a stabilization of the percentage of students that pass and do not pass. Regarding the global prediction of the model, note that the total percentage of *Bachillerato* students that will not pass is worrying because it will lie around 30% which still constitutes a high rate. In accordance with the model, the first stage of *Bachillerato* is the key level to begin to combat academic underachievement in *Bachillerato* not only because it has greater associated academic underachievement rates than second stage but also these rates tend to stabilization. From a pedagogical point of view this feature can be explained due to the considerable degree of difficulty of *Bachillerato* subjects with respect to ESO contents.

## 4 Conclusions

In this paper we have developed a continuous model to study academic underachievement in the last educational stage of the Spanish high school, called *Bachillerato*. The major novelty of this contribution is the treatment of school failure as a problem that is transmitted through social contact including its mathematical type-epidemiological modeling. We point out that the results we have obtained agree with conclusions derived from other studies but in addition the model allows seeing how things will evolve in the next future. Hence, it enables us to make an estimation of the number of pupils who will apply for higher education. This is a key issue since academic underachievement is a decisive factor for unemployment and also has a great influence in the future development of a country.

**Acknowledgement:** This work has been partially supported by the Spanish M.C.Y.T. grant MTM2009-08587 and the Universitat Politècnica de València grant PAID06-09-2588.

## References

- [1] <http://www.euractiv.com/en/enterprise-jobs/eu-youth-job-strategy-under-fire-news-497858> (Accessed on 12<sup>th</sup> of September 2011).
- [2] N.A. Christakis, J.H. Fowler, *Connected: The Surprising Power Of Our Social Networks And How They Shape Our Lives*, Brown and Company. Hachette Book Group, 2009.
- [3] N.A. Christakis, and J.H. Fowler. The spread of obesity in a large social network over 32 years *The New England Journal of Medicine*, 357(2007):370–379, 2007.
- [4] M. Fernández Enguita, L. Mena Martínez, J. Riviere Gómez. *Fracaso y Abandono Escolar en España (Academic Underachievement and School Abandon in Spain)* Colec. Estudios Sociales, Vol. 29, Fundación La Caixa, 2010. (In Spanish).
- [5] <http://www.educacion.es/mecd/jsp/plantilla.jjsp?id=310&area=estadisticas> (Accessed on 12<sup>th</sup> of September 2011). (In Spanish).

- [6] <http://www.educacion.es/mecd/jsp/plantilla.jjsp?id=3131&area=estadisticas>  
(Accessed on 12<sup>th</sup> of September 2011).
- [7] J.D. Murray. *Mathematical Biology*. Springer, 2002.
- [8] C. Fierro–Hernández, Patrón de rasgos personales y comportamiento escolar en jóvenes (Personal trait patterns and scholar behaviour), *Revista de Educación*, N.332, (2000) 291–304. (In Spanish).
- [9] <http://www.wolfram.com/products/mathematica>.

# Mathematical modeling of the spread of divorces in Spain

R.Duato\*, L.Jódar

(\*) Instituto Universitario de Matemática Multidisciplinar,  
Universitat Politècnica de València,  
Edificio 8G, Piso 2, 46022 Valencia, España.

November 30, 2011

## 1 Introduction

Over the last decade, family structure in Spain has changed remarkably. One of the main changes observed along this period has been the increase in the divorce rate together with the growth of unofficial couples, in part boosted by the so called social progress laws promoted by the Spanish Socialist Governments between years 2004 and 2011, by the general lack of resilience of the Spanish population and by the impact of a highly volatile economic period in Spain between years 2000 and 2011. For the sake of clarity in the understanding, the above quoted laws are: the express divorce law (2005), the gay marriage (2005) and the divorce terms law (2010) which guarantees free abortion decision to 16 year-old girls during their fourteenth pregnancy week. Furthermore, this change is also enhanced by the increase in parasite singles derived from the recent financial crisis and the above stated highly volatile economic period in Spain. The economy has definitely a high impact in divorce as marriage is a contract [5],[6]. Moreover, there is evidence that the difference in happiness within the couple is closely related with divorce [1]. Like other authors, we measure the influence of the Economy throughout the unemployment level.

Apart from the economic influence, we consider as the other relevant factor in the divorce spread, the social contagious, which is mainly propagated by personal contacts between divorced and married women, see [1],[2], and Letter XXVI of [7].

In this paper we propose a social epidemiology model of divorce regarded as a social disorder propagated by divorced women over married ones. In connection with this, we can state that a variation of divorce related rules, too much favored to women these days, will definitely reduce divorce claims, which are performed by women in almost nine out of

---

\*e-mail: ruben.duatoribera@gmail.com

ten occasions. Besides, this clear advantage in laws for women, makes that when problems arrive in the couple, apart from the above mentioned social contagious, a proportion of these women abuse of the measure of claiming for divorce. To our mind, divorce rules should be equitable for both sexes and not be favoring neither women nor men.

Taking into account the relationship between the gap of happiness within the couple and the divorce, see [1], a suitable measure to reduce the divorce growth derived from this fact could be to focus on girls education, even from childhood periods, teaching them how to fill effectively their leisure time.

The reduction of divorce growth is not only interesting for social stability and for the health of children, but also of general interest to keep the birthrate as well as the pensions of the entire population.

The discrete mathematical model constructed here, divides the Spanish population aged between 20 and 60 years old, into three categories: singles,S; married,M; and divorced,D. Because of the social contagious and the economic influence, populations at year  $n$  of singles  $S_n$ , married  $M_n$  and divorced  $D_n$ , are time dependent and there is a transfer of populations described by a linear-quadratic system of difference equations.

The paper is organized as follows. Section 2 deals with the mathematical model construction. Real data are collected from the Spanish Statistics Institute [3]. Since in the last decade the rate of registered marriages is progressively decreasing, the yearly economic effect is measured as its discrepancy from the average of the rate in the period 2000-2009. As the contagious rate is really tough to measure, we consider the real populations of each category at the end of years 2000 and 2009, and by using a fitting parameter approach, the contagious rate is estimated. In section 3, numerical simulations are performed in order to predict the divorced population at year 2015 by assuming different economic scenarios and also a predicted economic scenario after a possible arrival to the government of the conservative party in 2012. Finally, section 4 highlights the conclusions of the model.

## 2 Model Construction

Our approach may be regarded as an epidemiological model where the social disorder of divorce is propagated within the population from divorced women to married women. This hypothesis is based on [1],[2], and Letter XXVI of [7]. Apart from the social contagious we assume that the economy, which will be represented in our model with the parameter  $\alpha_e(n)$ , together with the decline of marriage [9],[5], has a significant influence in the marital dissolution, i.e. divorces are due to increase when we are living a favorable economic period within the country and are due to decrease when we are facing the opposite situation. Besides, we consider that the appropriate economic index to measure it is the unemployment level.

The entire population is divided into three subpopulations: singles,S married,M and divorced,D. We denote by  $S_n$ ,  $M_n$  and  $D_n$ , the number of singles, married and divorced, respectively at the end of year  $n$ . The population includes all Spanish citizen aged between 20 and 60, both extremes included, and deals with data collected from the Spanish Statistics

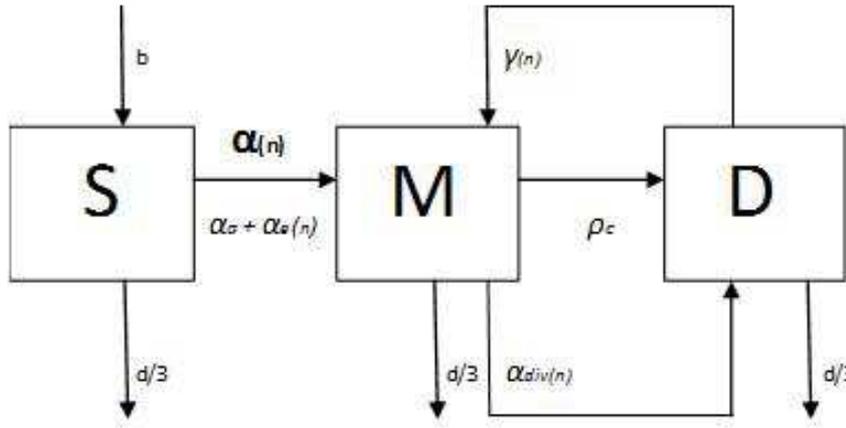


Figure 1: Flow diagram for the divorce model.

Institute [3].

The dynamic of the population can be described as follows (*n*, time in years):

$$S_{n+1} - S_n = (b - (d/3))S_n - (\alpha_a + \alpha_e(n))S_n \tag{1}$$

$$M_{n+1} - M_n = (\alpha_a + \alpha_e(n))S_n - (d/3)M_n - \alpha_{div}(n)M_n + \gamma(n)D_n - \rho_c M_n D_n \tag{2}$$

$$D_{n+1} - D_n = \alpha_{div}(n)M_n - (d/3)D_n - \gamma(n)D_n + \rho_c M_n D_n \tag{3}$$

where the parameters of the model are:

- *b*, average birth rate in Spain between 2000 and 2009.
- *d*, average death rate in Spain between 2000 and 2009.
- $\alpha_a$ , average rate of singles individuals got officially married between 2000 and 2009.
- $\alpha_e(n)$ , rate of singles individuals got officially married during year *n* due to the economic effect.
- $\alpha_{div}(n)$ , rate of married individuals got divorced during year *n*.
- $\gamma(n)$ , rate of divorced individuals got remarried during year *n*.
- $\rho_c$ , the contagious rate derived from contact between divorced and married women. This is based on the fact that divorced women enjoying a good post-divorce situation tend to get in touch and convince those married women facing any kind of doubts in their marriage, according to Letter XXVI of [7] and [2].

We evaluate  $\alpha_e(n)$  as the discrepancy, of the real rate of single individuals got officially married during year  $n$ , from the average rate  $\alpha_a$ .

Parameters  $b$ ,  $d$ ,  $\alpha_a$ ,  $\alpha_{div}(n)$  and  $\gamma(n)$  are obtained from the Spanish Statistics Institute [3].

Parameter  $\alpha_e(n)$  is obtained from  $\alpha_a$  and real data of year  $n$  are obtained from the Spanish Statistics Institute [3].

Parameter  $\rho_c$  will be obtained taking into account real data of divorced individuals at years 2000 and 2009 and fitting the model.

Although parameters  $b$  and  $d$  are yearly changing, the variation is so little that we are considering both as constants.

### 3 Results

Fitting the model with data of populations  $S_{2000}, M_{2000}, D_{2000}$  and  $S_{2009}, M_{2009}, D_{2009}$  one gets:

$$\begin{array}{ll} S_{2000} = 7.933.400 \text{ (34,80\%)} & S_{2009} = 9.673.800 \text{ (36,80 \%)} \\ M_{2000} = 14.188.400 \text{ (62,30\%)} & M_{2009} = 15.242.700 \text{ (57,50 \%)} \\ D_{2000} = 667.100 \text{ (2,90 \%)} & D_{2009} = 1.464.100 \text{ (5,70 \%)} \end{array}$$

The estimated value for the contagious parameter using Mathematica software and minimizing the square error through the Nelder-Mead algorithm [8]:

$$\rho_c = 0,094.$$

The percentage and the total number of individuals according to different economic scenarios are plotted in figs. 2 and 3, respectively.

As we can observe in figure 2, the percentage of divorced population raises from around 6% in 2010 up to an 8,5% - 9%, depending on the economic scenario we are facing at the end of year 2015.

As we can observe in figure 3, the total number of divorced population raises from about 1500000 individuals in 2010 up to 2050000 - 2150000 individuals, depending on the economic scenario we are facing at the end of year 2015.

In order to perform the sensitivity analysis, we have considered the response to variations in the contagious parameters and calculated the predicted divorced population in the year 2015.

### 4 Conclusions

In this paper we propose a social epidemiology model where divorce is regarded as a social disorder propagated by divorced women throughout meetings with married women. Apart from this social contagious, the other relevant factor considered is the Economy of the country measured in terms of unemployment level.

The case studied in this paper is the Spanish one, however, with minor specific changes of another country, the model can be applied to other nations, provided that data are available. In this paper, data are collected from the Spanish Statistics Institute (INE). Considering three possible

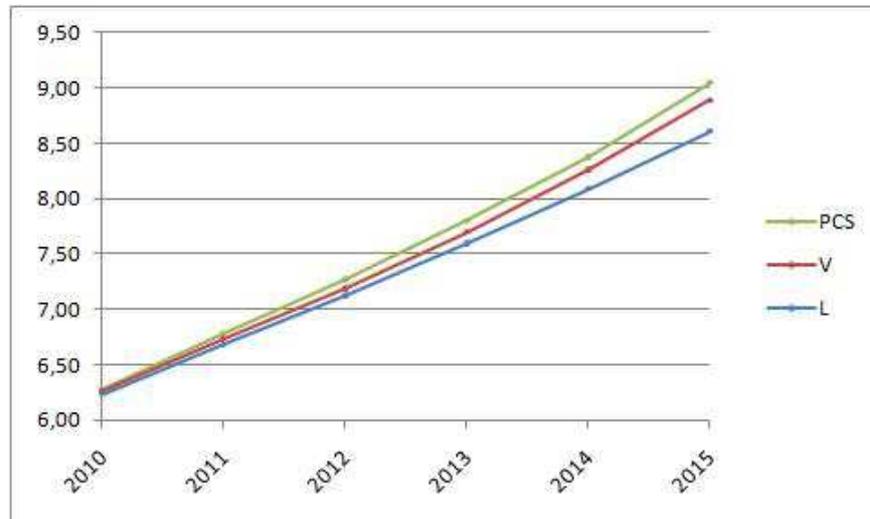


Figure 2: Percentage of divorced individuals according to different economic scenarios.

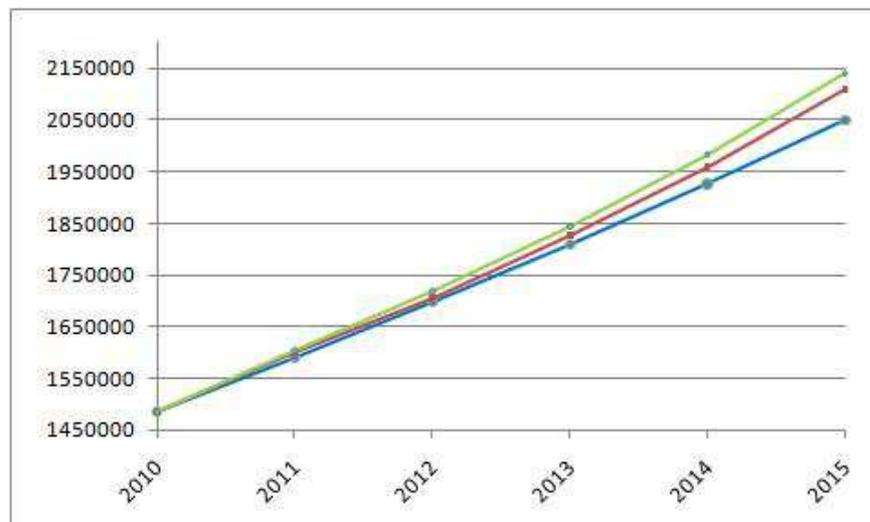


Figure 3: Total number of divorced individuals according to different economic scenarios.

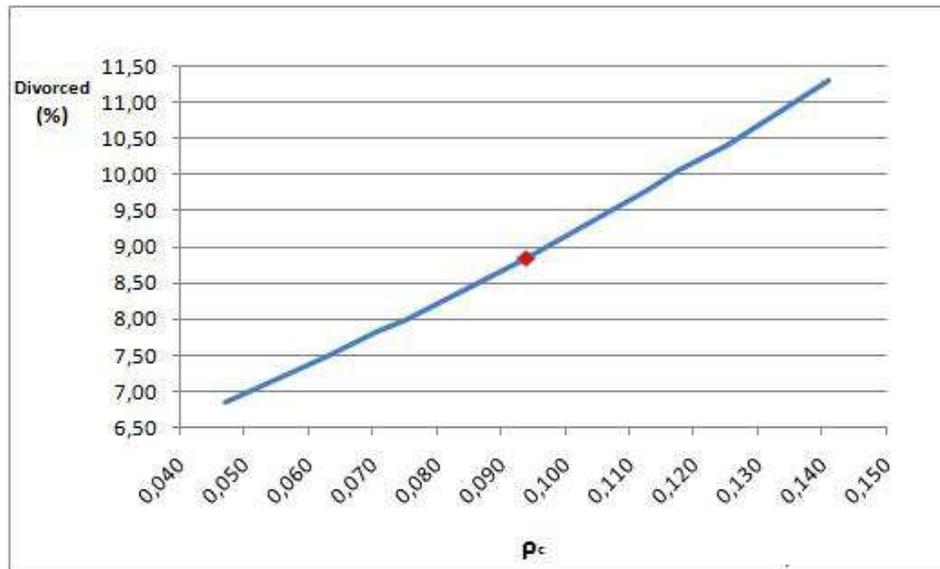


Figure 4: Sensitivity analysis of divorced population in December 2015 in response to variations in the contagious parameter.

different economic scenarios in Spain along the period [2010, 2015], we estimate the divorced population in December 2015. Results show that the divorced population at the end of year 2015 will increase around 50 per cent with respect to the existing divorced population of year 2010. A sensitivity analysis of the model with respect to the contagious parameter  $\rho_c$  proves the robustness of the model. The approach allows us the simulation of alternative scenarios changing the values of the parameters of the model, as well as, the study of the evolution of the divorced population in other forthcoming periods.

## References

- [1] C.Guven, C.Senik, H.Stichnoth, You can't be happier than your wife: Happiness Gaps and Divorce, SOEP papers 261, ZEW DP 10-007.
- [2] N.A.Christakis, J.H.Fowler, Connected: The Surprising Power of Our Social Networks and How they Shape Our Lives, Hachette Book Group, 2009.
- [3] Spanish Statistics Institute. Available at <http://www.ine.es>.

- [4] A.Trillas, De la V a la L, el abecedario de la salida de la crisis(in Spanish). Available at <http://www.elpais.com/articulo/economia/V/L/abecedario/salida/ /crisis/elpepieco/20090412elpepieco/mathaccent ' '0250/relax6/Tes?print=1>.
- [5] G.S.Becker, A Treatise on the Family, 1991. Harvard University Press, Cambridge, MA.
- [6] G.S.Becker, A Theory of Marriage: Part II, Journal of Political economy 82(2) (1974)S11-S26.
- [7] C.S.Lewis, The Screwtape Letters, 2001, Harper Collins, New York.
- [8] J.A.Nelder, R.Mead, A simplex method for functions minimization, Computer Journal, 7 (1964), 308-313.
- [9] Divorce and marriage, less than there used to be, The Economist, Blogs, Britain blighty, 2011. Available at <http://www.economist.com>

# Analysis of the Influence of Different Real Flow Effects on a CFD Anechoic Boundary Condition Based on the MoC

J. Galindo, A. Tiseira, P. Fajardo\*and R. Navarro †

CMT-Motores Térmicos,  
Universidad Politécnica de Valencia

November 30, 2011

## 1 Introduction

Modeling turbomachinery poses particular problems: their geometrical complexity requires the use of viscous compressible 3D codes to capture properly the flow behavior. If the objective of the computation is to obtain the acoustic response of a system the use of long ducts is needed in order to avoid interferences with the boundaries. This fact can be easily understood when compared to the experimental procedure that is commonly used to obtain the acoustic response in an impulse test rig as the one used in [1].

Including those long ducts in the computational domain leads to a dramatic increase of the computational cost [2]. In this frame, the idea of developing a non-reflecting boundary condition (NRBC) that behaves as a long duct seems interesting. Since the acoustic response of this long ducts is supposed to be the lack of wave reflection, several researchers [3, 4, 5] have tried to develop the so-called non-reflecting or anechoic boundary conditions (BC), to be able to eliminate these components without a worsening in the accuracy.

---

\*Corresponding author

†galindo, anti1, pabfape, ronagar1@mot.upv.es

## 2 1D Anechoic BC

An anechoic BC has been previously developed in [6]. An anechoic end is an outlet BC in which information does not travel upstream the boundary condition (anechoic means literally without echo). This non-reflecting BC should behave thus as a long duct, allowing to save computational cost. In order to achieve that the value of  $\beta$  along the characteristic line  $C^-$  will be modified every time a different entropy level reaches this line. Since the computation is discretized in time, the number of entropy variations along a characteristic line  $C^-$  will be finite. However, with this approach, one would still need to take into account the whole *history* of the flow in order to *remember* all the variations of entropy level. To avoid this, an approximation was developed to allow the computation of  $\beta_{BC}(T)$  from  $\beta_{BC}(T - \Delta T)$ :

$$\widetilde{\beta}_{in}^N = \frac{\beta_{in}^{N-1} + \frac{a_{N-1}}{2} \ln \frac{A_a^N}{A_a^{N-1}} + \frac{\lambda_N}{4} \ln \frac{A_a^N}{A_a^{N-1}}}{1 - \frac{1}{4} \ln \frac{A_a^N}{A_a^{N-1}}}. \quad (1)$$

In this way, one only needs to know  $\lambda$  and  $A_a$  (Riemann variables) arriving to the BC at given  $t$ , the value of  $\beta$  and  $A_a$  that arrived to the BC at  $t - \Delta t$  and the speed of sound at the BC at  $t - \Delta t$  in order to compute the anechoic BC. However, as pointed out before, this anechoic BC assumes the existence of an 1D inviscid and adiabatic flow of a perfect gas.

## 3 Non-Perfect Gas

The MoC was originally developed under the hypothesis of perfect gas [8]. However, if an ideal (but not calorically perfect) gas is considered, the MoC can still be used with some modifications. Following [7], if  $c_p = f(T)$ , the following variables replace the Riemann variables:

$$\begin{aligned} \lambda &= \int_{T_{ref}}^T \sqrt{\frac{c_p c_v}{RT}} dT + u \\ \beta &= \int_{T_{ref}}^T \sqrt{\frac{c_p c_v}{RT}} dT - u \\ s &= \int_{T_{ref}}^T \frac{q + uG}{T} dT \end{aligned} \quad (2)$$

In order to deal with non-perfect gas with the anechoic BC, a polynomial law is used for  $c_p(T)$ . Following a similar criteria, the first term of the Riemann variables defined at (2) can be expressed as a polynomial of temperature. It is interesting to remark that the relative importance of this effect depends on the the temperature variation of the given problem. In the particular case of the flow leaving the cylinders of an internal combustion engine this temperature variation are important.

## 4 Viscous Flow

The existence of viscosity in the flow has various effects introducing further complexity in the anechoic BC. First, the diffusive transport induced by viscosity cannot be modeled by the MoC, since it only deals with the hyperbolic part of the Navier-Stokes equations (i.e, the Euler equations). However, since the Reynolds number in the intended applications (flow in ICEs) is high, the diffusive transport should be negligible compared to the convection terms. Moreover, there is a loss in momentum due to the existence of friction. This effect is mainly concentrated in the surroundings of the wall. There exists a modifier that takes into account this effect in the value of the Riemann variables along its respective characteristic line:

$$d\beta|_f = -(\gamma - 1) \left[ 1 - (\gamma - 1) \frac{u}{a} \right] \frac{f}{D} u^2 \frac{u}{|u|} dt \quad (3)$$

Besides the previous two effects, friction has another effect that is relevant not because of the use of the MoC but due to the anechoic BC approach. In its development, it was stated that, for inviscid flows,  $\beta$  only had to take into account the entropy changes. However, in viscous flows, when a  $\lambda$ -wave ( $C^+$ ) arrives, the velocity raises, thus increasing the friction effect by means of the modifier and therefore beta increases. Since this same effect happens at every section, the flow velocity continue decreasing after the  $\lambda$ -wave has passed. When a wave reaches the anechoic BC, the change of beta that still happening at the virtual duct at the right of the boundary condition is sent backwards, so obviously the former anechoic approach can no longer be used to model the acoustic response of an infinite duct with friction, because that duct does echo. Therefore, since the information created at the right of the BC must be taken into account, one should model it, and a 1D-3D coupled simulation such as described in [2] could be helpful.

## 5 Swirling Flow

Among other applications, the anechoic BC is intended to be used as outlet boundary in a turbine simulations. However, the flow after a turbine is no longer 1D, because of the swirl created by the impeller. One of the main problems when using CFD commercial codes is that the BC available are mainly uniform flow BC. This is normally not a big issue if the main interest of the computation relays far from the boundary. However, in the case of the development of new BC, and particularly with the anechoic BC, being able to reproduce the flow behavior close to the boundary is necessary. If the flow is in radial equilibrium, the pressure radial profile satisfies the following equation:

$$\frac{dp}{dr} = \rho(r) \frac{v_{\theta}^2(r)}{r} = F(r) \quad (4)$$

Thus, if radial equilibrium is supposed to be achieved in every time-step, one can obtain the pressure profile at the BC at a certain time-step provided that the pressure at the axis ( $p_{axis}$ ), in order to obtain the integration constant, and  $F(r)$  profile are known. Regarding the value of  $p_{axis}$ , it can be assumed that the MoC can be still applied to the axis, because  $v_{rad} = 0$  due to axial symmetry, and therefore the axis could be assumed to be 1D. On the other side,  $F(r)$  is approximated by its value at the previous time-step. The integration process is started with the known value of  $p(r=0) = p_{axis}$  obtained by the application of the MoC to de 1D flow of the axis.

## 6 Conclusions

An anechoic BC was previously developed, based on the assumption of 1D inviscid and adiabatic flow of a perfect gas. In order to be able to use this anechoic BC in a turbomachinery's outlet, three different real flow behaviors have been studied in this paper. In section 3, the hypothesis of ideal instead of perfect gas has been reviewed. Then, in section 4, the implications of dealing with viscous flow have been investigated. On the one hand, the effect of friction with walls has been incorporated in the BC by means of a proper modifier of the Riemann variables and the use of a correlation for the friction factor. On the other hand, it has been shown that an infinite viscous duct does send information backwards when receiving and incoming pressure wave. Therefore, the anechoic approach would never be able to reproduce

reliably the behavior of a long duct. Finally, in section 5, the existence of swirl (characteristic of a turbine's outlet) has been analyzed. An approach to obtain the whole pressure profile at the outlet has been developed, assuming that the flow is in radial equilibrium.

## References

- [1] A. Broatch, X. Margot, A. Gil, F. D. Denia, A CFD Approach to the Computation of the Acoustic Response of Exhaust Mufflers, *Journal of Computational Acoustic* 13 (2) (2005) 301–316.
- [2] J. Galindo, A. Tiseira, P. Fajardo, R. Navarro, Coupling methodology of 1D finite difference and 3D finite volume CFD codes based on the Method of Characteristics, *Mathematical and Computer Modelling*.
- [3] F. Payri, J. Desantes, A. Torregrosa, Acoustic boundary condition for unsteady one-dimensional flow calculations, *Journal of Sound and Vibration* 188 (1) (1995) 85–110.
- [4] O. Atassi, Nonreflecting boundary conditions for the time-dependent convective wave equation in a duct, *Journal of Computational Physics* 197 (2) (2004) 737–758.
- [5] J. Dea, F. Giraldo, B. Neta, High-order non-reflecting boundary conditions for the linearized 2-d euler equations: No mean flow case, *Wave Motion* 46 (3) (2009) 210–220.
- [6] A. Torregrosa, P. Fajardo, A. Gil, R. Navarro, Non-reflecting boundary conditions based on the method of characteristics for its application in 3D CFD codes, SUBMITTED to *Engineering Applications of Computational Fluid Mechanics*.
- [7] D. E. Winterbone, P. Richard, *Theory of Engine Manifold Design : Wave Action Methods for IC Engines*, John Wiley & Sons, 2000.
- [8] R. Benson, R. Garg, D. Woollatt, A numerical solution of unsteady flow problems, *International Journal of Mechanical Sciences* 6 (1) (1964) 117–144.

# The Variational Principle in Transformation Optics Engineering and some Application

Carlos García-Meca,<sup>\*</sup> Michael M. Tung,<sup>‡</sup>  
Alejandro Martínez<sup>\*</sup>, and Juan Carlos Cortés<sup>‡</sup>

(<sup>\*</sup>) Valencia Nanophotonics Technology Center

(<sup>‡</sup>) Instituto de Matemática Multidisciplinar

Universidad Politécnica de Valencia, Valencia, Spain

November 30, 2011

## 1 Introduction

Transformation optics is an emergent field of engineering with a great impact on recent developments in advanced optical devices. With the help of differential geometry, transformation optics allows us to fully control electromagnetic fields in a previously unknown manner [1, 2], providing a way to engineer curved spaces for light by using suitable media. On the other hand, variational principles mathematically describe in a concise and elegant way a great variety of natural phenomena, independently of a particular choice of coordinate system. In this work, we extend this concept in a differential-geometric framework to formulate transformation optics in an inherently covariant and coordinate-independent form by postulating a Lagrangian function for the most fundamental description of electrodynamics in media. General expressions for transformation optics can be derived from the variational principle associated to this Lagrangian. As an application, we present the construction of a novel immersion lens of potential relevance to bio- and nano-technology. In particular, we focus on improving the shape and reflection properties of such optical devices.

---

<sup>\*</sup>e-mail: cargarm2@ntc.upv.es

## 2 Mathematical basics of transformation optics

Modern optics in its conventional approach is invariably based on Maxwell's macroscopic equations in differential form, describing the electric and magnetic fields produced by matter of charge density  $\rho$  and current density  $\mathbf{j}$  (symbols in bold represent vector quantities, whereas symbols in italics represent scalar quantities). In covariant form these equations read:

$$F_{[\lambda\mu;\nu]} = 0, \quad (1)$$

$$F^{\mu\nu}{}_{;\nu} = \frac{1}{\sqrt{-g}} (\sqrt{-g} F^{\mu\nu})_{,\nu} = \frac{4\pi}{c} j^\mu, \quad (2)$$

where  $c$  denotes the speed of light and we have used Gaussian units. We employ the Einstein summation convention with  $\mu = 0, 1, 2, 3$ , denoting by zero the temporal component. Eq. (1) represents the source-free (homogeneous) equation and Eq. (2) the source (inhomogeneous) equation of the electromagnetic field, where  $j^\mu = (c\rho, \mathbf{j})$  is the contravariant current four-vector and  $g < 0$  is the determinant of the underlying pseudo-Riemannian metric  $g_{\mu\nu}$  with indefinite signature  $(-, +, +, +)$ .  $F_{\mu\nu}$  is the electromagnetic field-strength or Faraday tensor, which can be expressed in terms of the four-potential  $A_\mu$  as

$$F_{\mu\nu} = 2A_{[\mu;\nu]} = A_{\mu;\nu} - A_{\nu;\mu} = A_{\mu,\nu} - A_{\nu,\mu} = 2A_{[\mu,\nu]}. \quad (3)$$

The Faraday tensor is related to the electric field  $\mathbf{E}$  and the magnetic induction  $\mathbf{B}$  by  $E_i = F_{i0}$  and  $B_i = \frac{1}{2} \epsilon_{ijk} F^{jk}$ . Latin indices only run over spatial values and  $\epsilon_{ijk}$  denotes the completely antisymmetric three-dimensional Levi-Civita symbol.

In order to describe electromagnetic processes, we introduce  $A^\mu$  on a smooth four-dimensional manifold  $M$  endowed with a Lorentzian metric  $\mathbf{g}$ . Following the terminology of Marsden *et al.* [3], we further identify *base space*  $B = M$ , which constitutes standard spacetime, and *ambient space*  $P$ , given by  $A^\mu : B \rightarrow M$ . Then,  $N = B \times P$  is the *configuration space* with coordinates  $(x^0, x^i, A^0, A^j)$ , so that any particular constellation or state of the system is uniquely determined by the mapping  $B \rightarrow N$ . We seek a general principle to find the evolution of  $A^\mu$  on the given manifold  $(M, \mathbf{g})$ . In geometric mechanics, the Lagrangian function  $L$  completely governs the behaviour of a deterministic system, and it suffices to define  $L$  on a tangent bundle  $TM$ , which is identical with the corresponding phase space. The tangent bundle  $TM$  consists of the manifold  $M$  and its tangent spaces  $T_p M$  for all  $p \in M$ . Figure 1 represents the tangent bundle  $TM$  consisting of manifold  $M$  and its tangent spaces  $T_p M$  for all  $p \in M$ . In this case, however, we

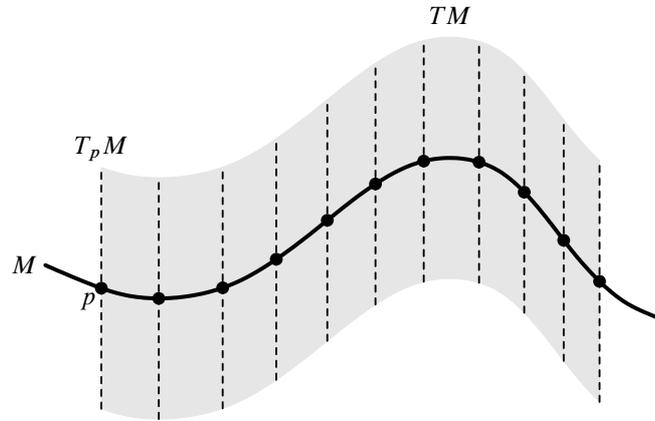


Figure 1: Standard tangent bundle  $TM$  consisting of manifold  $M$ , usually pseudo-Riemannian spacetime, and its tangent spaces  $T_pM$  as used in geometric mechanics.

will require the partial derivatives of a configuration with respect to all *spacetime* coordinates. This generalization leads to the jet bundle  $J^1N = B \times TP$ , see *e.g.* [4]. Figure 2 gives a schematic of the jet bundle  $J^1N$  consisting of configuration manifold  $N = M \times P$  and the tangent bundle  $TP$  of ambient space  $P$ . The corresponding Lagrangian function in the jet-bundle description of first-order classical field theories will therefore be a mapping  $L : J^1N \rightarrow \mathbb{R}$ .

Hamilton's principle states that the dynamics or field dynamics of a physical system is governed by a simple variational principle (see *e.g.* Refs. [5, 6] in the case of diffusion), whose solutions of equations of motion are determined by the extremum of the action functional  $\mathcal{A}$ . For electrodynamics the variation of the following action integral must vanish:

$$\delta \mathcal{A} = \delta \int_{\Omega} d^4x \sqrt{-g} L(A_\alpha, A_{\alpha;\beta}) = 0. \quad (4)$$

The invariant volume element is  $\sqrt{-g} d^4x = \sqrt{-g} dx^0 dx^1 dx^2 dx^3$ , since integration occurs over spacetime. We let  $\Omega \subset M$  be a bounded, closed set of spacetime and postulate the following form for the *Lagrangian density*

$$\mathcal{L} = \sqrt{-g} L = \sqrt{-g} \left( -\frac{1}{32\pi} X^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} - \frac{1}{c} j^\alpha A_\alpha \right) \quad (5)$$

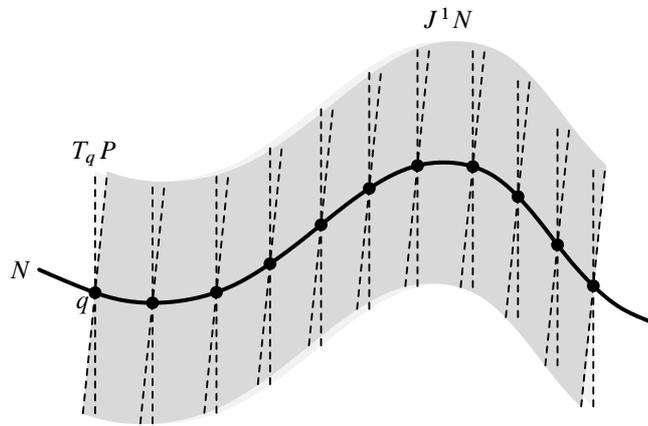


Figure 2: The jet bundle  $J^1 N$  consisting of configuration manifold  $N = M \times P$  and the tangent bundle  $TP$  of ambient space  $P$  as used in first-order classical field theory.

where  $X^{\mu\nu\rho\sigma}$  denotes the so-called *constitutive tensor*, which couples linearly to the fields to yield the quadratic kinetic term. It is customary to introduce the *excitation tensor*  $G^{\mu\nu} = \frac{1}{2} X^{\mu\nu\rho\sigma} F_{\rho\sigma}$ , with  $D_i = H_{i0}$  and  $H_i = \frac{1}{2} \epsilon_{ijk} G^{jk}$  ( $\mathbf{D}$  is the electric displacement and  $\mathbf{H}$  is the magnetic field). This last equality is known as the (*linear*) *constitutive relation*. The solution to Eq. (5) is provided by the corresponding Euler-Lagrange equations, which in combination with Eq. (1) fully control the dynamics of the underlying physical system. They can be determined by computing the associated functional derivative which, by requiring it to be zero, yields:

$$G^{\alpha\beta}{}_{,\beta} = \frac{4\pi}{c} j^\alpha. \quad (6)$$

These are just the two inhomogeneous Maxwell equations for optics in macroscopic media, which confirms the validity of the postulated Lagrangian.

Now we are in the position to derive the fundamental expressions of transformation optics from the Lagrangian formalism we have established. First, imagine some *virtual space* characterized by a metric  $g'$ , in which light propagates as desired. For simplicity, we assume that this space is void (the procedure can be straightforwardly extended to non-empty virtual spaces), so that the constitutive tensor is given by  $X^{\mu\nu\rho\sigma} = g^{\mu\rho} g^{\nu\sigma} - g^{\mu\sigma} g^{\nu\rho}$ . Thus, its corresponding Lagrangian

density reads

$$\mathcal{L}' = \sqrt{-g'} \left[ -\frac{1}{32\pi} \left( g'^{\mu\rho} g'^{\nu\sigma} - g'^{\mu\sigma} g'^{\nu\rho} \right) F'_{\mu\nu} F'_{\rho\sigma} - \frac{1}{c} j'^{\alpha} A'_{\alpha} \right]. \quad (7)$$

On the other hand, Eq. 5 represents the *physical space* we live in. We want light to propagate in physical space as it does in virtual space. For this, we reinterpret the coordinates of virtual space as those of physical space (in exceptional cases they may coincide from the onset) and take advantage of the fact that both Lagrangians (and therefore the resulting solutions for the fields) are formally identical, provided that we fill physical space with the following media and sources:

$$X^{\mu\nu\rho\sigma} = \frac{\sqrt{-g'}}{\sqrt{-g}} \left( g'^{\mu\rho} g'^{\nu\sigma} - g'^{\mu\sigma} g'^{\nu\rho} \right) \quad \text{and} \quad j^{\alpha} = \frac{\sqrt{-g'}}{\sqrt{-g}} j'^{\alpha}. \quad (8)$$

This powerful theory allows us to realize both, curved-space metrics (*e.g.*, that of black holes) in a flat physical space, and/or any deformation caused by coordinate changes [1, 2]. Naturally, for deformations arising exclusively from transformations,  $g$  and  $g'$  represent the same geometry expressed in different coordinate systems. We will use this last approach in the next section.

### 3 Application: Flat reflectionless immersion lens

An immersion lens made up of a dielectric medium with refractive index  $n$  can improve the diffraction-limited resolution of free space by a factor of  $1/n$ , see Ref. [7]. Its usual geometry is that of a hemisphere, although other curved surfaces can be used. Thus, one side of the system is flat, while the other one is curved. For some applications, it would be desirable that both surfaces were flat. In addition, there appear reflections at the boundary of the lens with free space due to the difference in the refractive indices of both media. In order to overcome both drawbacks, we used the previous theory to transform a dielectric slab into an flat immersion lens by gradually expanding the fields in the transverse direction to propagation. This results in a compression of the transverse wavevector component of the fields going through the lens. This way we increase the range of components of the angular spectrum that correspond to propagating waves in free space, improving the resolution of the system by the expansion factor. Moreover, it can be shown that an expanding device is reflectionless if the medium to be expanded has a refractive index equal to the expansion factor and

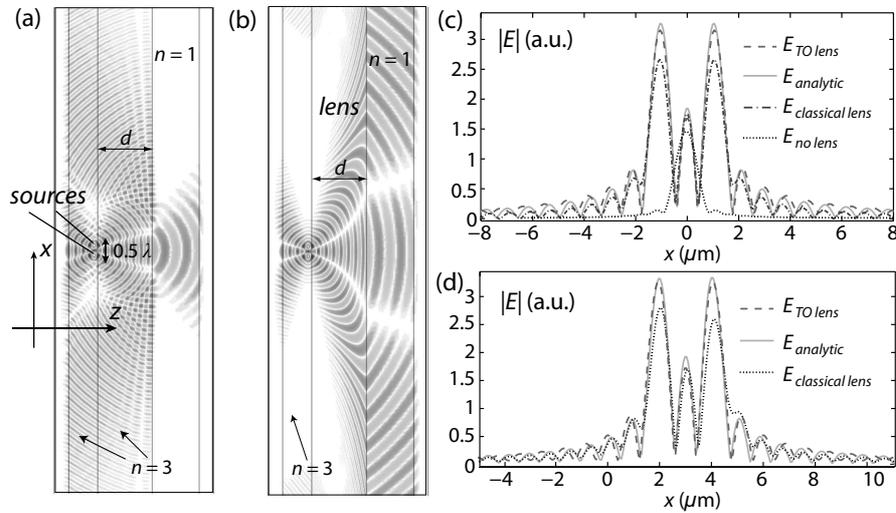


Figure 3: Electric field generated by two sources embedded in a dielectric with  $n = 3$  (a) without and (b) with the designed lens between the dielectric and air ( $n = 1$ ). (c) Reconstructed field amplitude for different cases. (d) Same as in (c) but with the sources  $1\mu\text{m}$  away from the lens center.  $\lambda_0 = 1.5\mu\text{m}$ .

the output medium is free space [8]. For instance, we can use the transformation  $t' = t, x' = x/(1 + Cz), y' = y, z' = z$ , where  $C$  is a constant such that  $F_x = 1/n_1 = 1 + Cd$  ( $F_x$  is the inverse expansion factor in the  $x$ -direction,  $d$  is the device thickness, and  $n_1$  is the original dielectric refractive index). Then, we use Eq. (8) to calculate the properties of our device. As an example, we simulated a lens with  $n_1 = 3$  in COMSOL Multiphysics. To test the device, two cylindrical sources separated by  $\lambda_0/2$  are placed inside a medium with  $n_1$ . When the fields radiated by these sources directly exit to air, *viz.* Figure 3(a), reflections appear and the radiation pattern in air is that of two sources separated by  $\lambda_0/2$ . When we use the proposed lens, no reflections appear and the radiation pattern is that of two sources separated by  $1.5\lambda_0$ , *viz.* Figure 3(b). The fields that an external optical system would reconstruct for the cases with and without lens ( $\mathbf{E}_{TO\ lens}$  and  $\mathbf{E}_{no\ lens}$ ) are depicted in Figure 3(c). In the first case the two sources can be clearly observed, while in the second one some components are lost and we only detect a broad unique source. In the image obtained with the lens there is a magnification factor of 3. We verified this by comparing the image reconstructed from the fields radiated by two sources in air separated by  $1.5\lambda_0$  ( $\mathbf{E}_{analytic}$ ) with  $\mathbf{E}_{TO\ lens}$ , observ-

ing an excellent agreement. We also include the fields we would reconstruct if we used a classical hemispherical immersion lens ( $\mathbf{E}_{\text{classical lens}}$ ), whose amplitude is lower than that of  $\mathbf{E}_{\text{TO lens}}$  due to reflections. Finally, we show in Figure 3(d) the effect of separating the sources from the lens origin. This does not affect the performance of the proposed lens, while the classical lens introduces distortion.

## 4 Conclusions

We have developed a conceptual framework to represent macroscopic electrodynamics in media by a Lagrangian on a pseudo-Riemannian manifold, which straightforwardly allows to identify any virtual space possessing especially designed optical properties with real, physical space. This novel approach encompasses all previous formulations of transformation optics in a compact and elegant way. We applied this technique to design an immersion lens with superior properties for possible application in bio- and nano-technology.

## References

- [1] J. B. Pendry, D. Schurig, and D. R. Smith, Controlling Electromagnetic Fields, *Science*, 312:1780–1782, 2006.
- [2] U. Leonhardt and T. G. Philbin, General Relativity in Electrical Engineering, *New Journal of Physics*, 8:247, 2006.
- [3] A. Lew, J.E. Marsden, M. Ortiz and M. West, Asynchronous Variational Integrators, *Archive for Rational Mechanics and Analysis*, 167:85–146, 2003.
- [4] A. Echeverría-Enríquez, M. C. Muñoz-Lecanda, and N. Román-Roy, Non-standard connections in classical mechanics, *Journal of Physics A*, 28:5553–5567, 1995.
- [5] M. M. Tung, Basics of a differential-geometric approach to diffusion: Uniting Lagrangian and Eulerian models on a manifold, *Mathematics in Industry*, 12:897–901, 2007.
- [6] M. M. Tung and A. Hervás, A differential-geometric approach to model isotropic diffusion on circular conic surfaces in uniform rotation, *Mathematics in Industry*, 15:1053–1058, 2010.

- [7] S. M. Mansfield and G. S. Kino, Solid immersion microscope, *Applied Physics Letters*, 57:2615–2616, 1990.
- [8] C. García-Meca, M. M. Tung, J. V. Galán, R. Ortuño, F. J. Rodríguez-Fortuño, J. Martí, and A. Martínez. Squeezing and expanding light without reflections via transformation optics *Optics Express*, 19:3562–3575, 2011.

# Evolutionary support vector regression algorithm applied to the prediction of the thickness of the chromium layer in a hard chromium plating process

P.J. García Nieto\*, F.J. de Cos Juez<sup>†</sup>,  
F. Sánchez Lasheras<sup>‡</sup> and J.A. Vilán Vilán<sup>+</sup>

(\*) Department of Mathematics,

Faculty of Sciences, University of Oviedo, 33007 Oviedo, Spain.

(†) Project Management Area, Mining Department,

University of Oviedo, 33004 Oviedo, Spain.

(‡) Department of Construction and Manufacturing Engineering,

University of Oviedo, 33204 Gijón, Spain.

(+) Department of Mechanical Engineering,

University of Vigo, 36200 Vigo, Spain.

November 30, 2011

## 1 Introduction

The hard chromium plating process consists mainly of three operations (see Figure 1) and their corresponding quality inspections. It has already been presented by the authors in previous research [1] and is briefly outlined below:

- *Vapour degreasing*: This is mainly a cleaning operation. It is performed on the piece in order to ensure the cleanliness of the surface.

---

\*Corresponding author. E-mail: lato@orion.ciencias.uniovi.es

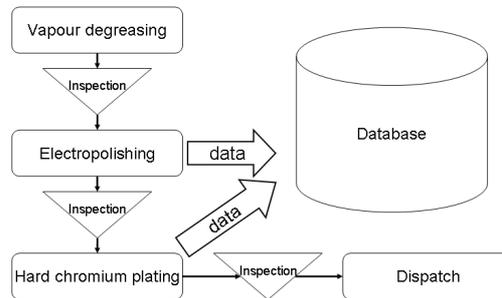


Figure 1: Real BMD versus predicted ANN BMD of the patients used for the validation.

- *Electropolishing*: An electrochemical process that removes material from the work piece. The performance of this operation before the hard chromium plating helps to ensure a good roughness of the surface to be coated.
- *Hard chromium plating*: This is the operation whereby a thin layer of chromium is deposited on the workpiece.

This research work uses evolutionary support vector regression algorithms for the prediction of the thickness of the chromium layer in a hard chromium plating process. Evolutionary support vector machines (ESVMs) [2] is a novel technique that assimilates the learning engine of the state-of-the-art support vector machines (SVMs) [3], but evolves the coefficients of the decision function by means of evolutionary algorithms (EAs) [1, 4].

## 2 Materials and methods

The support vector regression (SVR) model depends on the three hyper-parameters  $\epsilon$ ,  $C$  and  $\gamma$  [3]. The idea is to start with a coarse-grain grid search of parameters, and then carry out a refinement of the search using an evolutionary computation algorithm. The classical evolutionary programming (CEP) algorithm was first described in a paper by Bäck et al. [1, 4], and later analyzed by Yao et al. [1, 4] and Lee and Yao [1, 4]. It is used to optimize a given function  $f(\mathbf{x})$  (in our case, the validation error), i.e. obtaining  $\mathbf{x}_0$  so that  $f(\mathbf{x}_0) < f(\mathbf{x})$ , with  $\mathbf{x} \in [lim_{inf}, lim_{sup}]$  (note that in this

Table 1: Intervals in which the coarse-grain search was performed.

Variable	Intervals
$C$ parameter	$[0.01, 1]$ $[1, 10^2]$ $[10^2, 10^3]$ $[10^3, 10^4]$
$\varepsilon$ parameter	$[10^{-3}, 10^{-2}]$ $[10^{-3}, 10^{-2}]$ $[10^{-2}, 10^{-1}]$ $[10^{-1}, 1]$
$\gamma$ parameter	$[10^{-3}, 10^{-2}]$ $[10^{-2}, 10^{-1}]$ $[0.1, 1]$ $[1, 10^2]$

case, the values of  $[lim_{inf}, lim_{sup}]$  are set by the initial coarse grain).

### 3 Results

The SVR topology consists of seven input variables: (1) iron content in the electropolishing bath; (2) electropolishing time; (3) electropolishing temperature; (4) thickness removed by electropolishing; (5) chromic acid content hard chromium bath; (6) hard chromium time; and (7) hard chromium temperature. Some of them belong to the electropolishing operation and others to the hard-chromium plating. The output variable is the thickness of the hard chromium layer deposited on the piece. The training set was divided into 10 subsets, and a measure of cross-validation was calculated. The root mean square error (RMSE) in this cross-validation process was then used to guide the evolutionary search of SVR hyper-parameters. In order to perform the coarse grain search, the ranges listed in Table 1 were used as the initial step of the algorithm for the hyper-parameters  $C$ ,  $\varepsilon$  and  $\gamma$ . The performance of the SVR was studied for each interval in its middle value and afterwards the evolutionary programming algorithm was applied. This algorithm was run with 500 generations and 200 individuals, as in such generations the convergence criterion for the RMSE that was fixed in a value of  $10^{-6}$  had already been achieved. Once the SVR hyper-parameters optimization was completed, using the coarse-grain search and the evolutionary programming algorithm, the RMSE provided by the final model was analyzed and compared. The result of the RMSE for the best forecasting model obtained in this paper was 0.001939. It is important to note that the results presented by the authors in a previous study [1], where a neural network model tuned by means of DOE was employed, reached as optimum values of 0.00230255 for a Tansig transfer function and 0.00218176 for the sigmoid transfer function. A comparison of the results of the SVR-tuned model and the real values of

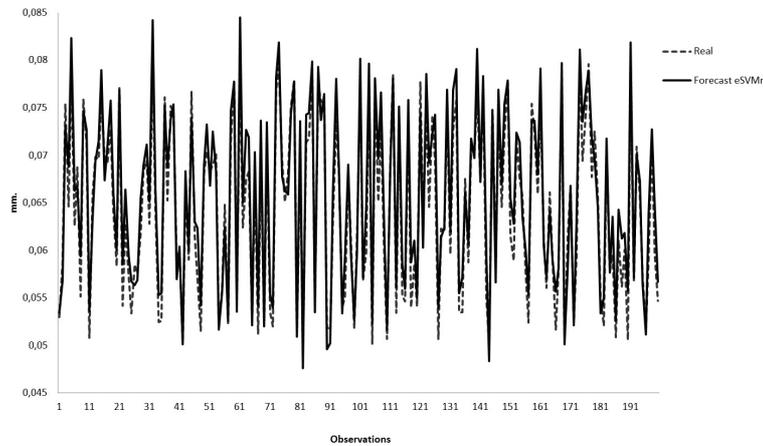


Figure 2: Results corresponding to the SVR-tuned model versus the real values of the hard-chromium layer.

the hard-chromium layer are presented in Figure 2. As can be observed, the forecasting model is able to predict, with high accuracy, the real value of the hard-chromium layer. Finally, Figure 3 presents not only the predicted and real values of the hard-chromium layer for the same subset of data, but also the best model obtained by means of a multilayer perceptron neural network model with a sigmoid transfer function, tuned by means of DOE [1] (the best model in our previous research).

## 4 Discussion and conclusions

In this research work, a support vector machine for regression (SVR) was implemented so that the hyper-parameters search was conducted by means of evolutionary computation. This new and innovative methodology is used for the prediction of the thickness of the hard chromium layer in a hard-chromium industrial process. The complete system for prediction, including the evolutionary SVR algorithm gave a very good performance in our industrial application, outperforming a previous model based on multilayer perceptron neural networks and optimized by means of DOE.

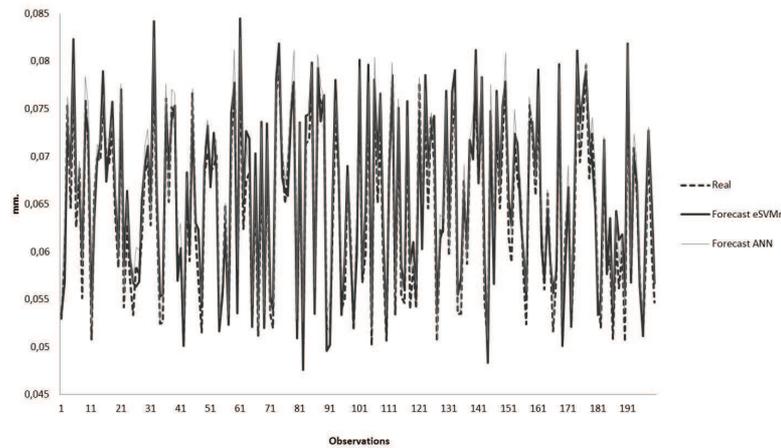


Figure 3: Results corresponding to the SVR-tuned model versus the real values of the hard-chromium layer, and the comparison with the best model obtained by means of a multilayer perceptron neural network model with a sigmoid transfer function.

## Acknowledgements

This work was partially supported by the Spanish Science and Innovation Ministry, (Grant reference AYA2010-18513). Furthermore, authors would like to express their gratitude to the Department of Education and Science of the Principality of Asturias for its partial financial support (Grant reference FC-11-PC10-19). We would also like to thank Anthony Ashworth for his revision of the English grammar and spelling of the manuscript.

## References

- [1] F. Sánchez Lasheras, J.A. Vilán Vilán, P.J. García Nieto and J.J. del Coz Díaz. The use of design of experiments to improve a neural network model in order to predict the thickness of the chromium layer in a hard chromium plating process *Mathematical and Computer Modelling*, 52(7-8):1169–1176, 2010.
- [2] S. Salcedo-Sanz, E.G. Ortiz-García, A.M. Pérez-Bellido, A. Portilla-Figueras. Title of the paper *Short term wind speed prediction based on*

*evolutionary support vector regression algorithms*, 38:4052–4057, 2011.

- [3] I. Steinwart, and A. Christmann, Support Vector Machines. New York, Springer, 2008.
- [4] A.P. Engelbrecht, Computational Intelligence: An Introduction. River Street, Hoboken, New Jersey, Wiley & Sons, 2007.

# The benefits of hybrid methodologies in Data Mining and Knowledge Discovery. Applications to Human Behaviour

K. Gibert\*

Knowledge Engineering and Machine Learning group & Dep. Statistics and Operations Research,  
Universitat Politcnica de Catalunya, Barcelona, Spain

November 30, 2011

## 1 Introduction

Data Mining, also known as Knowledge Discovery from Databases (KDD) is a discipline oriented to extract new and useful decisional knowledge from data [2]. The KDD results provide decision support and contribute to better management or control of highly complex organizations or processes. KDD is a generic umbrella providing a methodological framework in which: i) Pre and post-processing are also considered as important parts of the process as data exploitation itself; ii) Both qualitative and quantitative data are rigorously analyzed; iii) Prior domain knowledge should be used for improving quality of discovered knowledge.

The data exploitation itself is in the kernel of the KDD process and many different data mining techniques can be considered, depending on the target problem and data structure, among others. Statistics and Artificial Intelligence provide many suitable methods. However, in real applications, it is better to analyze numerical and qualitative data jointly, as they both can be relevant for decision making. In many applications, hybrid AI& Stats methodologies for KDD constitute very powerful tools to catch the structure of very complex domains and provide useful knowledge. Some of the methodologies we developed in the area of AI& Stats around clustering

---

\*e-mail: karina.gibert@upc.edu

problems, one of the most frequent problems addressed in the context of KDD [9] are overviewed. They have been used for better understand some phenomena related to human behaviour.

Problem addressed	AI& Stats tools	Collab. entity	Main results	Impact on target domain
Disability critical for elderly people	CIBR+ CPG+ PAG	Osped. S.Lucia, Rome (Italy)	At equal levels of physical impairments, depressed patients loose functionality [1]. In a second step nested binary logistic regressions were used to build a model for recognizing the profile of a new patient and the PAG was used to provide a visual predictive tool [8].	Standard treatments should be associated to every profile and the PAG helps to diagnoses in clinical everyday activities.
Dependency in mental pat. (preparation of Dependency Law, Catalunya)	CIBR+ CPG	Health Dep. (catalan gov.)	Dependency ows to volition and requires health care services (HCS) related to monitoring therapeutic adherence (MTA) or avoid suicide rather than help in dressing or standing-up [4]. An implicit profile was elicited of singles with chaotic use of HCS (51 assistances to emergency services in a week). Experts recognized it as a useless expensive profile.	Elegibility criteria for the accessibility to the benefits of the Law regarded the discovered profiles. Also, standard kits of HCS to be offered for every profile could be identified, like telephonic MTA for singles to avoid crisis, rationalize use of HCS and improving QOL.
Poor knowledge of Mental Health in under-dev. countries	CIBR+ CPG+ TLP	World Health Org. (WHO)	Mental health systems were tipified and they range from the most institutionalized where mainly inpatient policies in mental hospitals works to the most community oriented, including some mixt structures in poorest countries, where lack of resources produces evolution by other formulae [6].	Assessment of WHO for improving mental health in every country is currently being provided, according to the type of mental health system that the country currently shows.
Treatment response Acquired Brain Dammage	CIBR+ CPG+ CCEC	Institut Guttmann, Badalona (Spain)	Among the response patterns found, a group of patients with older lesions that resist to treatment, and another group of older patients with partial response to treatment, where executive functions were not reestablished [5].	Exercices overstimulating executive brain areas are proposed to partial respondents. Psycho-social support is offered to non-respondent to keep QOL.
Social inclus. of patients with Spinal Cord Lesion	CIBR.xS+ TLP+ TG	Institut Guttmann, Badalona (Spain)	Prototypical evolution of perceived QOL over time after discharge were found. As known, loosing optimism has long-term effects on physical impairment. Environmental factors associated were found, like dysfunctional couples [3] or inherent character of patient.	Social workers and Pshychologist could plan supporting activities oriented to prevent this negative evolution and to conserve or increase the QOL of the patients over time as much as possible.
Undesired side-effects of ECT in acute schizophrenia	KDSM	U. Psych. Hosp. Bellvitge Barcelona (Spain)	Loose of memory was not linked to particular patient characteristics, nor to the whole therapy, but to some specific interaction between patient and every single electroconvulsive therapy (ECT) session, invalidating the classic prepost analysis used for this topic from the 50s[11].	Further studies changed approach and considered intermediate measures after every ECT session in order to identify the reasons for loose of memory.

Table 1: Overview of applications

## 2 Applications and impact in target domain

Table 1 shows several applications with some hybrid AI& Stats methods and the final impact into the target domain. All of them have been developed in collaboration between the UPC and several entities indicated.

## 3 AI& Stats tools involved

**Clustering based on rules (CIBR):** generalizes hierarchical clustering to formally include prior expert knowledge into the process [1]. Final clusters are coherent with prior expert knowledge and easier to interpret.

**Class Panel Graph (CPG):** Post-processing tool designed to visualize the conditional empirical probability distributions of all variables against classes in a single/few page(s) [5]. It compacts much information, providing enough perspective for identifying the characteristic variables of a certain class. Powerful to support the conceptualization of the classes, but it requires some technical skills to be properly analyzed.

**Traffic lights panel (TLP):** A symbolic re-elaboration of the CPG. The *semantics* of the variables is introduced by means of a color coding for the main trend of the variable in each class. Color-coding follows the traffic lights symbology. Powerful to visualise class particularities for non technical end-users and supports conceptualization [5].

**Conceptual Characterization by Embedded Conditioning (CCEC):**

A rules-based method using the overlapping of the variables' probability distributions to detect the main class characteristics. A formal alternative to the CPG and TLP, based on logic inference [10]. Useful in simple cases with few variables, and can be automatized.

**Profile's assesment grid (PAG):** A post-processing tool that visualizes the class predictive (logistic) equations in a unitary cube. New objects can be places and classified [8], by using a minimum number of relevant variables. It showed to be very useful in clinical praxis.

**CIBR by States (CIBRxS):** Accepts repeated measures at discrete time or space points [7]. Static patterns are found at every measures wave by CIBR. The *trajectories graph* (TG) displays all static patterns with the object trajectories over time or space. A frequentist analysis permits selection of both typical trajectories (decision support) or the most rare trajectories (exceptions analysis). Direct interpretation of dynamic patterns.

**Knowledge Discovery in repeated very short serial measures (KDSMS):** Repeated very short series per object (or patient), measured

after a certain target event [11] are treated. Types of responses to the target event are elicited by considering the object as a blocking factor for the series, which are not independent. Baseline patterns provide prior knowledge.

## 4 Conclusion

Hybrid Ai& Stats tools are useful to provide understanding of very complex phenomena like disability or schizophrenia, where both classical data analysis, or pure knowledge-based methods couldn't elicit useful decisional knowledge, giving either confusing results or already well-known results in the area. They also help to elicit implicit expert knowledge relevant for decision making.

Proper preprocessing is crucial for correctly addressing the problem goals. It requires quite a lot of expertise and knowledge domain and concerns missing data treatment, feature selection, correct codification of variables, outliers treatment, transformation of variables when important, like in pre-post analysis (ECT) where differences eliminate the object effect.

Post-processing is also capital to convert the results into really useful decisional knowledge in the target domain. Post-processing is method dependent and no clear guidelines are established yet. Several postprocessing tools have been developed for clustering to increase the impact of KDD into the real target domain, as shown in table 1.

## References

- [1] R. Annichiarico, K. Gibert, and *et al.* *JRRD*, 41(6A):835–845, 2004.
- [2] U. Fayyad and *et alt.* *Advances in Knowledge Discovery and Data Mining*, chapter From Data Mining to Knowledge Discovery: An overview. AAAI/MIT Press., 1996.
- [3] K. Gibert, Garca-Rudolph A, and *et al.* *Studies in Health Technology and Informatics*, 150:579–583, 2009.
- [4] K. Gibert, C. Garca Alonso, and L. Salvador-Carulla. *Health Research Policy and Systems*, 8(28):1–16 (Highly Accessed), 2010.
- [5] K. Gibert, A Garca-Rudolph, and *et al.* *Acta Informatica Medica*, 16(4), 2008.
- [6] K. Gibert, Salvador-Carulla L, and *et al.* *JAMIA*, in press, 2011.

- [7] K. Gibert, G. Rodriguez-Silva, and I. Rodriguez-Roda. *EMS*, 25:712–723, 2010.
- [8] K. Gibert, R. Rodriguez-Silva, and R. Annicchiarico. doi:10.1016/j.mcm.2011.10.046.
- [9] K. Gibert and M. Snchez-Marr. *EMS*, 26:(983–985), 2011.
- [10] A. Pérez-Bonilla and K. Gibert. In *LNCS (4756): 653–663*. Springer, 2007.
- [11] J. Rodas, K. Gibert, and J. Rojo. *LNAI*, 2504:228–238, 2002.

# Using the proper generalized decomposition to compute the dominant mode of a nuclear reactor

S. González-Pintor<sup>‡</sup>, D. Ginestar<sup>\*</sup>, and G. Verdú<sup>‡</sup>

(<sup>‡</sup>) Departamento de Ingeniería Química y Nuclear, Universidad Politécnica de Valencia,  
Camino de Vera, 14. 46022 Valencia, Spain,

(<sup>\*</sup>) Instituto de Matemática Multidisciplinar, Universidad Politécnica de Valencia,  
Camino de Vera, 14. 46022 Valencia, Spain.

November 30, 2011

## 1 Introduction

The Proper Generalized Decomposition (PGD) method is a recently developed strategy [1] that uses the fact that some models encountered in science and engineering are defined in multidimensional spaces. When mesh-based discretization techniques are applied, these models exhibit what is called the curse of dimensionality, i.e., the number of unknowns increases exponentially with the spatial dimension of the geometry, leading to a very large amount of unknowns for high dimensional problems.

A differential eigenvalue problem useful in nuclear reactor physics is known as the Lambda modes problem [5]. This problem will be used to explain how we will use a PGD approximation to approximate the dominant eigenvalue. This problem in the approximation of one energy group is of the form

$$\mathcal{L}\Phi = \frac{1}{\lambda}\mathcal{M}\Phi, \quad (1)$$

---

\*e-mail:dginesta@mat.upv.es

where  $\mathcal{L}$  is the neutron loss operator and  $\mathcal{M}$  is the neutron production operator. This problem is assumed to be defined in a 2-dimensional separable domain  $\Omega = \Omega_x \times \Omega_y$ .

The proposed algorithm can be divided into an initialization step, and then an iteration process composed by a Rayleigh-Ritz procedure, another step for checking the convergence and a fourth step for enriching the approximation.

## 1.1 Initialization

The PGD method to compute the dominant eigenvalue of a differential eigenvalue problem and its corresponding eigenvector, is based on expressing the eigenvector  $\Phi$  associated with the dominant eigenvalue  $\lambda$  in the  $k$ -th step as

$$\Phi^k(x, y) = \sum_{i=1}^k \alpha_i F_1^i(x) F_2^i(y) . \quad (2)$$

For a first step,  $k = 1$ , the solution is completely separable, i.e.,  $\Phi^1(x, y) = \Phi_1(x)\Phi_2(y)$ , and the separable eigenvalue problem will be solved by means of an alternating direction method [1] using one-dimensional solvers as in [2].

## 1.2 Rayleigh-Ritz Procedure

Now we have the actual solution in the form

$$\Phi^{k+1}(x, y) = \gamma_1 \Phi^k(x, y) + \gamma_2 \delta\Phi(x, y) , \quad (3)$$

where  $\Phi^k$  and  $\delta\Phi$  are known normalized functions and  $\gamma_j$  are unknown coefficients. We need to approximate the value of  $\gamma_j$  to obtain an approximation for the eigenvector associated with the dominant eigenvalue. A Rayleigh-Ritz procedure is used to project the eigenproblem onto a small subspace, obtaining a  $2 \times 2$  projected eigenvalue problem

$$\lambda L\vec{\gamma} = M\vec{\gamma} . \quad (4)$$

The dominant eigenvalue  $\lambda$  of problem (4) is our approximated dominant eigenvalue, and the eigenvector is given by (3).

### 1.3 Checking convergence

From the solution  $\Phi^k$  at iteration  $k$ , given by Equation (2), we compute the residual  $Re$  related to Equation (1), and stop if  $Re < \varepsilon$ , where  $\varepsilon$  is the desired tolerance for the residual of our approximated eigenpair.

### 1.4 Enriching the approximation

Now we want to improve our approximation by means of adding a correction term [3]  $\delta\Phi(x, y)$ . Thus, we are interested in solving approximately a linearized Newton correction equation of the form

$$(\lambda\mathcal{L} - \mathcal{M})\delta\Phi + \delta\lambda\mathcal{L}\Phi = -(\lambda\mathcal{L} - \mathcal{M})\Phi . \quad (5)$$

where the second order term  $\delta\lambda\mathcal{L}\delta\Phi$  has been eliminated. This is, essentially, an iteration of the Newton method to compute eigenpairs for an eigenvalue problem [3]. To solve this correction equation, a PGD based technique will be used, obtaining a correction  $\delta\Phi$  composed of a sum of separable functions.

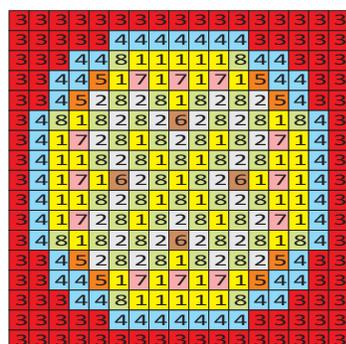
## 2 Numerical Results

The PGD method to compute the dominant eigenvalue and its corresponding eigenfunction for a differential eigenvalue problem, presented above, has been tested with a 2-dimensional reactor core, called Biblis 2D reactor [4], in one group of energy.

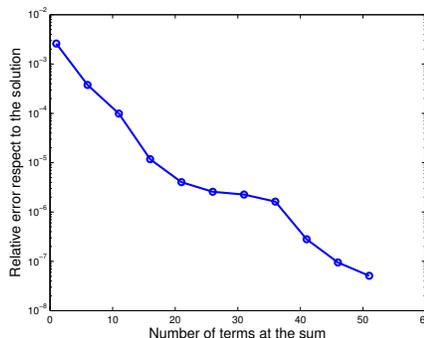
The geometry of the core has been extended to be a Cartesian product of 1-dimensional spaces. This geometry together with the fuel assembly distribution is shown in Figure 1(a). The nuclear cross sections of the different materials composing the core are the ones of the first energy group at [4].

The reference solution for this problem has been computed with a nodal collocation method [5] using  $K = 5$  polynomials in the expansions. The reference value for the dominant eigenvalue is  $\lambda_1 = 0.618502$ .

The convergence of the eigenvalue as a function of the steps  $k$  of the PGD method is shown in Figure 1(b). We observe that for this problem a small number of steps of the proposed method are enough to converge the eigenvalue. The convergence of this problem is also good for the eigenfunction using a small number of steps of the PGD method.



(a)



(b)

Figure 1: Geometry of reactor Biblis 2D (a) and Relative error evolution for the eigenvalue  $\lambda$  (b)

## References

- [1] A. Ammar, B. Mokdad, F. Chinesta, and R. Keunings. A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modeling of complex fluids. *Journal of Non-Newtonian Fluid Mechanics*, 139(3):153–176, 2006.
- [2] S. González-Pintor, D. Ginestar, and G. Verdú. Continuous pseudospectral methods for the neutron diffusion equation in 1d geometries. *Mathematical and Computer Modelling*, 50(5-6):783–793, 2009.
- [3] S. González-Pintor, D. Ginestar, and G. Verdú. Updating the lambda modes of a nuclear power reactor. *Mathematical and Computer Modelling*, In Press, Corrected Proof:–, 2010.
- [4] A. Hébert. Application of the hermite method for finite element reactor calculations. *Nuclear Science and Engineering*, 91:34–58, 1985.
- [5] G. Verdú, D. Ginestar, V. Vidal, and J.L. Muñoz-Cobo. 3D  $\lambda$ -modes of the neutron-diffusion equation. *Annals of Nuclear Energy*, 21(7):405–421, 1994.

# Modeling adaptative social behaviour on epidemics with dynamical networks

Gilberto González-Parra <sup>†\*</sup>;  
Jose L. Herrera <sup>†</sup>, and Abraham J. Arenas<sup>‡</sup>

(<sup>†</sup>) Grupo de Matemática Multidisciplinar, Universidad de los Andes, Venezuela,

(<sup>‡</sup>)Departamento de Matemáticas y Estadística, Universidad de Córdoba, Montería, Colombia

November 30, 2011

## 1 Introduction

Recently dynamical networks have been used in order to study evolutionary process such as disease propagation. These networks include real actions since they take into account for the mutual interaction between the network topology and the states of the nodes. Static networks also have been used but they adjust better to process where the structure does not change. There are several important studies that deal with epidemics using networks [4]. In addition, different type of network structures such small world and scale-free have been studied in order to understand epidemics [3, 2]. On the other hand recently some interesting research using adaptative networks to model disease spread have been presented [1]. For instance in [5] authors propose a interesting susceptibleinfectedsusceptible (SIS) epidemic model coupled with opinion dynamics to investigate the effects of social impact on the epidemic spreading in complex networks. The model consists of two types of nodes with different behavior patterns, active nodes and passive nodes.

In this work we model adaptative social behavior on epidemics with dynamical networks. These dynamical network models try to reproduce with

---

\*e-mail: gcarlos@ula.ve

accuracy the simultaneous time evolution of the disease and the underlying network topology. In this paper, we study the time evolution of the disease and the underlying network topology using a social disease behavioral paradigm. In order to obtain simulation results we apply Monte Carlo method to observe stationary network dynamics. In this research we are interested in studying the effect that rewiring (social contacts changes) has on the disease dynamics.

## 2 Epidemic network model

The classical *SIR* model considers a population with only three compartments, susceptible:  $S(t)$ , infected,  $I(t)$ , and recovered,  $R(t)$ . The compartmental deterministic mathematical model can be represented analytically by the following nonlinear system of ordinary differential equations,

$$\begin{aligned}\dot{S}(t) &= -\beta S(t)I(t), \\ \dot{I}(t) &= \beta S(t)I(t) - \nu I(t), \\ \dot{R}(t) &= \nu I(t).\end{aligned}\tag{1}$$

where the parameters of the model are:  $\beta$ , transmission rate due to contact and  $\nu$ , recovery rate from the infection. An important issue that need to be considered in the simulation of the networks, is that in the contagion process the value of the probability of contagion per unit time  $\beta$  needs to be modified for each  $\Delta t$  in order to have an equivalent model to the continuous one [6]. Here we are interested in studying the effect of an equivalent value of  $\beta$  and we use a parameter  $\hat{\beta}$  which it is not equal to the parameter  $\beta$  of the continuous model.

There are different algorithms to simulate the evolution of the disease under human behavior. In this work we assume that the number of nodes (individuals) and the total connections or links in the initial social network maintain invariant in the simulation. However, the connections between nodes changes over the simulation. It is important to remark that in some cases the simulation process finish and not every node is recovered since some nodes remain susceptible without getting infected.

The main idea in this model is that the social relationship rewiring is related with the disease transmission. In this approach we assume that if the disease transmission is high the individuals disregard the possibility of

changing their relationships with infected individuals since they think that any way they will get infected and with low infectivity they break their relationship since now they think can avoid the disease. We study in this case the effect that has the social relationship rewiring on the dynamics of the diseases. Additionally, we study the effect that has the connectivity degree  $\langle k \rangle$  on the evolution of the epidemics. The steps of the algorithm are:

- Create an initial network with  $N$  nodes where  $(N - 1)$  are susceptible and only one node is infected.
- Advance one time step  $\Delta t$  and choose one node randomly.
- If the selected node is susceptible then we select from his neighborhood nodes one randomly. If this node is infected then with probability  $\hat{\beta}$  the susceptible node is infected and with probability  $(1 - \hat{\beta})$  the susceptible node break the contact and links with another node randomly.
- If the selected node is infected then with probability  $\hat{\nu}$  the infected node changes his state to recovered, if not we select from his neighborhood one node randomly. If this node is susceptible then with probability  $\hat{\beta}$  the susceptible node is infected and then with probability  $\hat{\nu}$  the infected node changes his state to recovered.
- If the selected node is recovered, the network stays invariant.
- Based on the previous steps the network topology may change in each time step  $\Delta t$ . Then go to step 2 and this process finish when there are not infected nodes in the network.

### 3 Simulation of the dynamic epidemic network model

Simulations are made using the Monte Carlo method, assuming constant population and connections. Comparisons between different networks varying connectivity degree, disease transmission rate and relink probability are performed, but only some representative cases are shown. It is considered that the stationary regime is achieved when in the *SIR* network model there are only two classes remaining; susceptible and recovered. Initially it is allocated only one infected individual in one of the communities in order to

mimic the arrival of one infected individual to a region. Numerical results show the mean and standard deviations of each scenario.

On the left hand side of Fig. 1 it can be seen that the proportion of susceptible individuals in the stationary state approaches to zero when the probability of disease transmission  $\hat{\beta}$  increases. On the right hand side of Fig. 1 we present the simulation times to arrive to the stationary state (without infected). It can be observed that longer simulation times are obtained for higher values of  $\hat{\beta}$ . *This result seems contra intuitive* and this is why is important to remark to understand the disease propagation in the real world. One way to explain this fact is that when  $\hat{\beta}$  is low the probability to relink to avoid infected nodes is high. Therefore, many nodes do not get infected and the disease disappears faster. On the other hand, when  $\hat{\beta}$  is high, the disease is propagated to more nodes and the permanence on the system is higher. Additionally, in this last case the relink to avoid the disease has a lower probability.

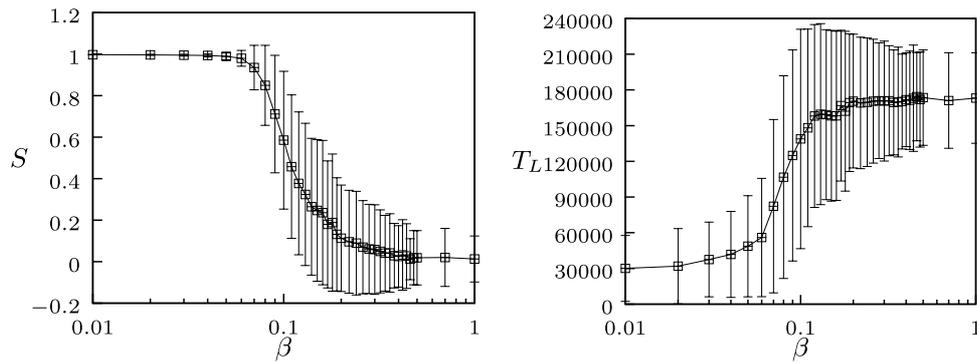


Figure 1: Simulation of the epidemics with different parameter values of  $\hat{\beta}$ ,  $\langle k \rangle = 8$ ,  $\nu = 0.02$  and 500 realizations . Proportion of susceptible individuals in the stationary state (Left side). Time to extinction ( $T_L$ ) and to arrive to the stationary state (Right side).

Another issue that is important to study the coevolution of network and disease is the connectivity degree  $\langle k \rangle$ . In order to study this effect we perform Monte Carlo simulations with  $\langle k \rangle = 4, 8$  y  $12$ . In the left hand side of Fig. 2 it can be observed that the maximum peak of the graph corresponding to the proportion of infected nodes is higher for  $\langle k \rangle = 12$ . Furthermore, this peak is achieved in less time than with the other networks with lower connectivity degree  $\langle k \rangle$ . In this way, we obtain that the diseases

propagate stronger and faster with higher connectivity degree. In the right side of Fig. 2 it can be seen that the extinction time of the epidemic for a fixed value of  $\hat{\beta}$  increases for higher values of the connectivity degree. In this way when the network has low connectivity the disease does not catch all the individuals and extincts faster. Thus, a society with low connectivity degree has the potential to eliminate faster the epidemics. On the other hand, for  $\hat{\beta} > 0.5$  the extinction time is approximately equal for different connectivity degrees.

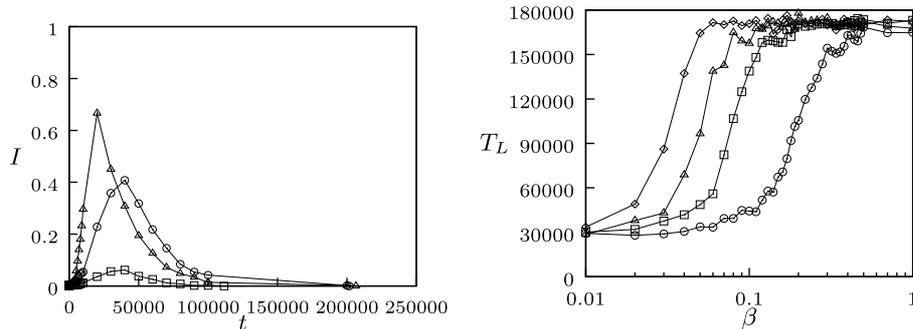


Figure 2: Proportion of infected for  $\langle k \rangle = 4$ (squares),  $\langle k \rangle = 8$  (circles) and  $\langle k \rangle = 16$  (triangles) (Left). Extinction time ( $T_L$ ) for different parameter values of  $\hat{\beta}$  using connectivity degree  $\langle k \rangle = 4$ (circles),  $\langle k \rangle = 8$  (squares),  $\langle k \rangle = 16$  (triangles) and  $\langle k \rangle = 24$  (rhomboids).

## 4 Discussion and conclusions

In this work, we proposed a behavior social paradigm under a epidemic scenario. In this paradigm we assume that if the disease transmission is high the individuals disregard the possibility of changing their relationships with infected individuals since they think that any way they will get infected and with low infectivity they break their relationship since now they think can avoid the disease.

The main result obtained is that with a high disease transmission probability the permanence of the disease on the population is larger. One way to explain this behavior is that for low values of the disease transmission the society has better protection and the disease extincts faster without catching all the individuals. On the other hand, with low rewiring probability the society has low protection and the disease permanence is longer affecting more

individuals. Another result is that with higher social connectivity degree the disease has higher and faster percolation on the society. Thus, a society with low connectivity degree has the potential to eliminate faster the epidemics.

From a practical point of view we can conclude that the social relationship rewire probability changing the structure of the network has an important influence on the evolution of the epidemics. As expected the rewire probability protects more individuals from the society against the disease and the epidemics extincts faster. Moreover, the results obtained in this research show that the transmission disease and network topology, play an important role in determining the evolution and outcome of a particular epidemic scenario on an adaptive network. In regard to future interesting researches it can be mentioned that networks with time variant connectivity degree and improving the techniques for estimating networks parameters are open issues.

## References

- [1] S. Funk, M. Salath, and V.A.A. Jansen. Modelling the influence of human behaviour on the spread of infectious diseases: a review. *J. R. Soc. Interface*, 7:1247–1256, 2010.
- [2] G. González-Parra, L. Acedo, R.-J. Villanueva Micó, and A.J. Arenas. Modeling the social obesity epidemic with stochastic networks. *Phys. A: Stat. Mech. Appl.*, 389(17):3692–3701, 2010.
- [3] T. Gross, C.J.D. D’Lima, and B. Blasius. Epidemic dynamics on an adaptive network. *Phys.Rev.Lett.*, 96(20):208701, 2006.
- [4] M.J. Keeling and K.T.D. Eames. Networks and epidemic models. *J. R. Soc. Interface*, 2(4):295–307, 2005.
- [5] S. Ni, W. Weng, and H. Zhang. Modeling the effects of social impact on epidemic spreading in complex networks. *Phys. A: Stat. Mech. Appl.*, 390(23-24):4528–4534, 2011.
- [6] Y. Wang, D. Chakrabarti, C. Faloutsos, C. Wang, and C. Wang. Epidemic Spreading in Real Networks: An Eigenvalue Viewpoint. In *SRDS*, pages 25–34, 2003.

# Credit risk management: A multicriteria approach to assess creditworthiness

V. Giménez<sup>(†)</sup>,  
F. García<sup>(†)</sup>\*, F. Guijarro<sup>(†)</sup>

(†) Departamento de Economía y Ciencias Sociales,  
Facultad de Administración y Dirección de Empresas,  
Universidad Politécnica de Valencia,  
Camino de Vera s/n, 46022, Valencia,

November 30, 2011

## 1 Introduction

Financial institutions can face different types of risk [1]:

- Market risk: unexpected changes in prices or rates.
- Credit risk: unexpected changes in value associated with changes in credit quality.
- Liquidity risk: the risk that the costs of adjusting financial positions will increase substantially or that a company will lose access to financing.
- Operational risk: associated to human factors: fraud, system failures, trading errors.
- Systemic risk: chain reaction crises affecting the whole market.

---

\*e-mail: fergarga@esp.upv.es

In recent years, credit risk has been a focal point, mainly due to the international financial crisis that has considerably affected a large number of financial institutions [2]. Banks classify their corporate clients in terms of default probability, a variable that is essential for ranking clients when making decisions regarding financing. Default probability is usually obtained as a linear function from a set of economic and financial variables that provide information about different aspects of corporate clients: size, liquidity, solvency, profitability, debt, etc. A scoring or ranking model combines these variables in order to obtain an accurate assessment of default probability, thus serving to automate the evaluation process of default risk measurement within a financial institution. In order to obtain the afore-mentioned function, a set of explanatory variables  $x$  and a binary variable  $y$  corresponding with the company situation are related:  $y$  takes the value of 1 if the company has defaulted, and 0 otherwise.

The problem can be summarised as finding a function (generally a linear one) that relates the dependent variable (default) with the set of explanatory variables (economic and financial information). The most common statistical techniques for establishing the previous relation are logistic regression and discriminant analysis [3, 4, 5, 6, 7]. A compendium of the main research work that has made use of these techniques can be found in [8], where their restrictions are also analysed. A common problem for logistic regression and discriminant analysis is that the sample needs to contain a relatively high number of companies in default situation so that the number of healthy and distressed companies is as balanced as possible. However, in the years preceding the current crisis, default companies made up a very reduced percentage as regards the total: less than 5% in all cases. Even in the case of the crisis getting much worse this percentage will in no case become comparable to that of healthy companies.

These problems greatly restrict the practical application of statistical techniques. Furthermore, by segmenting the sample of companies in two groups (solvent and insolvent companies) it is not possible to differentiate among the companies which belong to the same group. Therefore, although two companies may be currently solvent, that does not necessarily mean that their financial situation and solvency level are strictly equal. However, in the statistical models both shall be considered to be within the group of solvent companies, leaving aside their different degree of solvency.

This papers objective is to propose a model that, by means of a linear function, can relate the probability of default with the economic and financial

information of the company (1) whilst going beyond the previously indicated restrictions for statistical techniques. Unlike statistical approximations, in our proposal it will not be necessary to segment the sample into just two groups (healthy and distressed) but the degree of solvency/insolvency will be differentiated in a continuous manner. For this, the intervention of an expert or group of experts is needed, so objective information (company accounting information) is combined with subjective information (the opinion of the expert or experts). The solvency function to be obtained will take the usual linear form of (1):

$$solv(x) = a_1x_1 + a_2x_2 + \dots + a_nx_n \quad (1)$$

In this way, given companies  $p$  and  $q$ , company  $p$  is more solvent than  $q$  when  $solv(p) > solv(q)$ . That is, the probability of default for company  $q$  would be greater than the probability of default for company  $p$ . Therefore, the solvency function would allow us to create a global ranking of companies using their economic and financial information.

The objective of this paper is not to know the absolute values obtained with function , but only the relative ones. What is relevant is the comparison that can be made between the companies using this function. This can be highly relevant when creating a rating of companies. Hence the range of values of function turns out to be irrelevant and needs not be restricted to that of a probability [0..1] which is the assumption of the statistical models. The rest of the paper is structured as follows. In the second section the model is introduced. The proposed model is a goal programming model which obtains a company ranking according to its solvency level. The paper concludes with a summary and conclusions, as well as a list of the references consulted.

## 2 Default elicitation based on ordinal rankings

The model suggested in this paper is empirical in nature. For this reason it moves away from statistical foundations of other classic approximations. Below, we introduce the different phases that make up the proposal and enable us to obtain the coefficients of function (1):

Step 1. Database selection.

In this first stage a set of companies is selected from which function (1) will be deduced. The only requisite is that these companies must belong to the same sector, given that the relationship between accounting information and solvency can differ according to the sector analysed. The database will be made up of economic and financial variables and ratios. It is not necessary to distinguish between default and healthy companies. That is to say, unlike the statistical models, it is not necessary in our proposal to have a binary variable indicating whether or not the company is solvent.

Step 2. Partial ordinal ranking of companies according to their credit-worthiness.

In this phase the objective information of the economic and financial variables is combined with the opinion of an expert or panel of experts (from this point onwards and without losing generality the singular form shall be used). The accounting information of a group of  $n$  companies that have previously been selected by means of random sampling without replacing the database of step 1 is shown to the expert. The set of variables is collected in vectors  $\mathbf{x}$  for each of the  $n$  companies:

$$\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n \tag{2}$$

The expert creates a partial ordinal ranking of these from greater solvency (position 1) to lesser solvency (position  $n$ ). This way, after consulting the economic and financial information, the expert ranks the companies according to their solvency criteria. Doing this, we obtain the following ranking:

$$\mathbf{x}_1 \geq \mathbf{x}_2 \geq \dots \geq \mathbf{x}_n \tag{3}$$

where relation  $\mathbf{x}_i \geq \mathbf{x}_j$  implies that in the opinion of the expert, company  $\mathbf{x}_i$  is preferable to company  $\mathbf{x}_j$ , bearing in mind the economic and financial characteristics of both.

In (3) a relation of partial dominance between the companies has been taken instead of one of complete dominance. The reason is that the expert is shown a limited number of variables and he or she must rank the companies by only considering these variables. In the case of a draw between two or more of them, even though they may have different values in the variables, no new information (variables) is shown for the tiebreak, for which reason the option of considering two or more companies to be similar must be permitted. This phase is repeated during  $R$  rounds. The process is completed when all companies have been compared once at least.

The cardinal ranking of other techniques such as Direct Rating (DR), in which the expert values each company on a scale of 0 to 100, or Point Allocation (PA), where the expert shares out 100 points between all the companies, could have provided more information during the process [9, 10]. However, we believe that the experts precision in the comparisons would have been considerably reduced: with these techniques the expert must not only decide which company is more solvent than another, but must also quantify the exact degree of solvency. This makes the complexity of his or her analysis much greater. With our approximation where the expert provides the ranking, the model should be capable of quantifying the different degree of solvency of the companies. Doing this, part of the experts subjectivity is eliminated when it comes to quantifying the degree of company solvency in a more precise manner.

Step 3. Obtaining weights  $a_j$  of function (1).

In order to obtain function (1), a binary goal programming model is used from the ordinal rankings suggested by the expert in Step 2. The aim of this model is to estimate coefficients  $a_j$  of the solvency function, so the difference between the ranking suggested by the expert and the ranking estimated by function (1) is minimized. This difference is calculated as the number of inconsistencies between both rankings. For example, imagine that only three companies are used and the expert has ranked them as  $1 > 2 > 3$ .

$$\begin{aligned}
 & \sum_{r=1}^R \sum_{j=1}^{n-1} \sum_{l=j+1}^n b_{rjl} \\
 & \text{s.t.} \\
 & \sum_{i=1}^m a_i x_{rij} - \sum_{i=1}^m a_i x_{ri(j+1)} = d_{ij}^+ - d_{ij}^- \quad j = 1 \dots n - 1, r = 1 \dots R \\
 & \sum_{i=1}^m a_i x_{rin} \geq 1 \quad r = 1 \dots R \\
 & b_{rjl} \geq -1/M \sum_{k=j}^{l-1} (d_{rk}^+ - d_{rk}^-) \quad j = 1 \dots n - 1, l = j + 1 \dots n, r = 1 \dots R \\
 & d_{rk}^+, d_{rk}^- \geq 0 \quad r = 1 \dots R, j = 1 \dots n - 1 \\
 & b_{rjl} \in \{0, 1\} \\
 & M \text{ is a big positive number}
 \end{aligned} \tag{4}$$

If function (1) obtained by the model ranks them as  $2 > 1 > 3$ , then there

would only exist an inconsistency between companies 1 and 2. The greater difference would be given if the function would rank them as  $3 > 2 > 1$ , where 3 relations of dominance would be unfulfilled:  $1 > 2$ ,  $1 > 3$  y  $2 > 3$ . Therefore, the maximum number of inconsistencies in a round in which  $n$  companies are compared is  $n(n - 1)/2$ , whilst the minimum will be 0 when both rankings coincide. Another possibility consists of quantifying the cardinal difference. That means to count the inconsistencies not by the number of reversals in the ranking, but by the difference in the value they obtain using function (1). In order to obtain the coefficients of function (1) the (4) binary goal programming model is applied, where  $b_{rjl}$  is a binary variable referring to round  $r$ . It takes value 1 when the value of function (1) of company  $j$  is less than the value that company  $l$  takes in the function and 0 when the opposite is true. Therefore the binary variable will take value 1 when the estimated function inverts the order designated by the expert between companies  $j$  and  $l$ . The objective function minimizes the sum of these binary variables given that the aim is that the function estimated by the model respects, as far as possible, the rankings created by the expert. Coefficients  $a_i$  ( $i=1..m$ ) are the unknown variables of the model and correspond to coefficients  $a$  of function (1);  $x_{rij}$  is the value of the financial variable  $i$ -th for company  $j$ -th, bearing in mind that company  $j$ -th is referred to round  $r$  (due to the type of sample chosen, in each round company  $j$  will logically be different). Therefore, the function estimated by the model for company  $j$  in round  $k$  would be obtained as  $\sum_{i=1}^m a_i x_{rij}$ ;  $d_{rj}^+$  and  $d_{rj}^-$  represent, respectively, the positive and negative deviation variables, associated with company  $j$ -th in round  $r$ .

### 3 Summary and Conclusions

One of the restrictions of the statistical techniques used in determining credit risk is in the relative high number of companies in default that is required. Hence the default functions obtained using these techniques have a limited predictive capacity. Another problem is that when considering default to be a binary variable, information is lost as regards the different degree of solvency of the companies. This paper shows a model to estimate the solvency function, that allows for a set of companies to be ranked according to their solvency level, by considering a relevant set of economic and financial variables and going beyond the described restrictions of statistical techniques. To this end, the use of a binary goal programming model that incorporates not

only economic and financial information, but also the opinion of an expert or group of experts.

## References

- [1] D. Duffie, K.J. Singleton (2003) *Credit Risk: Pricing, Measurement, and Management*. Princeton University Press. New Jersey.
- [2] D. Cabedo, J.M. Tirado (2011) Sector concentration risk: A model for estimating capital requirements, *Mathematical and Computer Modelling* 54, 1765–1772.
- [3] E.I. Altman (1968) Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *The Journal of Finance* 23, 589-609.
- [4] R. Edmister (1972) An empirical test of financial ratio analysis for small business failure prediction, *Journal of Financial and Quantitative Analysis* 7, 1477-1493.
- [5] R.A. Eisenbeis (1977) Pitfalls in the application of discriminant analysis in business, *Journal of Finance* 32, 875-900.
- [6] G.V. Karels, A.J. Prakash (1987) Multivariate normality and forecasting of business bankruptcy, *Journal of Business Finance and Accounting* 14, 573-593.
- [7] C. Zavgren (1983) The prediction of corporate failure: the state of the art, *Journal of Accounting Literature* 2, 1–35.
- [8] S. Balcaen, H. Ooghe (2006) 35 years of studies on business failure: an overview of the classic statistical methodologies and their related problems, *The British Accounting Review* 38, 63–93.
- [9] P. Bottomley, J. Doyle, R. Green (2000) Testing the Reliability of Weight Elicitation Methods: Direct Rating Versus Point Allocation, *Journal of Marketing Research* 37, 508–513.
- [10] H. Alfares, S. Duffuaa (2009) Assigning Cardinal Weights in Multi-Criteria Decision Making Based on Ordinal Rankings, *Journal of Multi-Criteria Decision Analysis* 15, 125–133.

# Application of graph-spectral methods in the vulnerability assessment of water supply networks\*

J.A. Gutiérrez–Pérez<sup>‡†</sup>, M. Herrera<sup>‡</sup>, R. Pérez–García<sup>‡</sup>  
and E. Ramos–Martínez<sup>‡</sup>

(<sup>‡</sup>) Fluing–IMM, Universitat Politècnica de València,  
C. de Vera s/n, 46022 Valencia, Spain.

November 30, 2011

## 1 Introduction

There are various methods to approach a solution to graph clustering, among the most important are the *spectral methods*. The spectral methods are based on eigenvalues and eigenvectors of a block-diagonal matrix conveniently associated with the graph. Among spectral methods, graph-based ranking algorithms are essentially a way of deciding the importance of a vertex in a graph, by taking into account global information drawn from the graph structure. Addressing this issue, [5] and [7] investigated applications of graph theory and complex network principles in the analysis of vulnerability and ranking elements of WSNs.

The aim of this paper is to define importance areas by a method of semi-supervised clustering into WSNs, through ranking nodes. Our proposal uses

---

\*This work has been supported by project IDAWAS, DPI2009-11591, of the Dirección General de Investigación of the Ministerio de Ciencia e Innovación of Spain and the complementary support ACOMP/2011/188, of the Consellería de Educació of the Generalitat Valenciana.

<sup>†</sup>contact e-mail: joagupre@upv.es

spectral-based methods by the algebraical connectivity of the graph Laplacian matrix. In order to develop efficient models from these spectral methods, it will also be useful to adapt the PageRank [1] and HITS [6] algorithms. Both ranking algorithms provide relative importance measures of webpages.

## **2 Adapting the algorithms to WSN**

To adapt the algorithms to WSNs, we considered the physical structure of a WSN as a mathematical graph  $G = (V, E)$  in which the set of all graph vertices,  $V$ , represents tanks, water sources and nodes or junctions (which are the connection between pipes and points of water withdrawal). The set of the graph edges,  $E$ , represents pipes, valves and pumps. By understanding a WSN as a graph of special characteristics, it is possible to abstract the concept of web page looking at it a consumption node in a WSN. Links between pages are now understood as pipes connecting different nodes. Thus, both algorithms can be adapted. About it [3] demonstrated the possibilities of ranking algorithms through the adaptation of the PageRank algorithm to measure the importance of the nodes of a WSN.

## **3 Experimental study**

In order to show the methodology exposed we considered a real case, the WSN of the central area of Celaya, Mexico. This network is made out of 479 pipes and 339 nodes; 5 sources and 1 tank. Its total pipe length is 42.5 km, the node elevation average is 156 metres; and the total consumed flow rate amounts to 91 l/s.

As stated, our aim is to divide the WSN of our case study on importance clusters. To this purpose, we apply the procedure of semi-supervised clustering algorithm explained in [2]. After applying this process to the WSN data, the following results were obtained. In Figure 1 we can observe the division of the WSN into three importance clusters introducing the PageRank and HITS measures. In addition, this figure shows the distribution of the nodes with high values in each case. It is important to note that cluster 2 (Figure 1a) is the zone with high PageRank average. Therefore, we can say that this area is more critical than the others. It is because there are nodes with sensible connectivity and remove them could affect the performance of the

WSN. In the case of HITS algorithm (Figure 1b), the results obtained show that the division of the clusters was not homogeneous, because one of the clusters has more of the 50 percent of the nodes. In Table 1 we can observe a summary of each cluster regarding PageRank and HITS.

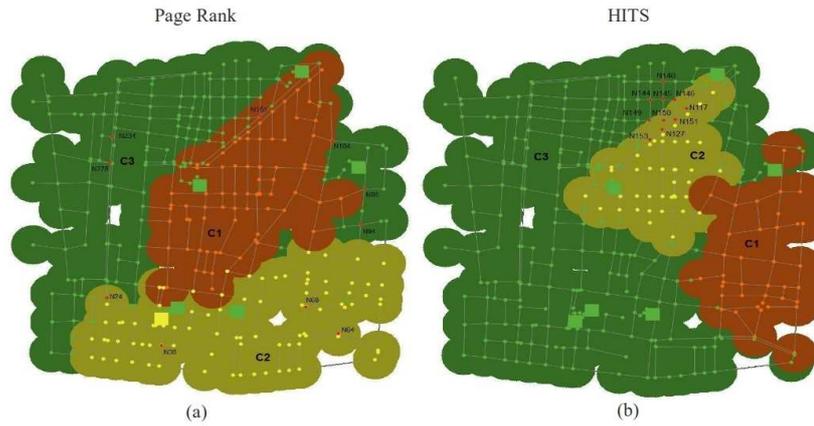


Figure 1: Scheme of the division clusters by PageRank and HITS algorithms.

Table 1: Description of PageRank and HITS clusters.

Cluster	n Nodes	PR avg.	Cluster	n Nodes	HITS avg.
C1	74	0.00298	C1	38	208.1E-7
C2	91	0.00304	C2	41	144.4E-6
C3	174	0.00289	C3	260	0.03243

## 4 Conclusions

In the present work, we propose to introduce PageRank and HITS algorithms as importance measures to form clusters into a WSN by a semi-supervised clustering methodology. This let to know critical nodes into the WSN, and how would they affect the nodes are connected with them.

It has demonstrated that PageRank and HITS algorithms can be an acceptable relative importance measures. These measures help us to understand the structure and performance of the network, identifying vulnerability points.

## References

- [1] Brin S. and Page L., 1998. The Anatomy of a Large-Scale Hypertextual Web Search Engine. <http://infolab.stanford.edu/backrub/google.html>.
- [2] Herrera M. Improving water network management by efficient division into supply clusters. PhD. thesis, Universitat Politècnica de València, Spain, 2011.
- [3] Herrera M., Gutiérrez-Pérez J. A., Izquierdo J., Pérez-García R., 2011. Ajustes en el modelo Page-Rank de Google para el estudio de la importancia relativa de los nodos de la red de abastecimiento. Proceedings, X Seminario Iberoamericano de Planificación, Proyecto y Operación de Sistemas de Abastecimiento de Agua (SEREA). Morelia, México.
- [4] Herrera M., Canu S., Karatzoglou A., Pérez-García R., Izquierdo J., 2010. An approach to water supply clusters by semi-supervised learning. International Environmental Modelling and Software Society (iEMSs) 2010 International Congress on Environmental Modelling and Software.
- [5] Izquierdo J., Montalvo I., Pérez-García R., 2008. Sensitivity analysis to asses the relative importance of pipes in water distribution networks. *Mathematical and Computing Modeling*. 48:268-278.
- [6] Kleinberg S., 1999. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 48:604-632.
- [7] Michaud D. and Apostolakis G. E., 2006. Methodology for ranking the elements of water supply networks. *Journal of Infrastructure Systems*, 12(4):230-242.

# Regularization Algorithm within Two-Parameters for one and two dimensions for the Identification of the Heat Conduction Coefficient in the Parabolic Equation

D. Hinestroza G.<sup>\* \*</sup>, J. Peralta<sup>\*†</sup> and L.E. Olivar<sup>† ‡</sup>

(\*) Universidad del Valle,

Ciudad Universitaria Meléndez Calle 13 No 100-00, Cali, Valle del Cauca, Colombia

(†) Universidad del Tolima,

Barrio Santa Helena, A.A. 546, Ibagué, Colombia

Noviembre 30, 2011

## 1 Introduction

We will consider two identification inverse problem in the parabolic equation in two and one dimensions. In the first case, we have two identify the heat-coefficient  $q(x)$  in the parabolic equation

$$\frac{\partial u}{\partial t} - \nabla \cdot (q(x)\nabla u(x, t)) = f(x, t), \quad (x, t) \in \Omega \times (0, T] \quad (1)$$

where  $\Omega$  is an open bonded domain in  $R^2$  with piecewise smooth boundary  $\partial\Omega$ .  $f$  is the heat source density in  $L_2(0, T; L^2(\Omega))$

Initial-boundary conditions

---

\*e-mail: doris.hinestroza@correounivalle.edu.co

†e-mail: jenifer.peralta@gmail.com

‡e-mail luiseduardo.olivar@hotmail.com

$$u(x, 0) = u_0(x), \quad x \in \Omega \tag{2}$$

and

$$u(x, t) = 0, \quad (x, t) \in \partial\Omega \times [0, T], \tag{3}$$

are known functions.

In the second case we consider the identification problem of finding the heat-coefficient  $a(x)$  in the parabolic equation

$$u_t = \frac{\partial}{\partial x}(a(x)u_x(x, t)) + f(x, t), \quad 0 \leq x \leq 1, \quad 0 \leq t \leq 1 \tag{4}$$

under the initial and boundary conditions

$$u(x, 0) = \varphi(x), \quad 0 \leq x \leq 1, \tag{5}$$

$$u(0, t) = h(t) \quad 0 \leq x \leq 1 \tag{6}$$

$$\partial_x u(1, t) = 0 \quad 0 \leq t \leq 1, \tag{7}$$

where  $\varphi(x)$ ,  $h(t)$  are known  $u(x, t)$  twice differentiable and the coefficient  $a(x)$  is suppose continuous.

We will suppose that the observation data  $u$  and  $f$  is given by

$$u^\varepsilon(x, t) = u(x, t) + noise, \quad t \in (t_1, t_2). \tag{8}$$

$$f^\varepsilon(x, t) = f(x, t) + noise, \quad t \in (t_1, t_2). \tag{9}$$

For the one dimensional problem, to recover  $a$  we considerer temperature measurements

$$g_i^\delta(t) = u_i(t) = u(x_i, t), \quad 0 \leq t \leq 1, \quad i = 1, \dots, n \tag{10}$$

It is well known that the inverse problems [7] related with identification of coefficients are ill posed and regularization technique have to be using. See [1, 3]. A number of important physical identification problems fall within the above framework. This partial differential equation is a basic model equation in oil reservoir simulation and ground water flow structure with permittivity coefficient  $a$  or the heat transfer of material with the conductivity  $a$ . See [5, 6]. In this kind problems the physical law is given, but the information of the parameters are unknown. In real situation is easier to measure the solution  $u$  in several points of its domain that to measure the coefficient  $a$ . Several papers about identification coefficients for the parabolic and elliptic equations have been proposed in [1 – 4].

## 2 Definition of the Functional that Depends of Two Parameters

To stabilize this problem for the two dimensional case, we will propose in this paper a regularization method that depends on two parameters.

Let us define the admissible set of coefficients in the set  $K$  given by

$$K = \{q \in H^1(\Omega) : \alpha_1 \leq q(x) \leq \alpha_2 \text{ in } \Omega\},$$

where  $\alpha_1$  and  $\alpha_2 > 0$ .

Let us define the operator

$$F : H^1(\Omega) \times [0, T] \longrightarrow H^1(\Omega) \times [L^2(\Omega)]^n \tag{11}$$

$$F(u) = \left( \int_{t_1}^{t_2} u(x, t) dt, \int_{t_1}^{t_2} \nabla u(x, t) dt \right) \tag{12}$$

and introduce

$$\Psi : K \subset H^1(\Omega) \longrightarrow H^1(\Omega \times [0, T])$$

In order to stabilize the problem of identification coefficient in the one dimensional case, we define the operators

$$\Psi(q)(x, t) = u(q, x, t), \tag{13}$$

$$F(\Psi(q)) = \left( \int_{t_1}^{t_2} u(q, x, t) dt, \int_{t_1}^{t_2} \nabla u(q, x, t) dt \right) \tag{14}$$

$$\Psi : D(\Psi) \subseteq X \rightarrow Y \tag{15}$$

$$a \rightarrow u$$

$$F : Y \rightarrow Z = \otimes_{i=1}^n L^2[0.1] \tag{16}$$

$$u \rightarrow (u_1, u_2, \dots, u_n, )$$

where  $u$  is the solution of the problem (1)-(4) and  $u_i(t) = u(x_i, t)$ ,  $i = 1, \dots, n$ . Therefore, we have the implicit operator

$$F \circ \varphi(a) = (u_1, u_2, \dots, u_n) = (u_i)_{i=1}^n \tag{17}$$

These two operators are weakly closed defined in the two and one dimensional case are bounded as we will see in the next two lemmas.

**Lema 2.1** *The operator  $F \circ \Psi$  and  $\Psi$  are bounded.*

**Lema 2.2** *The operators  $F \circ \Psi$  and  $\Psi$  are weakly closed.*

The facts that the operators  $\Psi$  and  $F \circ \Psi$  are weak closed and bounded, allow us to apply the results in [2]. To investigate the solutions of the implicit problems ( ) and ( ) we define for  $\alpha, \beta \geq 0$ , the functionals for the two dimensional case:

$$J(q) = M_{\alpha}^{\beta}[q, z, q^*, \Psi^*] = \|F \circ \Psi(q) - z\|^2 + \alpha \|q - q^*\|^2 + \beta \|\Psi(q) - \Psi^*\|^2$$

and consider the minimization problem

$$P_{\alpha}^{\beta} = \min_{q \in K} M_{\alpha}^{\beta}[q, z, q^*, \Psi^*], \tag{18}$$

where

$$z = \left( \int_{t_1}^{t_2} u(x, t) dt, \int_{t_1}^{t_2} \nabla u(x, t) dt \right). \tag{19}$$

For the one dimensional case we have to minimize the functional

$$J(a) = \|F \circ \Psi(a) - g^{\delta}\|_Z^2 + \alpha \|a\|_X^2 + \beta \|\Psi(a)\|_Y^2, \tag{20}$$

To minimize the functionals  $J$  we need to find the first variation of  $J$  and this is the main problem related with these functionals. To find this kind of derivative we will use the sensibility and adjoint equations. Ones we have the derivative of this functional we use the conjugate gradient method to find the approximation solutions. To determine the descend parameter we have to solve a problem of minimization in one dimension. All this calculations are made in the paper that will be consider for publication soon.

### 3 Numerical calculations

Will consider only one example for the one dimensional case. The data  $g$  was generated adding to the exact solution  $u$ , random uniformly distributed numbers.

**Example 3.1** *Let's consider the equation*

$$\frac{\partial}{\partial t}u(x, t) - \frac{\partial}{\partial x}(a(x)u_x(x, t)) = f(x, t)$$

$$u(0, t) = 0, \quad u(1, t) = 0, \quad u(x, 0) = \sin(\pi x)$$

*The exact solutions are :*

$$u = \exp(t) \sin(\pi x), \quad a(x) = ((40x^2)/(40x^2 + 1)) + 1.$$

*We will initially suppose the  $a(x) = 1.7$ . In the table below, we have the best approximation for  $\alpha = \epsilon = 10^{-5}$*

See pictures at the end of the paper.

## 4 Conclusions

The importance of this paper is that we are using a new method using two parameters of regularization, where we define adequate operators and show the problem as an implicit inverse problem which allows us to use the results obtained in [2]. Furthermore, we obtained the first variation of the functional through the sensitivity and adjoint equations. In the future, we will define a numerical scheme for numerically identifying the parameter in the parabolic initial boundary-value for the two dimension problem. The present work is accompanied with the support of Instituto Universitario de Matemática Multidisciplinar, Universidad Politécnica de Valencia, España and the Universidad del Valle in Cali, Colombia.

## References

- [1] .W. Engel and J. Zou, A new approach to convergence rate analysis of Tikhonov regularization for parameter identification in heat conduction, *Inverse Problems* 16:1907-1923, 2000
- [2] . Hinestroza, D. A. Murio and S. Zhan Regularization Techniques for nonlinear Problems. *Computers and Mathematics with Applications* 37:145-159, 1999.

[3] . Hinestroza and D.A. Murio Identification of transmissivity coefficients by mollification techniques. Part I: one-dimensional Elliptic and Parabolic Problems, *Computers Math. Applic.* 25: 59-79, 1993.

[4] .Kunish and G. Peichel, Estimation of a temporally and spatially varying diffusion coefficient in a parabolic system by an augmented Langrangian techniques, *Numer. Math.* 59:473-509, 1991.

[5] .W.-G.Yeh, Review of parameter identification procedures in groundwater hydrology: The inverse problem, *Water Resources Research* 22:95-108 1986.

[6] A. Kirschner, An Introduction to the Mathematical Theory of Inverse Problems.Second Edition. New York, Springer, 2011.

$\alpha$	$\epsilon$	0	$10^{10}$	$10^9$	$10^8$	$10^7$	$10^6$	$10^5$	$10^4$	$10^3$	$10^2$	$10^1$
0		0.4698	0.4303	0.4456	0.4360	0.4344	0.4520	0.4629	0.4493	0.5698	0.7264	0.8508
$10^{10}$		0.4403	0.4588	0.4396	0.4306	0.4376	0.4687	0.4526	0.4936	0.5735	0.7128	0.8514
$10^9$		0.4481	0.4324	0.4333	0.4266	0.4393	0.4274	0.4437	0.4487	0.5496	0.7205	0.8311
$10^8$		0.4396	0.4444	0.4903	0.4359	0.4538	0.4444	0.4406	0.4340	0.5741	0.7154	0.8568
$10^7$		0.4482	0.4245	0.4402	0.4848	0.4572	0.3970	0.4395	0.4329	0.5998	0.7207	0.8258
$10^6$		0.3678	0.3435	0.3773	0.3235	0.3863	0.3694	0.3683	0.3405	0.5595	0.7167	0.8247
$10^5$		0.1049	0.1113	0.1208	0.1006	0.1232	0.0893	0.1197	0.0977	0.2871	0.7052	0.8410
$10^4$		0.1769	0.1763	0.1733	0.1752	0.1763	0.1712	0.1770	0.1761	0.1679	0.4325	0.7910
$10^3$		0.2435	0.2362	0.2365	0.2363	0.2360	0.2372	0.2366	0.2347	0.2369	0.2366	0.2689
$10^2$		0.7218	0.7215	0.7219	0.7217	0.7218	0.7214	0.7218	0.7216	0.7226	0.6381	0.3254
$10^1$		1.1271	1.1272	1.1271	1.1272	1.1272	0.9918	1.1211	1.1273	1.1269	1.0761	0.8411

Table 2.

Figure 1:  $L^2 - Error$  example for  $\delta = 0.02$

,

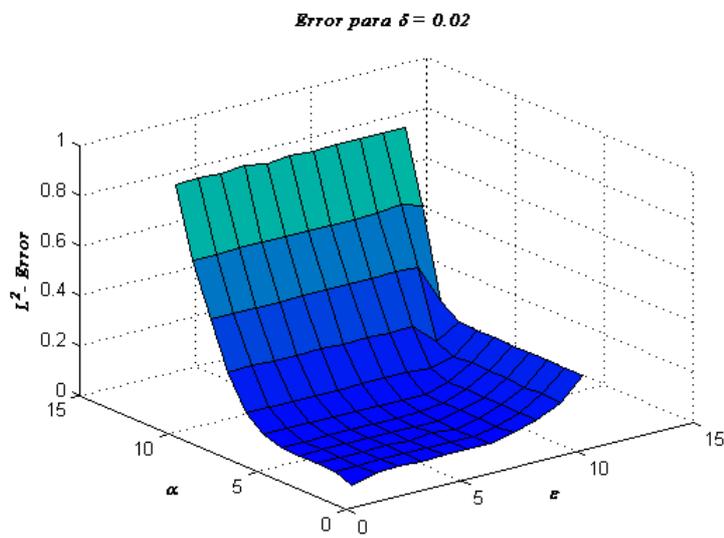


Figure 2: Surface of the error for  $\delta = 0.02$ , example

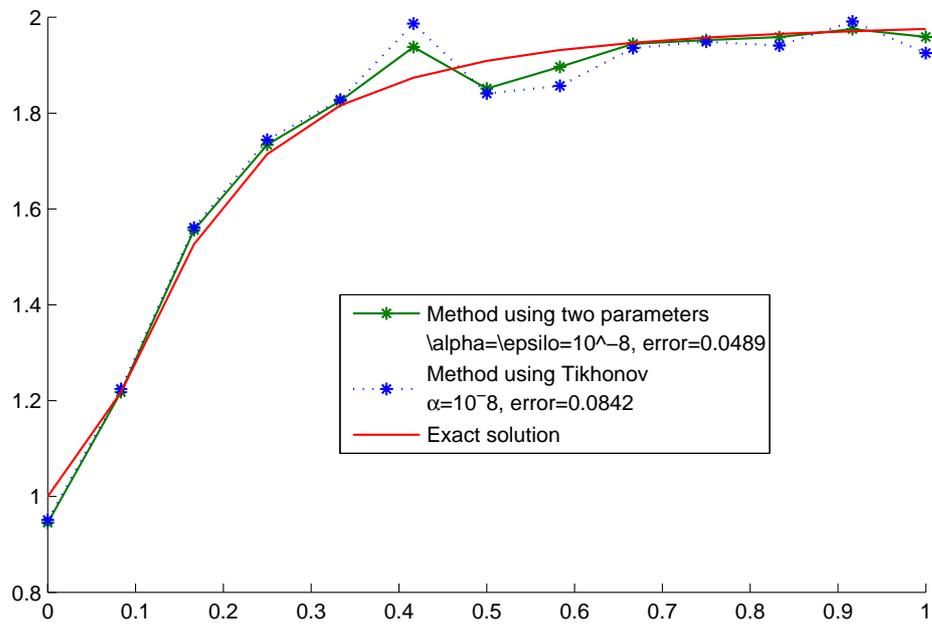


Figure 3: comparison between Tikhonov and the Method using two parameters

# Evaluation of the Eulerian-Lagrangian Spray Atomization (ELSA) model in spray simulations: 2D cases

S.Hoyas<sup>\*\*</sup>; Xandra Margot<sup>\*</sup>, Antonio Gil<sup>\*</sup>, Dung Khuong-Anh<sup>\*</sup>  
and Frederic Ravet<sup>†</sup>

(<sup>\*</sup>) CMT-Motores Térmicos, Universitat Politècnica de València,  
Valencia 46022, Spain,

(<sup>†</sup>) Renault, 1 Avenue du Golf 78288, Guyancourt, France

November 30, 2011

During the last fifteen years Computational Fluid Dynamics (CFD) has become one of the most important tools to both understand and improve the Diesel spray development in Internal Combustion Engine (ICE). Fuel injection process and subsequent fuel-air mixing formation play a major role on combustion and pollutant emissions in ICE. Still now, some of the processes taking part in this phenomena, such as primary atomization or nozzle cavitation, are not fully understood [1, 2, 3], even taking into account that the full set of equations describing the physical situation are known, [4], [5].

The goal of the Eulerian-Lagrangian Spray Atomization (ELSA) model is to realistically describe the dense zone of the spray and its atomization. Since the seminal work of Vallet et al. [6] it has been under development by several authors, including Blokkeel et al. [7], Beau [8] or Lebas [9]. ELSA model takes advantages of the Eulerian description of the near nozzle flow where some assumptions of standard spray models based on discrete droplet method (DDM) shows strong limitations. DDM approach is valid only when the liquid volume fraction is small inside the computational cells and when the drops are homogeneously distributed in the computational space, nei-

---

\*e-mail: serhocal@mot.upv.es

ther of them is satisfied in the near field of the spray. Therefore, in order to keep a low void fraction and assure numerical stability, it is necessary to use grid sizes larger than orifice diameter, which cannot adequately resolve the flow structures in this region by means of DDM. Additionally, it is also not required to assume any particular shape to represent drops and liquid ligaments on ELSA model, where the average area of the liquid-gas interface is introduced as a measure of the atomization extent. Moreover, the DDM method applies isolated drop based models in this region with strong interaction within the liquid phase, where its validity is hardly justified.

The main hypothesis of ELSA, is that the flow must be a high-speed turbulent spray, where Reynolds bulk number and Weber number should be high (See the PhD thesis of Beau [8] for a detailed study). Reynolds number, should be greater than 1500 [6], and Weber number,  $We > 350$  [10, 11]. Values of these numbers in actual Diesel injector are far greater than this threshold. The second hypothesis is that the turbulent mixing process between the liquid and surrounding gaseous phase is simulated as a single-phase turbulent fluid flow with mean properties. From this second hypothesis comes the main concern about ELSA model, which is that it does not give a detailed information about both phases separately in the near nozzle region.

According to the previous statements, the purpose of the present study is to perform a 2D preliminary validation of the ELSA spray model implemented in the Star-CD code by Renault. This work is part of a more ambitious project, with the general objective of developing and validating this spray model implementation for real-life applications on CFD engine calculations, including cavitating nozzles. There also exists a growing interest in this sort of nozzles [12], due to the general perception that cavitation can increase the turbulence level, thus improving the mixture of the spray and its combustion. As DDM cannot resolve the near nozzle region, ELSA which effectively resolves the vicinities of the nozzle, can be the right tool in order to couple the internal flow, and the spray.

In order to perform this 2D validation of the ELSA model, data obtained from an injection experimental test facility [13] are used. The geometry reproduces a chamber of  $80 \times 25mm$ . In this chamber, one diesel spray develops, coming from a non-cavitating single-hole injector (tapered nozzle), with an outlet diameter of  $112\mu m$ . An inert gas is considered to be enclosed in the chamber at an initial ambient pressure ( $P_A$ ) and temperature ( $T_A$ ) of  $3.53MPa$  and  $293K$  respectively. Fuel injection pressure is  $P_i = 80MPa$  and

the fuel density is  $\rho = 822\text{kg/m}^3$

The simulations are actually a 3D simulation with only a cell in the azimuthal direction, and modeling a  $5^\circ$  sector of the spray. This is shown in Figure 3, where is also depicted the boundary conditions imposed on the spray and chamber. The key zone in this sort of simulations is the vicinities of the nozzle, where the mesh size has to be small enough to capture the spray structure and droplets. The criterion used in this paper is to define the size of the first cell and then extrude the mesh, fixing the axial and radial ratio. The last to first ratio,  $R = (l_1/l_n)^{-1/(n-1)}$  is fixed in all cases to 0.006. In this formula,  $l_1, l_n$  are the length of the first and last cell, and  $n$  is the number of cells in the radial direction.

A typical calculation of a Diesel Injection, around  $t_{inj} = 2\text{ms}$ , using the ELSA model takes approximately one week on 8-12 last generation cores. So, it is crucial to choose the mesh as coarser as possible to reduce the number of computational time, as we need to validate many different factors afterward.

Selecting the right set-up and appropriate method have been done to determine the relative importance of few parameters, and the best case which we suggest to use in this article results very similar in profile to results obtained from Diesel spray experiments. The coarser mesh is able to describe accurately penetration and axial velocity, with a corrected value of the  $k - \varepsilon$  constant  $C_{\varepsilon 1} = 1.60$ . Apart from this, droplet generation and spray angle and shape show that ELSA model can describe accurately these features.

There are three major drawbacks to continue the validation. Firstly, it is necessary to increase the temperature at the chamber in order to take into account the evaporation of Diesel droplets. Secondly, 2D cases are especial simulations, and do not reflect the 3D structure and development of real life sprays. Last, but no least by any means, is it necessary to consider cavitating nozzles, as it seems that cavitation will be present in nozzles now and in the near future.

## References

- [1] A.H. Lefebvre, Atomization and sprays, Taylor and Francis, 1989.
- [2] R. Lebas, G. Blokkeel, P. Beau, F. Demoulin, Coupling Vaporization Model With the Eulerian-Lagrangian Spray Atomization (ELSA) Model in Diesel Engine Conditions, SAE Technical Paper 2005-01-0213, 2005.

- [3] R. Lebas, T. Menard, P.A. Beau, A. Berlemont, F.X. Demoulin, Numerical simulation of primary break-up and atomization: DNS and modelling study, *Int. J. Multiphase Flow* 35 (2009) 247260.
- [4] M. Ishii, *Thermo-fluid dynamic theory of two-phase flow*, 2nd edition, Springer, New York, 2011.
- [5] I. Kataoka, Local instant formulation of two-phase flow, *Int. J. Multiphase Flow*, 12 (1986) 745-758.
- [6] A. Vallet, A.A. Burluka, R. Borghi, Development of a Eulerian model for the atomization of a liquid jet, *Atom. and sprays*, 11 2001 619-642.
- [7] G. Blokkeel, B. Barbeau, R. Borghi, A 3D Eulerian Model to Improve the Primary Breakup of Atomizing Jet, *SAE Technical Paper 2003-01-0005*.
- [8] P.A. Beau, *Modelisation de l'atomisation d'un jet liquide - Application aux sprays diesel*, Ph.D. Thesis, University of Rouen. 2006
- [9] R Lebas, *Modelisation Eulerienne de l'Atomisation haute pression - Influences sur la vaporisation et la combustion induite*, Ph.D. Thesis, University of Rouen, 2007
- [10] S. Lee, R. Reitz, Effect of liquid properties on the breakup mechanism of high-speed liquid drops, *Atom. Sprays*, 11 (2001) 1-19.
- [11] F. Tanner, Development and validation of a cascade atomization and drop breakup model for highvelocity dense sprays, *Atom. sprays*, 14 (2004) 20-32.
- [12] F.J. Salvador, S. Hoyas, R. Novella, J. Martinez-Lopez, Numerical simulation and extended validation of two-phase compressible flow in diesel injector nozzles. *Proc. Inst. Mech. Eng. Part D-J. Automob. Eng.* 225 (2011) D4:545-563.
- [13] R. Payri, J.M. García, F.J. Salvador, J. Gimeno, Using spray momentum flux measurements to understand the influence of Diesel nozzle geometry on spray characteristics, *Fuel*, 84 (2005) 551-561.

# Optimal methods with approximated derivatives for nonlinear equations

A. Cordero\* , J.L. Hueso\*\*; E. Martínez†, and J.R. Torregrosa\*

(\*) Instituto Universitario de Matemática Multidisciplinar

(†) Instituto de Matemática Pura y Aplicada

Universitat Politècnica de València,

Cno. de Vera, 14, 46022 València, Spain.

November 30, 2011

## 1 Introduction

In this work we show a general procedure to obtain optimal derivative free iterative methods [3] for nonlinear equations  $f(x) = 0$ , applying polynomial interpolation to a generic optimal derivative free iterative method of lower order.

Let us consider an optimal method of order  $q = 2^{n-1}$ ,  $v = \phi_n(x)$ , that uses  $n$  functional evaluations. Performing a Newton step  $w = v - \frac{f(v)}{f'(v)}$  one obtains a method of order  $2^n$ , that is not optimal because it introduces two new functional evaluations. Instead, we approximate the derivative by using a polynomial of degree  $n$  that interpolates  $n + 1$  already known functional values and keeps the order  $2^n$ .

We have applied this idea to Steffensen's method, [4], obtaining a family of optimal derivative free iterative methods of arbitrary high order.

Different numerical tests confirm the theoretical results. We also compare the new family with other well-known family of similar characteristics.

---

\*e-mail:jlhueso@mat.upv.es

## 2 Optimal methods of increasing order

A variety of problems in different fields of science and technology require to find the solution of a nonlinear equation. Iterative methods for approximating solutions of nonlinear equation are the most used technique. The efficiency index, introduced by Ostrowski in [2], establishes the effectiveness of the iterative method. In this sense, Kung and Traub conjectured in [3] that a method is optimal if it reaches an order of convergence  $q = 2^n$ , using only  $n + 1$  functional evaluations per step. Nowadays the interest of iterative methods is their optimality. We can cite, among other, Newton and Steffensen's methods [4], optimal for  $n = 2$ , Ostrowski and Jarrat's methods, [2, 5], optimal for  $n = 3$ . Some optimal eighth order methods have been proposed (see for example [6, 7, 8, 9, 10, 11]), even optimal sixteenth order methods have been published such as [1].

General procedures to obtain families of optimal multipoint iterative methods for every  $n$  were given in [3, 12].

Newton's and Steffensen's iterative methods for the solution of the nonlinear equation  $f(x) = 0$  are both of second order and they are optimal, because they only need 2 functional evaluations per step. Steffensen's method has the advantage that it does not need the derivative. In this paper we focus on derivative free iterative methods, but our procedure could also be applied starting from Newton's method.

Our idea for getting optimal methods of increasing order is to compose a Steffensen's iteration,

$$\begin{aligned} y_0 &= x_k \\ y_1 &= y_0 + f(y_0) \\ y_2 &= y_0 - \frac{f(y_0)^2}{f(y_1) - f(y_0)}, \end{aligned}$$

with a Newton's step

$$y_3 = y_2 - \frac{f(y_2)}{f'(y_2)}.$$

The resulting iteration has convergence order 4, being the composition of two second order methods (see [13], theorem 2.4), but the method is not optimal because it uses 4 functional evaluations.

In order to get optimality, the value of  $f'(y_2)$  can be approximated by the derivative  $p'_2(y_2)$  of a suitable second degree polynomial obtained from

already computed functional values. It is natural to consider the interpolating polynomial of the points  $(y_0, f(y_0))$ ,  $(y_1, f(y_1))$  and  $(y_2, f(y_2))$ . This polynomial can be written as

$$p_2(t) = a_0^{(2)} + a_1^{(2)}(t - y_2) + a_2^{(2)}(t - y_2)^2,$$

and its coefficients are the solutions of a linear system

$$p_2(y_j) = f(y_j), \quad j = 0, 1, 2.$$

Therefore, we get a method,  $M_4$ , with  $n + 1 = 3$  functional evaluations that will be optimal provided its order is  $2^n = 4$ .

$$y_0 = x_k \tag{1}$$

$$y_1 = y_0 + f(y_0) \tag{2}$$

$$y_2 = y_0 - \frac{f(y_0)^2}{f(y_1) - f(y_0)} \tag{3}$$

$$x_{k+1} = y_3 = y_2 - \frac{f(y_2)}{p_2'(y_2)}. \tag{4}$$

We can improve  $M_4$  by the same procedure. Composing it with a Newton's step one obtains an order 8 method but loses optimality because there are two new functional evaluations. Then, we substitute the derivative of function  $f$  by the derivative  $p_3'(y_3)$  of the third degree polynomial that interpolates  $f$  at the points  $y_0, y_1, y_2, y_3$ . The order is maintained whereas the method becomes optimal.

In theory, this procedure can be indefinitely repeated, giving a family of optimal methods of arbitrarily high order. The iteration of the generic method  $M_q$ , for  $q = 2^n$  is  $x_{k+1} = y_{n+1}$  where

$$y_0 = x_k \tag{5}$$

$$y_1 = y_0 + f(y_0) \tag{6}$$

$$y_{j+1} = y_j - \frac{f(y_j)}{p_j'(y_j)}, \quad j = 1, 2, \dots, n, \tag{7}$$

and  $p_j$  is the polynomial that interpolates  $f$  in  $y_0, y_1, \dots, y_j$ .

We are able to prove the following result.

**Theorem 1** *Let  $\alpha \in I$  be a simple zero of a sufficiently differentiable function  $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$  in an open interval  $I$ . If  $x_0$  is sufficiently close to  $\alpha$ , then the method  $M_{2^n}$  defined by (5-7) has optimal convergence order  $2^n$ .*

We have checked the effectiveness of the new optimal iterative methods  $M_q$  comparing them with the optimal family  $K_q$ , for  $q = 2^n = 4, 8, 16$  introduced in [3],

$$\begin{aligned}y_0 &= x_k \\y_1 &= y_0 + \beta f(y_0), \quad \beta \neq 0, \\y_{j+1} &= Q_j(0), \quad j = 1, 2, \dots, n,\end{aligned}$$

where  $Q_j(t)$  is a polynomial of degree at most  $j$  satisfying the interpolation conditions  $Q_j(f(y_i)) = y_i, \quad i = 0, 1, \dots, j$ .

The numerical results show that the new family has a slightly better performance than the classical one, so it can be competitive.

## References

- [1] Y.H.Geum, Y.I. Kim, *A biparametric family of optimally convergent sixteenth-order multipoint methods with their fourth-step weighting function as a sum of a rational and a generic two-variable function*, Journal of Computational and Applied Mathematics, 235 (2011) 3178-3188.
- [2] A.M. Ostrowski, *Solutions of equations and systems of equations*, Academic Press, New York-London, 1966.
- [3] H.T. Kung, J.F. Traub, *Optimal order of one-point and multi-point iteration*, Journal of the Association for Computing Machinery, 21 (1974) 643-651.
- [4] J.M. Ortega, W.C. Rheinbolt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.
- [5] P. Jarrat, *Some fourth order multipoint iterative methods for solving equations*, Math. Comp, 20 (1966) 434-437.
- [6] W. Bi, H. Ren, Q. Wu, *Three-step iterative methods with eighth-order convergence for solving nonlinear equations*, Journal of Computational and Applied Mathematics 255 (2009) 105-112.

- [7] J.R. Sharma, R. Sharma, *A family of modified Ostrowski's methods with accelerated eighth order convergence*, Numerical Algorithms, 54 (2010) 445-458.
- [8] W. Bi, Q. Wu, H. Ren, *A new family of eighth-order iterative methods for solving nonlinear equations*, Applied Mathematics and Computation, 214 (2009) 236-245.
- [9] R. Thukral, M.S. Petkovic, *A family of three-point methods of optimal order for solving nonlinear equations*, Journal of Computational and Applied Mathematics, 233 (2010) 2278-2284
- [10] L. Liu, X. Wang *Eighth-order methods with high efficiency index for solving nonlinear equations*, Applied Mathematics and Computation 215 (2010) 3449-3454.
- [11] A. Cordero, J.L. Hueso, E. Martínez, J.R. Torregrosa, *New modifications of Potra-Pták's method with optimal fourth and eighth order of convergence*, Journal of Computational and Applied Mathematics, 234 (2010) 2969–2976.
- [12] B. Neta, *On a family of multipoint methods for non-linear equations*, International Journal for Computation and Mathematics, 9 (1981) 353-361.
- [13] J.F. Traub, *Iterative methods for the solution of nonlinear equations*, Prentice Hall, New York, 1964.

# Normal goniometric values to guide decision-making in lower-extremity rotational problems using support vector machine techniques\*

J.M. Bru-Juanes<sup>‡</sup>, J.M. Bru-Lázaro<sup>‡</sup>, M. Herrera<sup>◊</sup>, and J. Izquierdo<sup>◊†</sup>

(‡) Centro de Podología y Posturología, Valencia, Spain

(◊) Fluing-IMM, Universitat Politècnica de València, Valencia, Spain

November 30, 2011

## 1 Introduction

Torsional analysis of the lower extremities has become an integral part of the decision-making process when treating neuromuscular problems. Abnormal gait or torsional misalignments may be risk factors for degenerative pathological effects on joints of the lower extremities [1, 2]. Identification of the alterations and precise knowledge about the extent a given extremity is within or outside of some normal parameter value is a difficult task. Solid knowledge of the normal development of torsional relationships is essential for treating musculoskeletal problems, and normal references, that is to say, objective and quantitative standard measurements are of crucial importance for comparison prior to making a decision.

In the 1980s, Staheli [3, 4] provided a great deal of information on normal ranges of lower limb rotation. The aim of this paper is to update normal

---

\*Work supported by project IDAWAS, DPI2009-11591, of the Dirección General de Investigación of the Ministerio de Ciencia e Innovación of Spain and the complementary support ACOMP/2011/188 of the Consellería de Educació of the Generalitat Valenciana.

†contact e-mail: jizquier@upv.es

goniometric values that could be used as a reference for clinical analysis and to discover unknown compensating mechanisms between the involved segments. We focus on a non-parametric clustering methodology based on support vectors (SVC). The process of clustering is based on kernel methods that avoid explicit calculations in high-dimensional feature spaces and make clustering more efficient. In addition, our proposal has the computational advantage of relying on an SVM quadratic optimisation that reaches global solutions.

## **2 Materials and method**

The database has been compiled by the same clinician (JMBJ) during more than 15 years for the whole series of measurements. We consequently claim that the bias of the data is negligible. Moreover, the same goniometer was used in all cases. Records collect data from individuals with ages ranging from 2 to 60 years. Data was obtained by following the Stanley Hoppenfeld protocol [7] to measure both internal and external femoral and tibial rotations.

Support vector clustering (SVC) was introduced by Ben-Hur et al. [8] to cluster the data set based on the theory of support vector machine (SVM). SVM apply simple linear methods to the data in a high-dimensional feature space non-linearly related to the input space by a kernel function. One essential property of these kernel functions is that they can work with the data on this high-dimensional space without needing explicit computation. This so-called 'kernel-trick' adds simplicity to the SVM approaches, which combined with their adequate performance on a wide range of real-world learning problems (such as classification and regression) has added to their popularity. SVC expands SVM to consider the problem of clustering. In the SVC algorithm, data points are usually mapped from the data space to a highly dimensional feature space using a Gaussian kernel.

The R language [9] has been used in this study.

## **3 Results**

According to their two principal components in an SVM clustering process, three different clusters are clearly obtained. The first principal component

is explained by the age, while the second is inversely proportional to the femoral rotation. The first cluster gathers 93% of the registers, corresponds to normality, and is clearly related to high values of both left and right tibial external torsion.

In all cases the measured variables decrease with age, more rapidly during the early years of life and more steadily afterwards. Other interesting aspects can be observed, however. For example, the linear relationship between left and right rotations expresses the symmetry of individuals. Also, the compensating effect between external tibial and internal femoral rotations is clearly observed. Even though a decreasing trend is observed for all individuals for all variables, it is worth noting that values are different for males and females.

## **4 Discussion**

Our results confirm the decrease of rotations as a function of the age as in all the previous works [3, 5, 4, 6, 10]. During childhood, the decrease of internal femoral rotations is accompanied by a compensating inverse decrease of external tibial rotations. This compensating effect, which needs further characterization, is clearly supported by our data.

These data question the old conclusion by Staheli et al. [3] related to a persistent external tibial rotation. According to our graphs it can be easily observed that left and right tibial rotations do not suffer significant variation with age. Finally, rotations in both male and female follow a decreasing trend; it is greater for females than for males, the difference being on average of  $7^\circ$  at the age of 8, and  $5^\circ$  when adult.

## **5 Conclusions**

To the best of our knowledge there is no study regarding the subject this paper addresses using SVM and, at the same time, using a database with such a systematic and coherent volume of data that measures all the levels of the lower extremity. This makes it possible to develop a method for enhancing the applicability of our results. The method is based on a process of support vector data description (SVDD) for each cluster and this process enables our outlined targets to be achieved. The final accuracy of our proposed

method outperforms traditional statistical methods, and it may be used as a supplementary method for techniques such as computerised tomography.

## References

- [1] D. Goutallier, J.M. Garabedian, A. De Ladoucette, J. Bernageaue, The incidence of femoral and tibial torsion in the development of compartmental osteoarthritis, *J Pediatr Orthop* 6 (1997) 484-488.
- [2] M. Maussa, Rotational malalignment and femoral torsion in osteoarthritic knees with patellofemoral joint involvement. A CT scan study, *Clin Orthop Relat Res* 304 (1994) 176-83.
- [3] L.T. Staheli, M. Corbett, W. Craig, H. King, Lower-extremity rotational problems in children. Normal values to guide management, *J Bone Joint Surg Am* 67 (1985) 39-47.
- [4] L.T. Staheli, Torsion-treatment indications, *Clin Orthop* 247 (1989) 61-66.
- [5] W.B. Greene, J.D. Heckman, *Clínica del movimiento articular*, Edika Med. Barcelona, 1997.
- [6] Y.H. Li, J.S.Y. Leong, In-toeing gait in children, *Hong Kong Med J*. 5(4) (1999) 360.
- [7] S. Hoppenfeld, P. deBoer, R. Buckley, *Surgical Exposures in Orthopaedics: The Anatomic Approach* (Hoppenfeld, *Surgical Exposures in Orthopaedics*), Lippincott Williams & Wilkins; Fourth edition, 2009.
- [8] A. Ben-Hur, D. Horn, H.T. Siegelmann, V. Vapnik, Support Vector Clustering. *Journal of Machine Learning Research* 2 (2001) 125-137.
- [9] R Development Core Team, 2011. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. <http://www.R-project.org>
- [10] K. Hatayama, M. Terauchi, H. Higuchi, S. Yanagisawa, K. Saito, K. Takagishi, Relationship between Femoral Component Rotation and Total Knee Flexion Gap Balance on Modified Axial Radiographs, *J Arthroplasty* 26(4) (2011) 649-53.

# Parallelization of the Finite-Difference Time-Domain Method for Room Acoustics Modelling based on CUDA

José J. López<sup>†</sup>\*, Diego Carnicero<sup>†</sup>, Nestor Ferrando<sup>‡</sup>, José Escolano<sup>†</sup>

(<sup>†</sup>) Instituto de Telecomunicaciones y Aplicaciones Multimedia,

Universidad Politécnica de Valencia, Camino de Vera s/n, 46022 Valencia, Spain

(<sup>‡</sup>) Instituto de Instrumentación para Imagen Molecular,

Universidad Politécnica de Valencia, Camino de Vera s/n, 46022 Valencia, Spain

November 30, 2011

## 1 Introduction

During last years General Purpose Computing on Graphics Processing Units (GPGPU) have experienced a fast evolution, allowing inherently parallel algorithms to experience important performance gains. For an affordable cost, a wide variety of applications have benefited from this evolution. The flagship of GPGPU has been the CUDA architecture. CUDA lets the programmer access and use a variety of resources from the GPU and an easy way of porting serial codes to parallel codes without much effort. However, although naively ported codes can perform much faster than its serial versions, careful usage of the GPU hardware can increase the speed by one or two orders of magnitude compared to the serial code.

It has not been until recently, with the CUDA architecture, that investigation has been carried on room acoustic modelling. The latest work is presented in [1], where the performance of different implementation approaches when working with double precision model data is discussed.

---

\*e-mail: jjlopez@dcom.upv.es

This study presents a comparison between six representative approaches to FDTD parallelization divided into two main categories applied to room acoustics modelling to analyze which one provides the best performance, as well as a novel approach to model data access and storage.

## 2 The FDTD method for acoustics

Finite-Difference Time-Domain (FDTD) method is arguably the most popular numerical method for the solution of problems in electromagnetism and acoustics. Although the FDTD method has existed for over 40 years, its popularity continues to grow as computing costs keep on declining. The method consist on dividing a volume (the room interior in the case of room acoustics) into small cubic cells, whose size determines the maximum frequency to be simulated, and apply the wave propagation equation of pressure to each cell. Then the difference equation computes the next time step of pressure as a function of the pressure in the neighbor cells in the two previous time steps. More details about the FDTD method can be found in [2].

## 3 The CUDA architecture

From the perspective of CUDA, a GPU is abstracted as a device that contains a set of memories and an array of independent Streaming Multiprocessors (SMs), with each SM being able to execute a large number of threads simultaneously.

The threads, organized in blocks, and further clustered in a flat hierarchical structure called *grid*, are sent to the cores for execution. Inside each core, on-chip *shared memory* allows fast communication between the threads of a block. In addition, the GPU has a *global memory* that enables persistent data storage along application lifetime and a (slower) global scope communication between different thread blocks. Finally, each core has fast read access to a *texture memory cache*, optimized for 2D spatial locality.

Two hardware architectures have been evaluated in this study: Tesla [4] and Fermi [5]. For further information on the CUDA architecture refer to [3].

## 4 3D FDTD Implementations

The CUDA threading model offers great flexibility when working with 2D data thanks to its hierarchical nature as stated in the previous section, unfortunately it imposes certain limitations with 3D data since thread blocks can be three-dimensional but grids can only be two-dimensional. Therefore it is necessary to explore different possibilities on how to work with 3D data efficiently, applied in this paper to the simulation of room acoustics through the FDTD modelling technique. This paper explores two techniques: the *tiling* method [1] and the *slicing* method.

The tiling method consists in flattening the volume into a tiled plane where each tile corresponds to a 2D slice. This method maximizes parallel execution by assigning an execution thread to every data point and keeping the streaming multiprocessors busy, but at the cost of a higher computation and memory access count for every model node. The recent Fermi architecture can cope with the additional computational cost, but the fact that these operations need to be performed by every single thread impose an important performance penalty in previous hardware architectures.

The slicing method is presented in [6], and consists in slicing the volume along its *z-dimension*. This way a flat *threadfront* sweeps the volume from bottom to top every time step, with every single thread in the *threadfront* processing a given column of nodes. This technique allows for data reuse, because each thread only has to read five data values from the slow global memory, as the remaining two have already been read in the previous time steps.

Besides the described techniques, several additional improvements can be added to the kernel codes. The first one is the usage of an *in-place* approach to memory access, that is, the usage of only two memory structures to store model data instead of the three, increasing the memory usage and increasing cache efficiency. The second one is the usage of rectangular thread block dimensions instead of the typical square shape, to increase the thread execution efficiency. And the last improvement consist in a technique to effectively process arbitrarily shaped rooms. It is based on two main points: storage of boundary data in the same data structure as air pressure data, and computation of the FDTD only when useful data is present.

## 5 Conclusions

In this study we perform an exhaustive study on how different processing strategies behave in respect to the calculation speed of the FDTD algorithm applied to room acoustic simulation, providing at the same time several implementation guidelines to improve the performance of similar algorithms.

From the six developed implementations, it has been demonstrated that the sliced processing method is the fastest in both hardware architectures tested. It has been proved the usage of texture cache provides the best results when using the Tesla architecture. And, as for the Fermi architecture, the usage of shared memory offers the best performance when processing single or double precision floating point model data. Moreover, speedups of two orders of magnitude has been obtained over optimized serial implementations of the algorithm and almost a two-fold performance increase over recently published parallel implementations.

## References

- [1] C. Webb, S. Bilbao, Virtual Room Acoustics : A Comparison of techniques for computing 3D-FDTD schemes using CUDA, in 130th Audio Engineering Society Convention. London, UK, May 2011.
- [2] D. Botteldooren, Finite-difference time-domain simulation in a quasi-Cartesian grid, *J. Acoust. Soc. Am.* 95 (5) (1994), pp. 2313-2319.
- [3] Nvidia Corp, CUDA C Programming Guide v3.2,  
[http://developer.download.nvidia.com/compute/cuda/3\\_2\\_prod/toolkit/docs/CUDA\\_C\\_Programming\\_Guide.pdf](http://developer.download.nvidia.com/compute/cuda/3_2_prod/toolkit/docs/CUDA_C_Programming_Guide.pdf)
- [4] E. Lindholm , J. Nickolls , S. Oberman, J. Montrym, NVIDIA Tesla: A unified graphics and computing architecture, *IEEE Micro* 28 (2) (2008), pp. 39-55.
- [5] N. Leischner, V. Osipov, P. Sanders. Fermi Architecture White Paper, 2009.
- [6] P. Micikevicius, 3D finite-difference computation on GPUs using CUDA, in GPGPU-2: Proc. of the 2nd Workshop on GPGPU, Washington, DC, USA, 2009, pp. 79-84.

# Singular value decomposition and Tikhonov regularization in hyperbolic passive location

I. A. Mantilla-Gaviria<sup>\*</sup>, M. Leonardi<sup>†</sup>,  
J. V. Balbastre-Tejedor<sup>\*</sup>, and E. de los Reyes Davó<sup>\*</sup>

(<sup>\*</sup>) Instituto ITACA - Universidad Politecnica de Valencia

Camino de Vera s/n - 46022 - Valencia (Spain)

(<sup>†</sup>) DISP - Tor Vergata University

Via del Politecnico 1 - 00131 - Rome (Italy)

November 30, 2011

## 1 Introduction

The hyperbolic passive location systems are a powerful surveillance tools for air traffic control and airport management (taxiing, taking off / landing, approach or enroute) and in this field they are based on the reception, by a number of stations, of signals emitted by a standard on-board device (called "transponder"): the Secondary Surveillance Radar (SSR) Selective Mode or Mode-S signals. These location systems are intended to inform air traffic controllers the location and identification of aircrafts or vehicles equipped with an operational SSR transponder [1], and perform the localization calculations by solving a system of hyperbolic equations based on TDOA (Time Difference of Arrival) technique; the algorithms for solving them run at real time in a central computer. The unknown target position (i.e. of the transponder antenna) is non-linearly related to the measurements; the location of the target usually calls for iterative solution techniques where the equations are linearized and the first iteration is based on a suitable "guess". In

---

\*e-mail: imantilla@itaca.upv.es

some scenarios, especially when the target is outside the perimeter of a set of stations or when its height is near to those of the stations, it is common to find a problem; i.e., the system of equations is ill-posed [2]. The consequence of this is that, when the system of equations is solved, the solution (target position) has a big error or may diverge [3]. As it is shown in [4], this problem has negative effects in the hyperbolic passive location process that can significantly degrade the system accuracy.

## 2 Ill-posed hyperbolic passive location and its solution by SVD and Tikhonov regularization

In hyperbolic passive location, every station measures the Time of Arrival (TOA) of a signal emitted by the target (aircraft or vehicle), then each one sends this value to a central computer where the target position is calculated. This calculation is achieved by the Time Difference of Arrival (TDOA) technique, where the intersection of multiples hyperboloids, which have been created with the relative time differences, are computed. Mathematically, each of these hyperbolas can be expressed in the following form,

$$TDOA = \frac{\sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2} - \sqrt{(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2}}{c} \quad (1)$$

where  $c$  is the speed of light,  $(x, y, z)$  the unknown target position and  $(x_i, y_i, z_i)$  is the known position of the  $i$ th station ( $i = 1$  denotes a reference station). In this sense, a set of  $N_s - 1$  equations in the form of (1), where  $N_s$  is the number of stations, must be solved to estimate the unknown target position. As it can be observed, the relation between the measurements and the unknown target position is highly non-linear. Nevertheless, the most established strategy to solve the system of hyperbolic equations is to linearize (1) by a Taylor series expansion and solve the resulting (linear) system of equations in the Least Squares (LS) sense [5][6].

The solution provided by the LS location estimator is the minimum residual norm solution (that is the solution of  $\operatorname{argmin} \{\|A\Theta - \hat{m}\|_2^2\}$  where  $A$  is the Jacobian Matrix coming from (1) called *coefficient matrix*,  $\hat{m}$  is the measurement vector composed with the measured TDOAs and  $\Theta$  is the estimated position vector  $(x, y, z)$  to be found) when the resulting system of equations is solved by the pseudoinverse matrix. However, as it is presented in [4], this strategy does not always provide acceptable accuracies because, in many operational conditions, the

coefficient matrix, of the system of hyperbolic equations, may have some linearly-dependent equations and in these cases the solution provided by LS may have a big error. The numerical reason of this is that the pseudoinverse matrix does not always satisfy all Hadamard's conditions [2][3][7] and the problem becomes ill-posed. This problem arises in MLAT scenarios due to: (a) the system geometry relative to the target position, (b) the measurement noise and (c) the quality of the guess point for Taylor linearization.

In this paper two regularization methods to mitigate the accuracy degradation due to the ill-posedness are presented. The first method is based on the Truncated - Singular Value Decomposition (T-SVD) [8] whose main idea consists of construct a modified coefficient matrix by truncating the SVD spectrum of the initial coefficient matrix (i.e. that one used for the LS estimator). The initial coefficient matrix  $A$  is decomposed in its SVD spectrum, then the first  $k_\epsilon$  components of this are retained and the remaining ones are neglected to construct a new coefficient matrix. This procedure is introduced in the iterative method of Taylor linearization instead the pseudoinverse matrix.

The second method is based on the Tikhonov regularization [9][10] that does not neglect any singular value component of the SVD spectrum, but introduces a priori assumptions about the size and smoothness of the solution. This additional information is in the form of a seminorm defined by  $\lambda^2 \|L\Theta\|_2^2$ , where  $\lambda$  is called Tikhonov regularization parameter,  $L$  is called regularization matrix and  $\Theta$  is the desired solution. In this way, the problem to be solved becomes  $argmin \{ \|A\Theta - \hat{m}\|_2^2 + \lambda^2 \|L\Theta\|_2^2 \}$ . One of the most critical factors, to apply the methods commented above, is the choice of the corresponding regularization parameters ( $k_\epsilon$  or  $\lambda$ ), because these define the "amount" of regularization introduced in the location problem. In hyperbolic passive location, if a large "amount" of regularization is introduced, then the regularized solution, either that of the T-SVD or of Tikhonov, tends to the "guess" point used in Taylor based algorithm and, by contrast, if a small "amount" of regularization is introduced, then the regularized solution tends to the LS solution and thus it is clearly affected by the ill-posed problem. In the literature there exist a considerable number of methods and procedures to estimate an approximated regularization parameter value, which are based on the solution of an optimization problem [11] or find a parameter that minimizes some special functions [12][13]. Due to the nature of these methods, they introduce a significant computational load and therefore the computation time can be not suitable for real-time location in Multilateration operations. In this paper, we propose the evaluation of the problem for a few number of regularization parameters (one, two or three) and then choose that solution with the

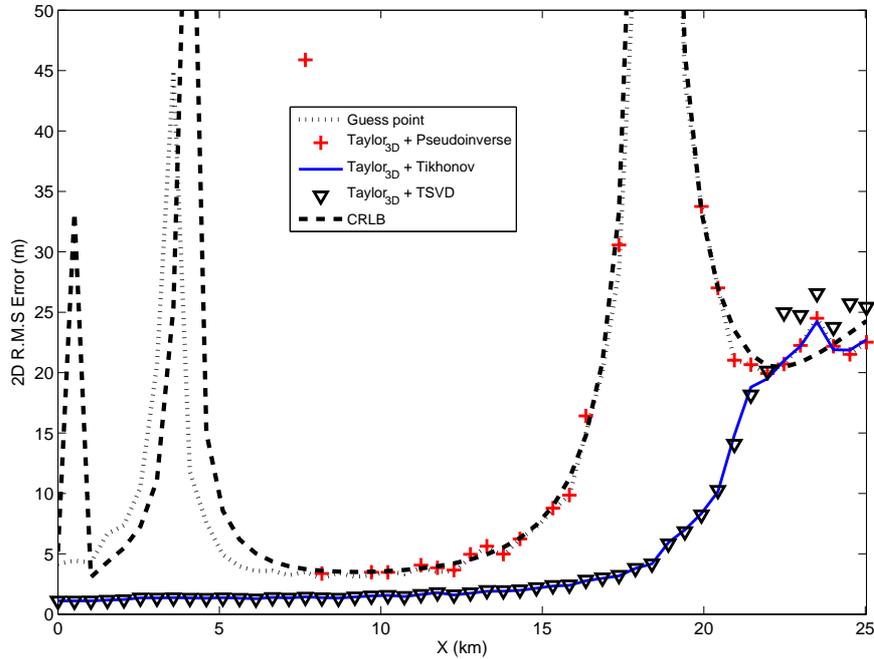


Figure 1: 2D r.m.s error.

minimum residual error. Moreover, we also propose the evaluation of the residual error by using the non-linear problem described by (1) instead of using the classical expression of matrix-vector product.

In the following we have simulated a Multilateration scenario (see [4] for details) to validate the improvements obtained by the implementation of the regularization methods introduced above. In this simulation, the  $2D$  *r.m.s* error of the estimator is computed and compared with the Cramr-Rao Lower Bound (CRLB) estimated accuracy [4].

In Fig. 1 it is shown the  $2D$  *r.m.s* error for the simulated scenario. From this figure it can be observed how the CRLB predicts poor accuracy values for several points (those ones around 1, 4 and 17 km). Then, the accuracy of the LS solution, as it is expected, reaches the CRLB and also diverges in other points. These two first results shown the problem is highly affected by the ill-posedness. On the other hand, it is evident to see how the application of T-SVD and Tikhonov regularization significantly mitigates the ill-posed problem over the entire scenario,

reaching a maximum *r.m.s* error of  $25m$ , which is a value within the boundaries stipulated by the regulatory bodies [1]. Additionally, at the end of the takeoff line, Tikhonov provides slightly better accuracy values than those of the T-SVD.

### 3 Conclusion

T-SVD and Tikhonov regularization to solve hyperbolic passive location problems have been introduced and significant improvements in the location accuracy have been found in particular where the problem results hard ill-posed.

This work has been partially supported by a FPU scholarship (AP2008-03300) from the Spanish Ministry of Education.

### References

- [1] ED-117, MOPS for Mode S Multilateration Systems for Use in Advanced Surface Movement Guidance and Control Systems (A-SMGCS), in: *EUROCAE (Ed.), EUROCAE*, November 2003.
- [2] G.H. Golub, C.F.V. Loan, *Matrix Computations*, Third ed., Baltimore, the Johns Hopkins University Press, 1996.
- [3] J. Hadamard, *Lectures on Cauchy's Problem in Linear Partial Differential Equations*, New Haven, Yale University Press, 1923.
- [4] G. Galati, M. Leonardi, M. Tosti, Multilateration (Local and Wide area) as a distributed sensor system: Lower bounds of accuracy, in: *European Radar Conference, EuRAD*, Amsterdam, 30-31 Oct. 2008, pp. 196-199.
- [5] W.H. Foy, Position-Location Solution by Taylor-Series Estimation, *IEEE Transactions on Aerospace and Electronic System*, AES-12, March 1976: 187-194.
- [6] D.J. Torrieri, Statistical Theory of Passive Location Systems, *IEEE Transactions on Aerospace and Electronic System*, AES-20, March 1984: 183-198.
- [7] A. Ribs, F. Schmitt, Linear Inverse Problem in Imaging, *IEEE Signal Processing Magazine*, 25 July 2008: 84-99.

- [8] P.C. Hansen, Regularization, GSVD and truncated GSVD, *BIT*, 29 September 1989: 491-504.
- [9] A.N. Tikhonov, Solution of incorrectly formulated problems and the regularization method, *Sovieth Math. Dokl.*, 4 (1963) 1035-1038.
- [10] P.C. Hansen, Analysis of discrete ill-posed problems by means of the L-curve, *SIAM Review*, 34 (1990) 561-580.
- [11] V.A. Morozov, On the solution of functional equations by method of regularization, *Sovieth Math. Dokl.*, 7 (1966) 414-417.
- [12] H. Gfrerer, An a posteriori parameter choice for ordinary and iterated Tikhonov regularization of ill-posed problems leading to optimal convergence rates, *Math. Comp.*, 49 (1987) 507-522.
- [13] G.H. Golub, M.T. Heath, G. Wahba, Generalized cross-validation as a method for choosing a good ridge parameter, *Technometrics*, 21 (1979) 215-223.

# Difference schemes for some dual-phase-lag models of heat conduction

J. Cabrera,\* M.A. Castro, F. Rodríguez, J.A. Martín

Dep. Matemática Aplicada, Universidad de Alicante,

Apdo. 99, 03080 Alicante, Spain.

November 30, 2011

## 1 Introduction

Non-Fourier models of heat conduction may account for phenomena that appear in transient responses at the microscale level, such as thermal “inertia”, heat waves and delayed responses to thermal disturbances (see [1]) and they are increasingly being considered in the modelling of microscale heat transfer in engineering and biomedical heat transfer problems (see, e.g., [2, 3]).

In the dual-phase-lag (DPL) model [1, 4], classical Fourier law relating the heat flux vector  $\mathbf{q}(\mathbf{r}, t)$  and the temperature gradient  $\nabla T(\mathbf{r}, t)$ , is replaced by

$$\mathbf{q}(\mathbf{r}, t + \tau_q) = -k\nabla T(\mathbf{r}, t + \tau_T), \quad (1)$$

where  $\tau_q$  and  $\tau_T$  are the corresponding phase lags. Combining Fourier law and the conservation of energy principle, the classical diffusion equation is obtained. In a similar way, using (1) instead of the Fourier law, a delayed partial differential equation is obtained [5]. However, it is common to use first-order approximations in (1),

$$\mathbf{q}(\mathbf{r}, t) + \tau_q \frac{\partial \mathbf{q}}{\partial t}(\mathbf{r}, t) \cong -k \left( \nabla T(\mathbf{r}, t) + \tau_T \frac{\partial}{\partial t} \nabla T(\mathbf{r}, t) \right), \quad (2)$$

---

\*e-mail: [jesus.cabrera@ua.es](mailto:jesus.cabrera@ua.es)

referring to the derived equation as the DPL model [4],

$$\frac{\partial}{\partial t}T(\mathbf{r}, t) + \tau_q \frac{\partial^2}{\partial t^2}T(\mathbf{r}, t) = \alpha \left( \Delta T(\mathbf{r}, t) + \tau_T \Delta \frac{\partial}{\partial t}T(\mathbf{r}, t) \right). \quad (3)$$

Higher order approximations in (1) can also be considered [6]. In this work, model (3) will be denoted DPL(1,1), while corresponding models resulting from second order approximation in  $\tau_q$  and first or second order approximations in  $\tau_T$  will be denoted, respectively, as DPL(2,1), and DPL(2,2).

Finite differences schemes to obtain numerical approximations for DPL(1,1) models and variations, in different spatial domains, have been proposed (see, e.g., [7, 8, 9]). In this work, numerical schemes for DPL(2,2) models in one spatial dimension, with appropriate initial and boundary conditions are developed, characterizing their properties of convergence and stability.

In particular, consider a finite rod with insulated ends at  $x = 0$  and  $x = l$ , so that

$$T(0, t) = T(l, t) = 0, \quad t \geq 0, \quad (4)$$

and initial temperature distribution

$$T(x, 0) = \phi(x), \quad \frac{\partial}{\partial t}T(x, 0) = \varphi(x), \quad \frac{\partial^2}{\partial t^2}T(x, 0) = \psi(x), \quad x \in [0, l], \quad (5)$$

with heat conduction, for  $t > 0$  and  $x \in [0, l]$ , given by the general equation,

$$\begin{aligned} \frac{\partial}{\partial t} \left( AT(x, t) + B \frac{\partial}{\partial t}T(x, t) + C \frac{\partial^2}{\partial t^2}T(x, t) \right) \\ = \frac{\partial^2}{\partial x^2} \left( T(x, t) + D \frac{\partial}{\partial t}T(x, t) + E \frac{\partial^2}{\partial t^2}T(x, t) \right), \quad (6) \end{aligned}$$

where DPL(2,2) is obtained with

$$A = \frac{1}{\alpha}, \quad B = \frac{\tau_q}{\alpha}, \quad C = \frac{\tau_q^2}{2\alpha}, \quad D = \tau_T, \quad E = \frac{\tau_T^2}{2},$$

while DPL(2,1) is the particular case with  $E = 0$ .

## 2 Finite difference scheme

Introducing the two new variables

$$v(x, t) = BT(x, t) + C \frac{\partial}{\partial t}T(x, t) \quad (7)$$

and

$$u(x, t) = AT(x, t) + \frac{\partial}{\partial t}v(x, t), \tag{8}$$

Eq. 6 can be expressed as

$$\frac{\partial}{\partial t}u(x, t) = \frac{\partial^2}{\partial x^2}(aT(x, t) + bv(x, t) + cu(x, t)), \tag{9}$$

where  $a = (C^2 + EB^2 - BDC - ACE)/C^2$ ,  $b = (DC - BE)/C^2$ , and  $c = E/C$ .

Hence, since from (7) and (8) one gets

$$\frac{\partial}{\partial t}T(x, t) = \frac{1}{C}(v(x, t) - BT(x, t)), \tag{10}$$

$$\frac{\partial}{\partial t}v(x, t) = u(x, t) - AT(x, t), \tag{11}$$

the mixed problem given by Eq. 6 with boundary conditions (4) and initial conditions (5) is equivalent to the system consisting of equations (9)-(11), with corresponding boundary and initial conditions.

Consider a bounded domain  $[0, l] \times [0, T_M]$ , for some fixed  $T_M > 0$ , together with the uniform mesh  $\{(x_j, t_n), j = 0 \dots P, n = 0 \dots N\}$ , defined by the increments  $h = \Delta x$  and  $k = \Delta t$ , such that  $l = Ph$  and  $T_M = Nk$ . The value of any function  $w$  at a point  $(x_j, t_n)$  in the mesh will be denoted  $w_j^n$ .

The proposed finite difference scheme for DPL(2,2) model will be derived from the following finite difference approximations in equations (9)-(11),

$$\begin{aligned} \frac{u_j^{n+1} - u_j^n}{k} &= \frac{1}{2}\delta_x^2(aT_j^n + bv_j^n + cu_j^n) + \frac{1}{2}\delta_x^2(aT_j^{n+1} + bv_j^{n+1} + cu_j^{n+1}), \\ \frac{T_j^{n+1} - T_j^n}{k} &= -\frac{B}{2C}(T_j^{n+1} + T_j^n) + \frac{1}{2C}(v_j^{n+1} + v_j^n), \\ \frac{v_j^{n+1} - v_j^n}{k} &= \frac{1}{2}(u_j^{n+1} + u_j^n) - \frac{A}{2}(T_j^{n+1} + T_j^n), \end{aligned}$$

where  $\delta_x^2 w_j^n = (w_{j-1}^n - 2w_j^n + w_{j+1}^n)/h^2$ , with appropriate discretized boundary and initial conditions. Writing  $r = h/k^2$ ,  $x_k = 4C + 2Bk + Ak^2$ ,  $y_k = 4C - 2Bk - Ak^2$ ,  $z_k = 4C + 2Bk - Ak^2$ ,

$$\alpha_k = \frac{r}{2} \left( c + \frac{x_k + z_k}{x_k} \frac{bk}{4} + \frac{ak^2}{x_k} \right), \quad \beta_k = \frac{r}{2} \left( a + \frac{ay_k}{x_k} - \frac{x_k + y_k}{x_k} \frac{bAk}{2} \right),$$

and

$$\gamma_k = \frac{r}{2} \left( b + \frac{4ak}{x_k} + \frac{bz_k}{x_k} \right),$$

and introducing, for  $n = 0 \dots N$ , the vectors  $U^n$ ,  $T^n$ , and  $V^n$ , where  $U^n$  stacks up the  $P - 1$  values  $u_j^n, j = 1, \dots, P - 1$ , and similarly for  $T^n$  and  $V^n$ , and writing  $M$  for the tridiagonal  $(P - 1) \times (P - 1)$  matrix  $M = \text{tridiag}(1, -2, 1)$ , the equations of the scheme can be given in matrix form as follows,

$$(I - \alpha_k M) U^{n+1} = (I + \alpha_k M) U^n + \beta_k M T^n + \gamma_k M V^n, \quad (12)$$

$$T^{n+1} = \frac{k^2}{x_k} U^n + \frac{y_k}{x_k} T^n + \frac{4k}{x_k} V^n + \frac{k^2}{x_k} U^{n+1}, \quad (13)$$

and

$$V^{n+1} = \frac{x_k + z_k}{4x_k} k U^n - \frac{x_k + y_k}{2x_k} A k T^n + \frac{z_k}{x_k} V^n + \frac{x_k + z_k}{4x_k} k U^{n+1}. \quad (14)$$

Hence, starting from the initial values given and taking into account the boundary conditions, the method can progress from step  $n$  to step  $n + 1$ , solving for  $U^{n+1}$  in (12), which can be done very efficiently, since the coefficient of  $U^{n+1}$  is also a tridiagonal matrix, and substituting into equations (13) and (14), to obtain  $T^{n+1}$  and  $V^{n+1}$ . It can be proved, writing the scheme as a two-level scheme, that the method is unconditionally stable, for any finite value of  $r = h/k^2$ , and that it is consistent with the original partial differential equation. Thus, from Lax equivalence theorem [10], the convergence of the scheme follows. It can also be proved that the truncation error is  $O(k^2) + O(h^2)$ , as shown in Figure 1.

## References

- [1] D.Y. Tzou, Macro- to Microscale Heat Transfer: The Lagging Behavior. Washington, Taylor & Francis, 1996.
- [2] L. Wang, and X. Wei. Heat conduction in nanofluids, *Chaos Solitons Fractals* 39: 2211–2215, 2009.
- [3] J. Zhou, J.K. Chen, and Y. Zhang. Dual-phase-lag effects on thermal damage to biological tissues caused by laser irradiations, *Comput. Biol. Med.*, 39: 286–293, 2009.

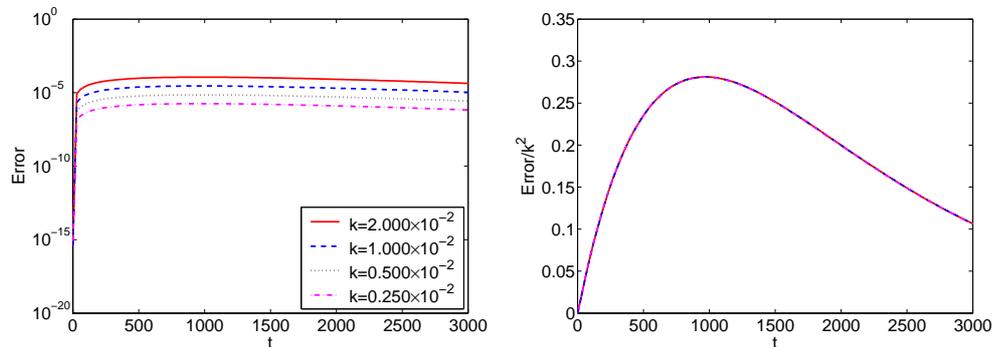


Figure 1: Maximum errors of the approximate solutions, in terms of time (left) and relation with the size of the mesh (right).

- [4] D.Y. Tzou. The generalized lagging response in small-scale and high-rate heating, *Int. J. Heat Mass Transfer* 38: 3231–3240, 1995.
- [5] M. Xu and L. Wang. Dual-phase-lagging heat conduction based on Boltzmann transport equation, *Int. J. Heat Mass Transfer* 48: 5616–5624, 2005.
- [6] D.Y. Tzou. A unified approach for heat conduction from macro to micro-scales, *ASME J. Heat Transfer* 117: 8–16, 1995.
- [7] W. Dai and R. Nassar. A finite difference scheme for solving the heat transport equation at the microscale, *Numer. Methods Partial Differential Equations* 15: 697–708, 1999.
- [8] W. Dai and R. Nassar. A compact finite difference scheme for solving a one-dimensional heat transport equation at the microscale, *J. Comput. Appl. Math.* 132: 431–441, 2001.
- [9] H. Wannng, W. Dai, R. Nassar, and R. Melnik. A finite difference method for studying thermal deformation in a thin film exposed to ultrashort-pulsed lasers, *Int. J. Heat Mass Transfer* 49: 2712–2723, 2006.
- [10] R.D. Richtmyer and K.W. Morton, *Difference Methods for Initial-Value Problems*. New York, John Wiley & Sons, 1967.

# Classifying Slate Tile Quality Using Automated Learning Techniques

J. Martínez\*, M. López†, J.M. Matías‡, and J. Taboada§

(\*) Centro Universitario de la Defensa,

Academia General Militar, Carretera de Huesca, 50090, Zaragoza, Spain

(†) Department of Mechanical Engineering,

University of Vigo, 36310 Vigo, Spain

(‡) Department of Statistics,

University of Vigo, 36310, Vigo, Spain,

(§) Department of Environmental Engineering,

University of Vigo, 36310, Vigo, Spain.

November 30, 2011

## 1 Introduction

Slate quality studies are typically addressed from a deposit perspective, through core sampling in the operational phase [1] or through experimental analyses. Given the research gap, we tackled slate quality evaluation using an objective methodology based on artificial vision and machine learning techniques.

The application of machine learning techniques to the ornamental rock field has been studied by Lopez et al. [2], who developed a granite classifier based on spectral colour data. In the particular case of slate, Araujo et al. [3], [4] used machine learning techniques to evaluate the quality of slate

---

\*e-mail: jmtorres@unizar.es

blocks; however, no research to date has developed an objective classifier of the quality of roofing slate slabs.

The classification problem was resolved using machine learning techniques applied to data captured by a hybrid 3D-2D laser scanner system developed by Lopez et al. [5], for which a set of slate characterization variables was established.

## 2 Machine Learning Techniques

To describe the machine learning techniques implemented in our research, we first briefly describe the classification problem.

### 2.1 The Classification Problem

Classification can be interpreted as partitioning the feature space in mutually exclusive regions in such a way that each region is associated with a class  $C_j$ ,  $j = 1, \dots, c$ , and, given a particular pattern, deciding to which class of the available  $c$  it belongs. Therefore  $\mathcal{Y} = \{1, \dots, c\}$ , and the classification rule can be defined as that function  $g : \mathcal{X} \rightarrow \mathcal{Y}$  such that:

$$g(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in \mathcal{X}_1 \\ \vdots & \vdots \\ c & \text{if } \mathbf{x} \in \mathcal{X}_c \end{cases} \quad \text{where } \bigcup_{i=1}^{i=c} \mathcal{X}_i = \mathcal{X}; \mathcal{X}_i \cap \mathcal{X}_j = \emptyset \text{ if } i \neq j.$$

In other words, each element  $\mathbf{x}$  is assigned to the class  $C_j$  if  $\mathbf{x} \in \mathcal{X}_j$ .

**Cluster Analysis** Cluster analysis aims to group items into homogeneous groups based on similarities between them. These clusters should have no prior information about the classes, as these are suggested by the data itself. For the purposes of this detection, the  $k$ -means [6] algorithm is usually implemented.

**Self Organizing Maps** A SOM model is a set of neurons organized in a regular grid of a smaller dimension than the input space. It requires a set  $\{\mathbf{x}_i\}_{i=1}^n = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ , of  $n$  training elements  $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{im})$ . Each neuron in the map, output layer or competitive layer (in a row  $k$  and column  $j$ ) is represented by a weights vector (prototype vector or codebook

vector) of the same dimension as the input vectors ( $m$ -dimensional) of the form  $\mathbf{w}_{kj} = (w_{kj1}, w_{kj2}, \dots, w_{kjm}) \in \mathbb{R}^m$ .

**Support Vector Machines** Regarding the classification problem, the linearly separable  $\mathbf{x}_i \in \mathcal{X} \subset \mathbb{R}^d, y_i \in \mathcal{Y} = \{-1, 1\}, i = 1 : n$  is separable by a decision function  $f_{\mathbf{w},b}(\mathbf{x}) = \text{sign}(\langle \mathbf{w}, \mathbf{x} \rangle + b)$ ,  $\mathbf{w} \in \mathbb{R}^d, b \in \mathbb{R}$ ; the optimal separating hyperplane, meanwhile, is defined as the maximum-margin separating hyperplane that maximizes the distance to the classes.

The optimization problem of the SVM satisfies the Kuhn-Tucker conditions and so can be solved by solving the dual problem. For the multi-classification problem see [7].

**Multi-layer perceptron neural networks** The multi-layer perceptron is a neural network where the architecture is acyclic, so that there is no feedback to the nodes (with a perceptron in each node) with specific activation functions and fixed weights; the network implements a function  $\mathbf{f} : \mathcal{X} \subset \mathbb{R}^m \rightarrow \mathcal{Y} \subset \mathbb{R}^c$ .

The general form of the MLP is thus:

$$f(\mathbf{x}, \theta) = \sum_{j=1}^h c_j \psi(\mathbf{w}_j^T \mathbf{x} + w_{j0}) + c_0$$

where  $\mathbf{w}$  and  $\mathbf{c}$  are the weights of the input and hidden layer respectively and the function  $\psi$  and  $h$  are the activation function and the number of neurons of the hidden layer respectively.

### 3 Results and Conclusions

Slate tiles classification was based on a set of variables, obtained using a hybrid 3D-2D laser scanner [5], and capable of characterizing each tile objectively. We generated a sample of 75 observations with 29 variables, with one variable corresponding to the slate commercial classification: top quality (1), intermediate quality (2) or reject quality (3).

To obtain the optimal classifier, we tested both unsupervised and supervised machine learning techniques: cluster analysis and SOM as unsupervised techniques and SVM and MLP as supervised techniques. The supervised techniques produced significantly better results than the unsupervised

techniques. Furthermore, the lower error rates obtained by the SVM and MLP techniques (around 1%), in comparison to those obtained by the cluster analysis and SOM techniques, reveal the presence of nonlinearity in the classification problem studied.

## References

- [1] Rivas T., Matías J.M., Taboada J. and Argüelles A. Application of Bayesian networks to the evaluation of roofing slate quality. *Engineering and Geology*, 94:27-37, 2007.
- [2] López M., Martínez J., Matías J.M., Taboada J. and Vilán J.A. Functional classification of ornamental stone using machine learning techniques. *Journal of Computational and Applied Mathematics*, 234:1338-1345, 2010.
- [3] Araújo M., Matías J.M., Rivas T. and Taboada J. Machine learning techniques applied to the construction of a new geomechanical quality index. *International Journal of Computer Mathematics*, 88(9):1830-1838, 2010.
- [4] Taboada J., Matías J.M., Araújo M. and Ordóñez C. Assessing the viability of underground slate mining by combining an expert system with a GIS. *Engineering Geology*, 87(1-2):75-84, 2006.
- [5] López M., Martínez J., Matías J.M., Vilán J.A. and Taboada J. Application of a Hybrid 3D-2D Laser Scanning System to the Characterization of Slate Slabs. *Sensors*, 10:5949-5961, 2010.
- [6] MacQueen J.B. Some Methods for Classification and Analysis of Multivariate Observations. *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press*, 1:281-297, 1967.
- [7] Hsu C.W. and Lin C.J. A comparison of methods for multiclass support vector machines *IEEE Transactions on neural networks*, 13(2):415-425, 2002.

# Onshore Wind Farms Maintenance Optimization Using a Stochastic Model

S.Carlos<sup>\*\*</sup>; I.Marton<sup>\*</sup>, A. Sanchez<sup>†</sup>, and S. Martorell<sup>\*</sup>

(<sup>\*</sup>) Departament d' Enginyeria Química i Nuclear,

Universitat Politècnica de València. Camí de Vera, 14. 46022 València. Spain.

(<sup>†</sup>) Departament d' Estadística i Investigació Operativa Aplicades i Qualitat,

Universitat Politècnica de València. Camí de Vera, 14. 46022 València. Spain.

November 30, 2011

## 1 Introduction

The operational conditions of wind farms are quite different from traditional power stations. Thus, a specific feature of wind power generation is the stochastic behaviour of wind velocity, what determines the energy produced, and also influences the turbine degradation process due to the stochastic load suffered by the wind turbine[1][2]. Such loads influence the failure rate of wind turbine components accelerating the aging process. Nowadays, the maintenance plan scheduled in wind turbine farms include routine checks, testing and maintenance, focused on minimizing the degradation effects. Moreover, maintenance plan is determined by the meteorological conditions. Thus to guarantee safe working conditions, maintenance activities can not be undertaken under adverse meteorological conditions. The maintenance tasks have to be planned to maximize the annual energy generated and minimize the costs associated.

---

<sup>\*</sup>e-mail: scarlos@iqn.upv.es

## 2 Mathematical models

### 2.1 Reliability model

Reliability evaluation is developed using a sequential Monte Carlos simulation. The times to failure and times to repair over a year period of time are generated by sampling the adequate probability distribution. In this paper the failure time distribution is assumed to be a Weibull distribution. The probability density function (pdf) is modified to include imperfect maintenance. So, the pdf depends on the chronological time ( $t$ ) and on the preventive maintenance effectiveness ( $\epsilon$ ), [3].

### 2.2 Power model

The power generated can be evaluated using:

$$P(v(t)) = \begin{cases} 0 & v < v_{ci} \text{ or } v > v_{co}, \\ \frac{a}{1+b \exp(-cv(t))} & v_{ci} < v < v_r, \\ P_r & v_r < v < v_{co} \end{cases} \quad (1)$$

where  $v(t)$  is a random variable, which can be modelled using a two parameters Weibull distribution,  $v_{ci}$  is called the cut-in velocity,  $v_{co}$  is the cut-out velocity and  $P_r$  is the rated power, this is the power generated when the wind velocity varies between rated velocity,  $v_r$ , and  $v_{co}$ . The Annual Energy Produced ( $AEP$ ) can be calculated as:

$$AEP = \sum_{j=1}^{365} (\Delta t - \sigma_j - \mu_j) (P(v(\Delta t)_j)), \quad (2)$$

where  $\sigma_j$  and  $\mu_j$  are the corrective and preventive maintenance duration, respectively. The maximum annual energy produced  $AEP_{max}$  could be obtained when  $\sigma = \mu = 0$ .

### 2.3 Cost model

The total cost,  $c_t$ , has been modeled using the cost contribution of preventive,  $c_{pm}$ , and corrective maintenance,  $c_{cm}$ , using the following expression:

$$c_t = c_{pm} + c_{cm} = \sum_{i=1}^n \sigma_i c_{hpm} + \sum_{i=1}^k \mu_i c_{hpm}, \quad (3)$$

where,  $c_{hpm}$  and  $c_{hcm}$  are the hourly personnel cost to perform the preventive and corrective maintenance activities, respectively,  $n$  is the number of preventive maintenance tasks performed to the interval time considered, and  $k$  is the number of failures observed in the same period.

### 3 Problem formulation

Using the models exposed in section 2 the Multiobjective Optimization Problem can be transformed into a Single objective Optimization Problem using the effectiveness concept:

$$y = f(x) = wE_p + (1 - w)(1 - E_c), \tag{4}$$

where  $w$  is the weighting coefficient with range interval  $[0, 1]$  and  $E_p$  and  $E_c$  are the power and cost effectiveness, defined as:

$$E_p = \frac{AEP_{max} - AEP}{AEP_{max}}, \tag{5}$$

$$E_c = \frac{c_{pm}}{c_t}, \tag{6}$$

### 4 Methodology

The optimization process is developed using the following procedure:

Step 1. The parameters associated to the reliability function are obtained using available wind turbine failure data. It is assumed that the failure process can be modeled using the two-parameter Weibull distribution. So, to characterize the reliability function, the scale and shape parameters must be estimated. These parameters have been obtained using the Maximum Likelihood Estimation method.

Step 2. Preventive and corrective maintenance duration, the failure times and wind speed are generated using a Monte Carlo sampling and the energy and cost are evaluated for every trial, as follows:

- a) Suppose that the wind turbine is operating in  $t=0$ .
- b) Simulate the first failure time. Using the reliability function obtained in step 1 it is possible generate the first random failure using the inverse transformation method.

c) Repeat step b) in a specific time span (in this case one year).

d) For each sample the daily wind velocity is simulated using a Weibull distribution and a chronological daily energy and cost are obtained.

Step 3. The optimization is performed using a direct search algorithm based on the Nelder Mead Simplex method.

## 5 Case of application

The case of application is focused on optimizing the maintenance plan of a wind turbine of 2MW. The behaviour of the two stochastic variables, failure time and wind velocity, considered in this study have been modeled using Weibull distribution.

Other stochastic variables considered in the study are the duration and the effectiveness of the maintenance tasks. Preventive and corrective maintenance durations is modeled as an uniform distribution in the range [6, 10] and [10, 20] hours, respectively. Maintenance effectiveness is also considered to be a uniform distribution in the range of interval [0.7, 1]. Monte Carlo simulation with 500 sampling has been used to generate random values of power, wind and failures. The optimization has been performed using a Nelder-Mead simplex algorithm implemented in MatLab [4]. The optimum maintenance interval obtained from optimization process is 113 days. Fig.1 shows the annual evolution of daily energy produced obtained from the optimization process. In this figure, a comparison between the behaviour of the  $AEP_{max}$  and the  $AEP$  is showed.

## 6 Conclusions

In the last years, the use of wind energy has been increased and this makes necessary to consider the reliability of energy produced that can be obtained and the associated costs. The model proposed in this paper can be used to find the optimum maintenance preventive strategy that maximize the annual power and minimize the cost taking into account the stochastic behaviour of wind velocity, reliability, maintenance duration and effectiveness. This study considers each wind turbine isolated, but preventive maintenance should be analyzed considering the whole wind farm.

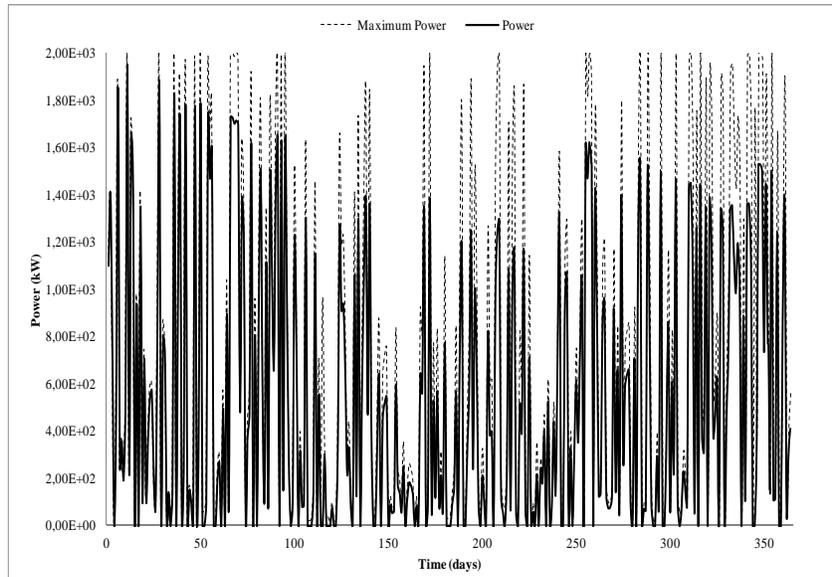


Figure 1: Annual energy produced.

## References

- [1] Tar, K., Szegedi,S. A statistical model for estimating electricity produced by wind energy. *Renewable energy*, Vol 36, pp 823-828, 2011.
- [2] Kusiak,A.,Zheng,H.,Shong,Z.Models for monitoring wind farm power. *Renewable energy*, Vol 34, pp 583-590, 2009.
- [3] Martorell, S., Sanchez,A., Serradell, V.Age-dependent reliability model considering effects of maintenance and working conditions. *Reliability Engineering and System Safety*, Vol 64, pp 19-31, 1999.
- [4] Nelder-Mead simplex direct search, written for Matlab, <http://sdetoolbox.sourceforge.net>

# Low-rank Update of Preconditioners for the nonlinear Richard's Equation.

L. Bergamaschi\*, R. Bru<sup>†</sup>, A. Martínez\*, J. Mas<sup>†</sup>, and M. Putti\*

(\*) Department of Mathematical Methods and Models for Scientific Applications

University of Padua,

via Trieste 63, 35121 Padova, Italy,

(†) Instituto de Matemática Multidisciplinar,

Universitat Politècnica de València, Spain,

November 30, 2011

## 1 Introduction

In this work we are mainly concerned with the efficient preconditioning of the linear system arising in the Picard iteration for the solution of a general system of nonlinear equations:  $\mathbf{F}(\mathbf{x}) = 0$ , which is usually written as

$$\begin{cases} J(\mathbf{x}_k)\mathbf{s}_k &= -\mathbf{F}(\mathbf{x}_k) \\ \mathbf{x}_{k+1} &= \mathbf{x}_k + \mathbf{s}_k \end{cases} \quad (1)$$

We employ the preconditioned Conjugate Gradient (PCG) method for the solution of the linear system, so that two nested iterative procedures need to be implemented, the outer iteration formed by the Newton steps and the inner iterations within the PCG method.

Since in each Newton step a new system has to be solved, we are dealing with the construction of a sequence of preconditioners  $P_k \approx J_k^{-1}$  which are

---

\*Work supported by the Spanish DGI grant MTM2010-18674 e-mail: {berga,acalomar,putti}@dmsa.unipd.it

<sup>†</sup>e-mail: {rbru,jmasm}@mat.upv.es

“optimal” in the sense that they would minimize the constant  $C$  of:

$$\|I - P_k J(\mathbf{x}_k)\| \leq C. \quad (2)$$

Our aim is to solve the Picard SPD system (1) with the PCG method, starting with an initial preconditioner, IC(0), AINV preconditioners, computed from the initial Jacobian, and to update this preconditioner using low-rank matrices. A sequence of SPD preconditioners  $P_k$  can thus be defined by imposing the secant condition, as used in the implementation of quasi-Newton methods.

## 2 BFGS-update of preconditioners and implementation

The idea is to start with a preconditioner  $P_0 \approx J_0^{-1}$ . Let us define  $\mathbf{y}_k = \mathbf{F}(\mathbf{x}_{k+1}) - \mathbf{F}(\mathbf{x}_k)$  and recall that  $\mathbf{s}_k$  is the solution of the  $k$ th Newton system. The preconditioner is updated by the recurrence formula

$$P_{k+1} = \left( I - \frac{\mathbf{s}_k \mathbf{y}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} \right) P_k \left( I - \frac{\mathbf{y}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} \right) + \frac{\mathbf{s}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k}. \quad (3)$$

If the Jacobian matrices are SPD and so is  $P_0$ , then  $P_k$  is also SPD under the condition  $\mathbf{s}_k^T \mathbf{y}_k > 0$ . It can also be easily proved that the sequence of matrices just defined satisfies the secant condition  $P_{k+1} \mathbf{y}_k = \mathbf{s}_k$ , and that  $\|I - P_k J(\mathbf{x}_k)\|$  can be tuned to any fixed accuracy by suitable choices of the initial guess  $\mathbf{x}_0$  and the initial preconditioner  $P_0$ . Note that this makes our preconditioner almost ideal in the sense of (2). The linear iteration is stopped as soon as the following test is satisfied

$$\|J(\mathbf{x}_k) \mathbf{s}_k + \mathbf{F}(\mathbf{x}_k)\| \leq \eta_k \|\mathbf{F}(\mathbf{x}_k)\|. \quad (4)$$

At a certain nonlinear iteration level,  $k$ , and given a vector  $\mathbf{z}_k^{(l)}$ , we want to compute  $\mathbf{c} = P_k \mathbf{z}_k^{(l)}$ . where with superscript  $l$  we indicate the linear iteration index. Let us suppose we compute an initial preconditioner  $P_0$ . Then, at the initial nonlinear iteration  $k = 0$ , we simply have  $\mathbf{c} = P_0 \mathbf{z}_0^{(l)}$ . For  $k \geq 0$ ,  $P_{k+1}$  is given inductively by (3).

We define  $k_{\max}$  as the maximum number of rank two corrections we allow. When the nonlinear iteration counter  $k$  is larger than  $k_{\max}$ , the vectors

$\mathbf{s}_i, \mathbf{y}_i$ ,  $i = k \text{ MOD } k_{\max}$  are replaced with the last computed  $\mathbf{s}_k, \mathbf{y}_k$  and a new preconditioner  $P_0$  is computed. Vectors  $\{\mathbf{s}_i, \mathbf{y}_i\}$  are stored in a matrix  $V$  with  $n$  rows and  $2 \times k_{\max}$  columns.

**RESTARTED NEWTON-BFGS (RNBFGS) ALGORITHM**

**Input:**  $\mathbf{x}_0, \mathbf{F}, k_{\max}, nlmax, tol$

- $k_{aux} := 0, k := 0$
- WHILE  $\|\mathbf{F}(\mathbf{x}_k)\| > tol \|\mathbf{F}(\mathbf{x}_0)\|$  AND  $k < nlmax$  DO
  1. IF  $k_{aux} = 0$  THEN Compute  $P_0$  approximating  $J_0^{-1}$   
     ELSE Update  $P_k$  from  $P_{k-1}$  using the  
     columns of  $V$ .
  2. Solve  $J(\mathbf{x}_k)\mathbf{s}_k = -\mathbf{F}(\mathbf{x}_k)$  by a Krylov method with  
     preconditioner  $P_k$  and tolerance  $\eta_k$ .
  3.  $\mathbf{x}_{k+1} := \mathbf{x}_k + \mathbf{s}_k$
  4.  $V(*, 2k_{aux} + 1) := \mathbf{s}_k, V(*, 2k_{aux} + 2) := \mathbf{y}_k,$
  5.  $k := k + 1$   $k_{aux} = k \text{ MOD } k_{\max}$
- END WHILE

**Output:**  $nlit := k, \mathbf{x}_k, \|\mathbf{F}(\mathbf{x}_k)\|$

### 3 Numerical Results

Here we give the numerical performance of our sequence of preconditioners in solving a realistic test case resulting from a 3d FE discretization of the Richard's equation that considers water infiltration from a surface drain into an initially dry soil connected on one side to a channel. The domain is formed by a flat 50 m×50 m terrain overlying a 2 m deep sandy soil. The drain is simulated using a constant flux Neumann boundary condition, while the channel is kept at a constant pressure head. No flow boundary conditions are imposed in all the other parts of the frontier. The sandy soil is characterized by a uniform and isotropic saturated conductivity value  $K_s = 10^{-3}$  m/s and a porosity  $\phi = 0.3$ . The steady state is reached by infiltration from the top and from the channel.

The simulations were run on a Intel I7 Core 2 QUAD workstation running at 2.4 GHz with 16 Gb RAM. The CPU times are measured in seconds. In the solution of the systems (1) we employed the PCG iterative method and stop the iteration whenever the exit test (4) with constant  $\eta_k = 10^{-4}$  is fulfilled. The nonlinear iteration is stopped whenever  $\|\mathbf{F}(\mathbf{x}_k)\| \leq 10^{-8}\|\mathbf{F}(\mathbf{x}_0)\|$  (i. e.  $\text{tol} = 10^{-8}$ ). In the subsequent tables we report the total number of nonlinear and linear iterations and CPU time. When  $k_{\max} = 0$  we assume that the initial preconditioner is computed, and not updated, at each nonlinear iteration.

Note that our Restarted Newton BFGS algorithm provides an improvement in terms of number of iterations, irrespective on the  $k_{\max}$  value and on the initial preconditioner. In almost all the runs there is also a reduction of the CPU time. The iteration number reduction appears to be monotone with  $k_{\max}$ , however high values of this parameter may lead to increased CPU time due to the increasing cost per iteration. The advantage of the proposed preconditioner is particularly evident when the computation of the initial preconditioner is costly.

$P_0$	$k_{\max}$	NLIT	LINIT	CPU
ILU(0)	0	12	1619	98.55
ILU(0)	1	12	1522	88.30
ILU(0)	2	12	1496	87.22
ILU(0)	3	12	1475	88.58
ILU(0)	no restart	12	1428	82.76
AINV(0.05)	0	12	4075	191.27
AINV(0.05)	1	12	3815	189.77
AINV(0.05)	3	12	3663	191.20
AINV(0.05)	5	12	3577	193.99
AINV(0.02)	0	12	2840	282.12
AINV(0.02)	2	12	2610	244.88
AINV(0.02)	5	12	2519	227.47
AINV(0.02)	no restart	12	2481	204.14

Table 1: Number of linear iterations (LINIT) and CPU time for the solution of the Picard-linearized systems with different initial preconditioners and values of the restart parameter  $k_{\max}$ . Steady state case.

# Predicting the development of anxiolytic addiction: a discrete mathematical model approach

E. de la Poza<sup>†</sup> \*; N. Guadalajara<sup>†</sup>, L. Jódar<sup>‡</sup> and P. Merello<sup>‡</sup>

(<sup>†</sup>) Departamento de Economía y Ciencias Sociales,

(<sup>‡</sup>) Instituto Universitario de Matemática Multidisciplinar,

Universitat Politècnica de València

November 30, 2011

## 1 Introduction

The most prescribed psychotropic drugs are anxiolytics (AX) belonging to the group of benzodiazepines [1] N05B, code defined as the classification system Anatomical Therapeutic Chemical (ATC). The main causes of the prescription of AX by the general practitioners (GP) of Primary Health Care Centers (PHCC) are the presence of anxiety and/or insomnia [2].

In Spain, the consumption of AX has increased significantly since the 90's. In fact, in 2007, it represented the 16% of the total pharmaceutical expenditure. In 2009, 15.5% of Spanish population consumed AX in a regular bases and the public expenditure of it summed up over 200 million Euros.

These drugs are only suitable for short-term treatment (between 15-30 Defined Daily Doses (DDD) $\approx$ 1 month). However, several studies show that GP often prescribe treatments for longer, which increases the risk of patients developing tolerance and physical or psychological addiction [3]. The consequences of developing drug dependence may lead to brain damage resulting in poor concentration, reduced ability to react, memory loss, and accidents.

---

\*elpopla@esp.upv.es

It should seek alternative therapies to psychopharmacology, to reduce not only the economic cost, but also the social cost [4].

The aim of this paper is to develop a mathematical model to predict addictive behaviours to AX for the next four years (2011-2014) in the province of Castellón considering as main variables that can influence the AX consumption: the social contact [7], [8], the economic situation [6], and the marital termination [5]. This will lead us to estimate the social effects, the pharmaceutical expenditure, and possible public health recommendations.

## 2 Mathematical model and results

The study population in this work is made up of subjects who were registered as inhabitants in the Census of the province of Castellón in the Autonomous Community of Valencia (Spain) and also were registered in the Population-based Information System (SIP) in the province of Castellón in 2009 and 2010.

To build the mathematical model, the population is divided into three subpopulations according to the AX (number of DDD) prescribed by year [9]: *A*: long-term users (more than 180DDD  $\approx$  more than 6 months); *R*: regular users (31-180DDD  $\approx$  2-6 months); *I*: incidental users (0-30 DDD  $\approx$  1 month). The subpopulation of incidental consumers (*I*) includes all people that have never consumed for the period considered or they have done just occasionally.

The total population (*P*) at any time *n* is expressed as follows:

$$P_n = I_n + R_n + A_n . \quad (1)$$

The factors considered in the development of addiction to AX consumption were: the economic scenario, the marriage termination (embracing marital dissolution through divorce, separation and annulment) and the social contact. The model was built following the assumptions of previous studies of social contagious [7], [10]. The values of all parameters were estimated from different sources of information and hypothesis with the exception of the transit rates ( $\mu_1, \mu_2$ ) that were adjusted by the model with the Nelder-Mead algorithm.

The dynamic of the population can be described by the following system

of difference equations ( $n$ , time in years):

$$\begin{aligned} I_{n+1} - I_n &= \beta(I_n + R_n + A_n) - d_i I_n - \alpha I_n + \gamma_1 R_n + \gamma_2 A_n - \mu_1 I_n A_n - \psi I_n, \\ R_{n+1} - R_n &= -d_r R_n + \alpha I_n - \alpha R_n - \gamma_1 R_n + \gamma_1 A_n + \mu_1 I_n A_n - \mu_2 R_n A_n + \psi I_n - \psi R_n, \\ A_{n+1} - A_n &= -d_a A_n + \alpha R_n - \gamma_1 A_n - \gamma_2 A_n + \mu_2 R_n A_n + \psi R_n. \end{aligned} \quad (2)$$

Different simulations are developed considering that the value of  $\alpha$  (rate of unemployed people that consume AX) evolves during the next four years assuming different economic scenarios [19]: L, V,  $\sqrt{\quad}$ , International Monetary Fund (IMF) and optimistic scenario.

For the particular case of Castellón, the percentage of consumers will double from 2010 to 2014. As consequence, the public expenditure of AX will increase for all possible economic scenarios.

### 3 Conclusions

The proposed model let us predict the development of the addicts rate to AX for the next four years as well as the public healthcare expenditure derived from the prescription of AX. Considering demographic, economic and cultural similarities between provinces, the same methodology could be applied to any Spanish province or the whole country.

The results obtained should be considered as a valuable source of information for Public Authorities in order to control and reduce the AX prescription by GP through different policies, mainly alternative therapies to the pharmacological ones since the results obtained show how the GPs' prescribing patterns are not in accordance with recommended dosage and duration of AX treatment.

### References

- [1] A. Bocquier, K. Bezzou, S. Nauleau, P. Verger, Dispensing of anxiolytics and hypnotics in southeastern France: demographic factors and determinants of geographic variations, *Fundamental and Clinical Pharmacology* 22 (2008) 323-333.
- [2] C. Vedia, S. Bonet, C. Forcada, N. Parellada, Estudio de utilización de psicofármacos en atención primaria, *Atención Primaria* 36 (5) (2005) 239-247.

- [3] R. Van Hulten, A.B. Bakker, A.C. Hubert, G. Leufkens. Determinants of change in the intention to use benzodiazepines. *Pharm World Sci* (2001); 23(2)70-75.
- [4] B.M. Verdel, P.C. Souverein, T.C.G. Egberts, T.P. van Staa, H.G.M. Leufkens, F. De Vries, Use of antidepressant drugs and risk of osteoporotic and non-osteoporotic fractures, *Bone* (2010), 47 (3): 604-609.
- [5] L. Manthey, T. Van Veen, E.J. Giltay, J.E. Stoop, A. K. Neven, B. W. J. H. Penninx, F. G. Zitman. Correlates of (inappropriate) benzodiazepine use: the Netherlands Study of Depression and Anxiety (NESDA). *British Journal of Clinical Pharmacology* (2011), 71, (2): 263-272.
- [6] A. Dragun, A. Russo, M. Rumboldt. Socioeconomic Stress and Drug Consumption: Unemployment as an Adverse Health Factor in Croatia. *Croatia Medical Journal* (2006), 47 685-692.
- [7] I. García, L. Jódar, P. Merello, F.J. Santonja. A discrete mathematical model for addictive buying: Predicting the affected population evolution, *Mathematical and Computer Modelling*, (2011), 54 (7-8): 1634-1637.
- [8] N.A. Christakis, J.H. Fowler, *Connected: The Surprising Power of Our Social Networks and How they Shape Our Lives*, Hachette Book Group, (2009).
- [9] R. Van Hulten, H.G. Leufkens, A. Bakker. Usage of benzodiazepines in a Dutch community: a 10-year follow up. *Pharmacy World and Science* (1998) 20, 78-82.
- [10] F.J. Santonja, R. Villanueva, L. Jódar, G. González-Parra, G. Mathematical modelling of social obesity epidemic in the region of Valencia, Spain. *Mathematical and Computer Modelling of Dynamical Systems*, (2010), 16 (1): 23-34.
- [11] A. Trías, De la V a la L, el abecedario de la salida de la crisis (in Spanish). Available at: [www.elpais.com/articulo/economia/V/L/abecedario/salida/crisis/elpepieco/20090412elpepieco/mathaccent"0250/relax6/Tes?print=1](http://www.elpais.com/articulo/economia/V/L/abecedario/salida/crisis/elpepieco/20090412elpepieco/mathaccent).

# LGEM: A Lattice Boltzmann Economic Model for income distribution and tax regulation

J. Cerdá\*, C. Montoliu, and R. J. Colom

Instituto de Instrumentación para Imagen Molecular (I3M)  
Centro mixto CSIC ? Universitat Politècnica de València ? CIEMAT,  
camino de Vera s/n, 46022 Valencia, España

November 30, 2011

## 1 Introduction

Econophysics is an interdisciplinary research field which applies statistical physics methods for solving problems in economics and finance. The term “econophysics” was first introduced by the theoretical physicist Eugene Stanley[1]. This novel discipline uses mathematical methods developed in statistical physics to study statistical properties of complex economic systems consisting of a large number of humans, and it can be considered as a branch of applied theory of probabilities [2]. In this sense, econophysics has much common ground with agent-based modeling and simulation, as it studies mathematical models of a large number of interacting economic agents [3]. Econophysics distances from the classical approach of economics, mainly representative-agent based, which ignores statistical and heterogeneous aspects of the economy.

---

\*e-mail:joacerbo@eln.upv.es

## 2 A Simple Lattice Gas Automata to Model Money Distribution

Following the principles of econophysics, we have developed a Lattice Gas Automata to model money distribution in a multi-agent system. Economic agents are considered to be molecules of a gas, and each molecule has an internal state variable  $m$ , which to our purposes, represents the quantity of money for this particular agent. Agents are allowed to move in a  $d$ -dimensional lattice  $L$ . In this particular model,  $d$  is taken to be 2 because it seems to represent in a more natural way the human movement on a territory.

The performance of the automata is taken in two phases. The first phase corresponds to a simple movement dynamics: each agent moves in the lattice. For this movement, we have selected a simple random walk dynamics. Time advances in discrete time steps of  $\delta t$ . For each time step, each agent randomly selects one of its neighbouring empty sites and performs its movement.

The second phase of the automata is properly the economic transition. When two agents are placed in two neighboring sites, an economic transaction takes place. In order to simplify the dynamics, a very simple model for an economic transaction has been considered. For each economic transaction, there is a probability  $1 - p_t$  that the final result of the transaction is that both agents finish with the same amount of money. But there is a probability  $p_t$  that one of them wins money from the other one, in order to keep the total amount of money constant. In that case, the agent that “wins” money is selected at random. The quantity of money transferred from one agent to the other has been taken to be fixed, as this has no influence over the stationary distribution of the system.

The proposed model does not consider the possibility of having debt. In the practice, whenever an agent with  $m = 0$  enters in an economic transaction, it can only have two results, instead of the three previously given results. Agents with no money can only win or stay unchanged.

## 3 Adding Features to the Model

LGEM can be slightly modified to increase its usefulness introducing a new phase, between phase 1 and 2, in the global behavior of the system. In this phase, money is redistributed by going from richer agents to poorer ones.

Several strategies can be designed for this phase. We will focus in the

strategies which we can call “charity strategies”: whenever an agent interacts with a poorer agent, there is a certain probability for the richer to voluntarily give some money to the poorer.

Two parameters have been introduced for each model:  $p_c$  and  $a_c$ .  $p_c$  always controls the probability of a richer agent deciding to give some money to a poorer one, while  $a_c$  controls the amount of money involved in the process.

### **MODEL 0: Constant Probability, Constant Magnitude, On poorest ( $m = 0$ ) agents**

In this model, whenever a richer agent ( $m > 0$ ) collides with a poorest one ( $m = 0$ ), there is a constant probability  $p_c$  for him to voluntarily give a fixed amount of money  $a_c$ . The only constraint of this process is that the money of the richer agent must be greater than  $a_c$ .

The result of this strategy is an important decrease of the poorest agents, a fact that can be seen in figure 1 (a).

### **MODEL 1: Constant Probability, Linear Magnitude, On poorest ( $m = 0$ ) agents**

In this model, whenever a richer agent ( $m > 0$ ) collides with a poorest one ( $m = 0$ ), there is a constant probability  $p_c$  for him to voluntarily give a variable amount of money.

The richest agent of the set (with money  $m_{max}$ ) gives an amount of  $a_c$ . An agent with lesser money will give an amount of money given by

$$a = \frac{m}{m_{max}} a_c \quad (1)$$

Again, the money of the richer agent must be greater than  $a_c$ .

The result of this strategy can be seen in figure 1 (b). One remarkable fact regarding to this strategy is the long-range effect for reducing not only the poorest agents, but also those who have less money than the average value.

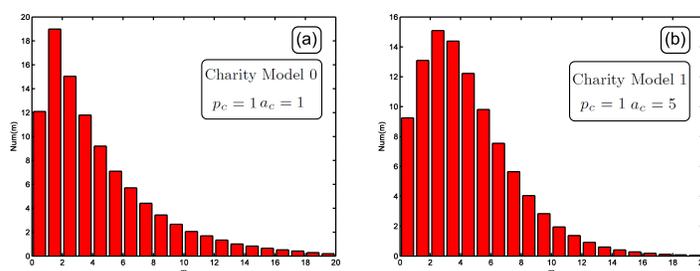


Figure 1: Results for a system of 200 agents in a 50x50 square lattice. Each agent has been initialized to have  $m_0 = 4$ .

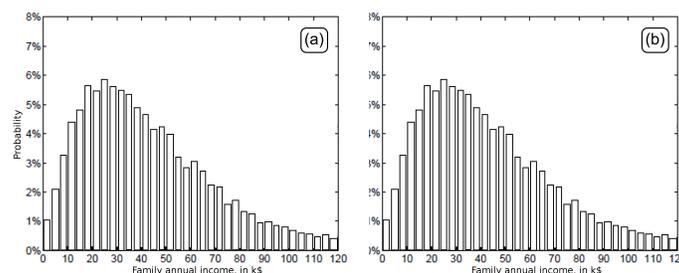


Figure 2: Probability distribution of income in the US in 1996. (a) Families with two adults. (b) All families.

## 4 Conclusions and Further Work

The previously presented models lead to a global behavior different from the exponential distribution. The main feature is a reduction on the number of agents with zero money. It turns out that these distributions adapt well to real market distributions. For example, in [4], the probability distribution of income for US families with two adults in 1996 is given. In the same paper, the probability distribution of income for all US families in 1996 is also given. Both distributions can be seen in Figure 2. The similarity between this distribution and the distributions obtained by LGEM is remarkable.

Considering the distributions obtained by LGEM models, and real data from economic systems, it turns out that these models can be very useful from a practical point of view. They can be considered as “toy models” to experiment with tax formulae or income redistribution strategies, before implementing in real systems.

The main feature of LGEM is its simplicity and the universality of the results obtained. Just by adjusting a few parameters, almost any distribution can be reproduced. Additionally, the process by which each agent performs its evolution is a parallel, local process, with the only requirement, in some particular redistribution strategies, of knowing a unique global value. All these facts lead to a very easy implementation.

LGEM can be expanded in several ways, to study new phenomena. As it has been explained, the money redistribution process has been designed to take place just in the cases when a richer agent meets a poorest ( $m = 0$ ) one. New strategies can be designed to consider the possibility of a “charity transaction” to take place linearly with the difference of money between the two agents. On the other hand, economic transactions have been considered to take place completely at random. More interesting models can be designed, which consider different probabilities for winning and losing, as a function of the money of each agent.

Finally, there is a last expansion that we are working on, and we believe to be the most interesting aspect of LGEM: spatiality. One of the best features of LG is the fact that topology has a direct influence on the performance. A very interesting study would consist of having two different economic systems capable of interacting with each other in some particular sites. We strongly believe that the study of the interactions between systems can lead to a deeper understanding of global economy dynamics.

## References

- [1] B. K. Chakrabarti, *Econophys-Kolkata: a short history*. Chapter: *Econophysics of Wealth Distributions*, pp:225–228, Springer, 2005.
- [2] R.N. Mantegna , and H.E. Stanley. *An introduction to econophysics: correlations and complexity in finance Cambridge University Press*, 1999.
- [3] V.M. Yakovenko, *Encyclopedia of Complexity and Systems Science*. Chapter: *Econophysics, Statistical Mechanics Approach To*, pp:2800–2826, Springer, 2009.
- [4] A. A. Dragulescu, and V. M. Yakovenko *Evidence for the exponential distribution of income in the USA Europ Phys J B*, 20:585–589, 2001.

# Development of a Non-Reflective NSCBC for the Multi-Dimensional non-linear acoustic simulation of silencers for internal combustion engines

A. Montorfano \*, F. Piscaglia, and G. Ferrari

Dipartimento di Energia, Politecnico di Milano  
via Lambruschini 4, I-20156 Milano (Italy)

November 30, 2010

## 1 Introduction

Non-reflecting boundary conditions have a great variety of applications in the field of Computational Fluid Dynamics (CFD); in particular, they are often used in acoustic applications when the open-end of a piping system has to be modeled as an ideal anechoic termination. A typical problem needing non-reflecting boundary condition is the calculation of the Transmission Loss (TL) of a silencer for internal combustion engines. TL is defined as the ratio between incident and transmitted sound pressure levels measured, respectively, upstream and downstream the component. To avoid pollution of the transmitted waves by undesired reflections from the outlet, an anechoic termination has to be employed.

In this work, simulations of acoustic behavior of different silencers for ICE are presented. A non-reflecting boundary condition based on the Navier-Stokes characteristic analysis [2, 3] has been developed in the framework of

---

\*Corresponding author. Tel. (+39 02 2399 3909), e-mail: andrea.montorfano@polimi.it

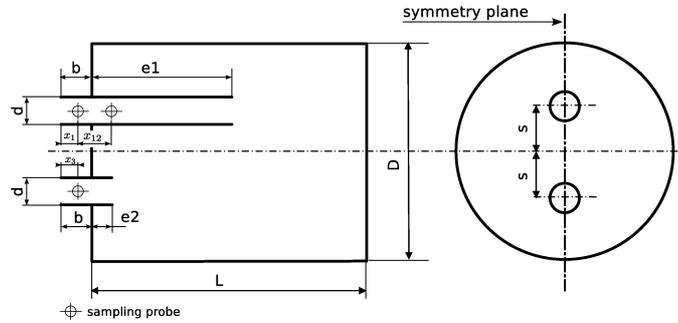


Figure 1: Geometric features of the silencer with flow reversal. Geometric details and experimental measurements are courtesy of AVL.

the open-source code OpenFOAM<sup>®</sup>, and it has been applied to a standard compressible solver to carry out acoustic simulations.

Results are validated against experimental data for two classes of silencers: reverse flow chambers and single-plug perforated mufflers.

## 2 Case setup

The first set of tests has been performed on a reverse flow chamber configuration. The generic geometry is represented in Fig. 1, where the main dimensions are specified as follows. Chamber diameter  $D$  is 197 mm for all cases, while chamber length  $L$  is 494 mm for cases RC-11 and RC-12, 377 mm for case RC-m and 127 mm for case RC-s. Length of inlet channel ( $e_1$ ) is 17 mm for cases RC-11 and RC-s, while for cases RC-12 and RC-m it is 257 mm and 167 mm respectively. Outlet channel ( $e_2$ ) is 17 mm long for all cases. Both inlet and outlet pipes have a diameter  $d = 50$  mm.

On the inlet section of the chambers a pulse function with a frequency content ranging from 20 to 2000 Hz has been employed as acoustic source.

Incident wave amplitude has been determined by sampling the pressure at two point along the inlet channel and then applying the two-sensor method described by Seybert [1]. The transmitted wave amplitude was directly measured by a sampling probe near the outlet section.

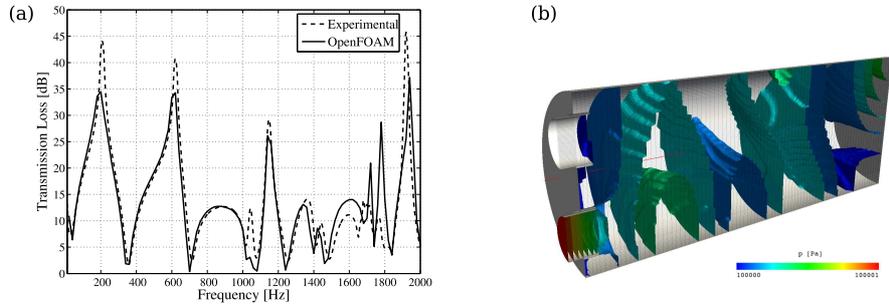


Figure 2: (a) TL of the RC-11 silencer. (b) Pressure isosurfaces at  $t = 0.44$  s.

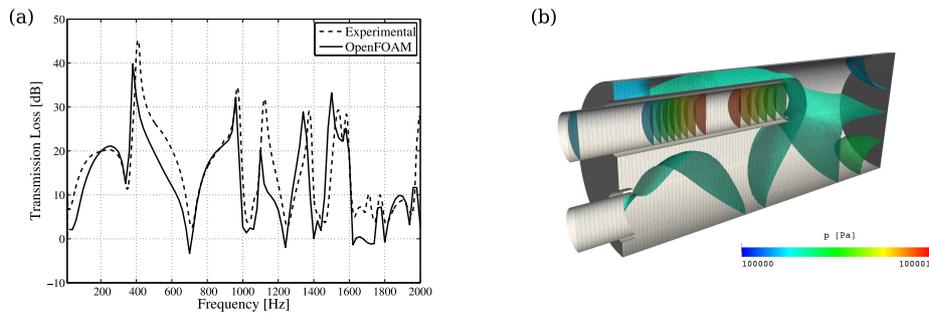


Figure 3: (a) TL of the RC-12 silencer. (b) Pressure isosurfaces at  $t = 0.44$  s

### 3 Results

Figures 2 to 5 show comparisons between computed and measured TL and snapshot of pressure waves at  $t = 0.44$  s for the four configurations already mentioned. Agreement between simulations and experiments looks very good for three cases (Figs. 2-a, 3-a and 4-a) across the whole frequency range, while the last geometry (Fig. 5-a) has some discrepancies in the high-frequency part of the graph. This is probably due to the non-planar shape of the pressure waves in the measurement region: this can be easily seen in Fig. 5-b.

By means of the proposed method, complex configurations like perforated mufflers can be simulated in details. Fig. 6 shows the instantaneous flow field and the TL of a single-plug perforated muffler with no mean flow. Comparison between measured and computed TL show a quite good agreement up to 1000 Hz. Beyond that point, the fit between simulations and experiments deteriorates, even though it is difficult to draw definite conclusions since the

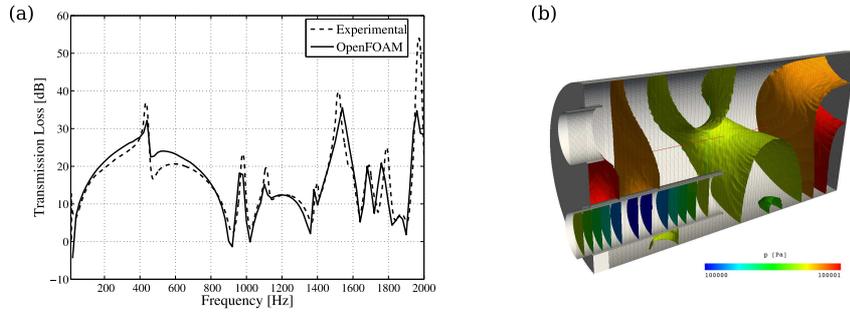


Figure 4: (a) TL of the RC-m silencer. (b) Pressure isosurfaces at  $t = 0.44$  s

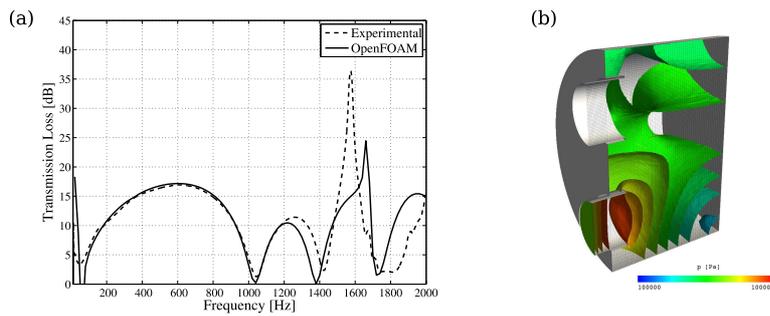


Figure 5: (a) TL of the RC-s silencer. (b) Pressure isosurfaces at  $t = 0.44$  s

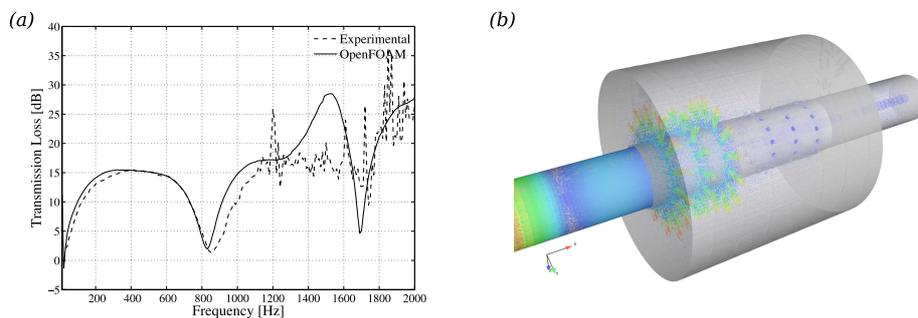


Figure 6: (a) TL of the perforated silencer. (b) Snapshot of the instantaneous velocity field

accuracy of the experimental data is unknown.

## 4 Conclusions

In the present work numerical simulations of TL for complex silencers geometries like reverse chambers and perforated mufflers have been carried out. Scope of this work was the validation of the non-reflecting outlet boundary condition implemented in OpenFOAM<sup>®</sup> to apply to acoustic simulations.

Results shown that the proposed approach has a good capability in predicting the acoustic behavior of complex silencers configurations, thus making it promising as a tool for industrial applications.

## References

- [1] A. F. Seybert, Two-sensor methods for the measurement of sound intensity and acoustic properties in ducts *Journal of Acoustic Society of America*, 82(6):2233–2239, 1988.
- [2] D. H. Rudy, J. C. Strikwerda, A nonreflecting outflow boundary condition for subsonic navier-stokes calculations, *Journal of Computational Physics*, 36(1):55–70, 1980.
- [3] T. J. Poinso and S. K. Lele, Boundary conditions for direct simulations of compressible viscous flows *Journal of Computational Physics*, 101(1):104-129, 1992.
- [4] G. Lodato and P. Domingo and L. Vervisch, Three-dimensional boundary conditions for direct and large-eddy simulation of compressible viscous flows *Journal of Computational Physics*, 227:5105-5143, 2008
- [5] M. L. Munjal, Acoustics of ducts and mufflers: with application to exhaust and ventilation system design, New York, John Wiley and Sons, 1987

# Artificial Neural Model for Classification of Volunteers

Jeronimo Mora-Pascual\*,  
Higinio Mora-Mora\*,  
Maria Teresa Signes-Pont\*,  
and Patricia Soler-Javaloy†

(\*) Specialized Processor Architectures Laboratory, Department of Computer Technology,  
University of Alicante,

(†) Sociologist and social worker in the Student Support Center,  
University of Alicante

November 30, 2011

## 1 Introduction

Voluntary tasks are also a good choice to occupy free time, though it does not mean that volunteers are necessarily influenced by this motivation. This type of committed aid conduct needs certain stimulating elements. From the Social Sciences, the main theories which base their analysis on the predicting elements of the altruistic conduct distinguish five types of explications: firstly, biological motivations; secondly, personal motivations and those which are proper to the history of socialisation itself; thirdly affective (empathy) motivations; fourthly, cognitive motivations and fifthly, condition motivations. The multidisciplinary approach of our study should make it easier the development of decision support models and expert systems of automatic detection of basic qualities for the practice of the volunteer work in areas of

---

\*e-mail:jeronimo@dtic.ua.es

specific intervention, which would help the staff selection processes guarantee the choice of suitable personnel for determined areas of social intervention. In short, we consider that the use of Self-Organizing Maps (SOM) will contribute favourably to the development of the selection model and to our capacity of generalization of the influential factors in the volunteer.

## **2 Analysis of the social voluntary work**

The study of volunteering is a very complex task and shows a rich diversity of questions and perspectives. At the basis there is a common fascination with the experience, which appears as a contradiction in certain disciplines, while others see it as a natural part of social life or a matter of personality. For example, given the underlying assumptions of the self-interested rational thinking in economics, why would any rational individual make a great effort to accept the costs of an activity that provides no material gains? Thus, economists set out to do a cost-benefit analysis of volunteering for individuals, paying attention to material and nonmaterial benefits that may compensate for the cost of volunteering to resolve this otherwise irrational behaviour [1]. Sociologists and political scientists, on the contrary, view volunteering as an expression of nucleus social philosophy such as freedom, social cohesion and solidarity [2]. And social psychologists have identified a pro-social personality type, that is, a durable set of predispositions that discern volunteers from non-volunteers [3]. According to the different performance fields there are some classification criteria for the voluntary work, such as international cooperation, aids and emergencies, environmental work, cultural, community service and social voluntary enlistment, that it is the kind of voluntary work we are going to study in this paper. Some basic qualifications are required for its success: responsibility and commitment, lack of prejudices, good social skills and empathy [4]. The study of social volunteering has resulted in a number of explanatory theories: for instance, one of the most agreed-upon aspects of volunteer study is that people with higher economic and social position have a tendency to volunteering [5]. Other reasons are related to the sensitivity of people: they develop more empathy and/or emotional intelligence [6]. Other theories explain that volunteers have learned it from their closest social circle, such as, parents, siblings, friends, etc.: theory of socialization and social cognitive learning [7]. Finally, other authors focus on the fact that volunteers live in a social environment that allows them to:

theory of the resource mobilization, social networks.

### 3 Data analysis objectives and hypothesis

We examine volunteering motivation and influence to sort out uncertainties about reasons that make people to help others: influence of individual variables like personality features (e.g. empathy), socialization variables (e.g. parents, friends, special interest groups, etc.) and environmental variables (e.g. chances to join associative organizations); our research is focused on the university context. We have formulated the following hypotheses:

- H1: People come to voluntary work through a socialization process.
- H2: Ability for empathy predisposes them to carry out help activities for other people.
- H3: The personal willingness to accomplish voluntary tasks does not reach a formal commitment with an organization without the influence of external factors (e.g. access to information, advertising campaigns, educational environment, ...).

In order to achieve our outlined objectives it is designed a questionnaire, which allows to discriminate between the answers from the group of volunteers and the group of non-volunteers. With regards to the kind of questions which the designed questionnaire poses, we can point out that all of them are close questions implying either yes/no answers (those referring to personality or socialization) or answers based upon Likert scales (mainly, those referring to empathy and self-acceptance). Moreover, question 1: *Do you belong to a social volunteer group? If not, what is the reason?* is considered to be of vital importance since it functions as a filter question, because it classifies the answers into three groups of people: current volunteers, former volunteers and students who have never volunteered before. In developing the experimental basis, we have taken into account matching with the population universe, being a representative sample of socio-demographic characteristics of the students at the University of Alicante, for example gender, university degree, etc. From the total sample  $N = 1100$  university students representing the whole range of university courses, aged 17-30 years old, extreme cases were selected and it was created an experimental database comprised of 310 youngsters with the following structure:

- University students who belong to a social volunteer group (  $N_1 = 145$ ).
- University students who do not belong to a social volunteer group due to their laziness (  $N_2 = 165$ ).

## 4 Analysis model of the volunteer behaviour by Self-Organizing Maps

The SOM has revealed to be exceptionally successful in arranging high dimensional input data along its two-dimensional output space in such a way that similar inputs are mapped onto neighbouring regions of the map. In other words, the spatial relationship of the input data is preserved as closely as possible within the representation space of the SOM [8]. SOMs are an interesting tool for exploratory data analysis and visualization [9, 10], which map from a high-dimensional input space to a low-dimensional grid, preserving the topology of the data set as closely as possible. Our proposal is to train a SOM (square grid with 8 x 8 neurons) with the answers to an opinion poll obtained from 310 students of the University of Alicante as the input data set and to exploit the SOM capacity to self-classify and to visualize hidden relations between high dimensional data. Each vector consists of 13 components, each one is the numerical answer of a question in the questionnaire. We propose the following specific objectives in our research of the human behaviour modelling about volunteer enlistment in a non-profit organization:

- First one, to analyze whether the SOM is capable of providing, automatically, a clustering of volunteers and non-volunteers data.
- Second one, to study individual questions and identify regions on the SOM map where they are influential.
- Finally, we want to know if it is feasible to answer the question: Is it possible to predict and categorize volunteers correctly, by means of the fill-out forms of opinion polls?.

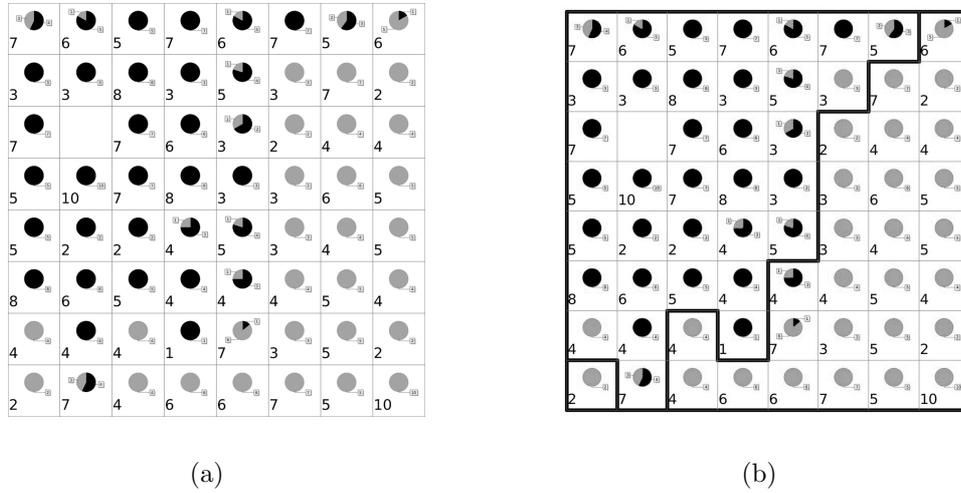


Figure 1: Trained SOM with answers of students. (a) self-organized BMUs, (b) k-means clustering with  $k = 2$ .

## 5 Experiments

In Fig. 1(a), we illustrate the SOM map we have obtained after the training step with data obtained by the questionnaires. As we can see, the SOM self-clustering is unambiguous.

Fig. 1(b) shows an automatic k-means clustering of the questionnaire data map with  $k = 2$  that validates previous observations. In Fig. 2(a), we analyze the influence of the socialization variable, for example as a response to the question: *Have your parents spent any of their spare time by helping other people?* We can see that the non-volunteer zone is darker (answered as *never*) than the volunteer zone (answered by *yes*) which is lighter grey.

In Fig. 2(b), we analyze the influence of another socialization variable as a response to the question 7: *Have your fiends spent any of their spare time by helping other people?* In this case, the component visualization is more impressive; so, this socialization variable is much more sensitive to categorize volunteers and non-volunteers than the previous one, as we can see. Finally, we are going to test the SOM capability to predict if a student is volunteer or not, only by means of that student s answers to the opinion poll and considering that this input vector was not used in the SOM training phase. So, we are going to develop a test of automatic classification for non-trained

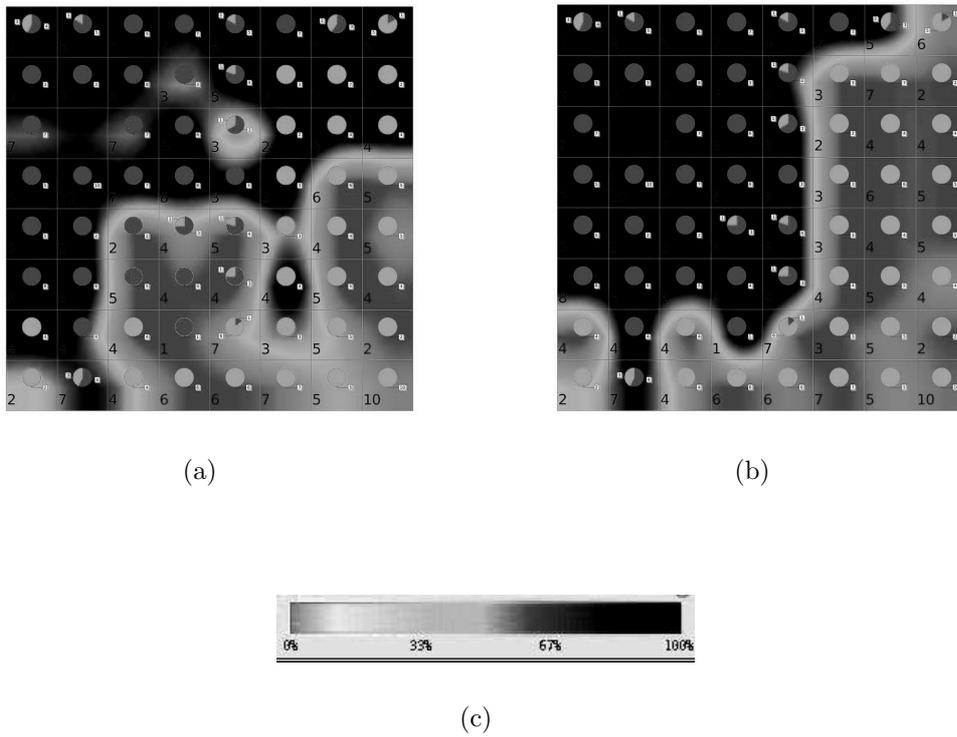


Figure 2: Component visualizations of socialization variables, (a) question 6, (b) question 9, (c) answer legend.

	# Correctly predicted	% Correctly predicted
$SOM_1$	36 of 40	90 %
$SOM_2$	37 of 40	92.5 %
$SOM_3$	34 of 40	85 %
$SOM_4$	33 of 40	82.5 %
$SOM_5$	38 of 40	95 %

Table 1: Automatic SOM classification for non-trained vectors

vectors. We have trained five different SOM with only 270 input vectors selected randomly from the total of 310, with empathy, socialization and environmental variables and the test set consisted of the forty remainders vectors. Table 1 shows the results of the forecasting developed for each trained map.

## 6 Conclusions

In this paper, we have studied aspects of special interest about the profile of the volunteer: empathy, socialization, environment by using connectionist methodology. SOM seems to be a good tool for multidimensional data analysis and to visualize hidden relations between high dimensional data vectors and it will contribute favourably to the development of a selection model and to our capacity of generalization of the influential factors in the volunteer.

## References

- [1] Handy, F., Cnaan, R. A., Brudney, J. L., Ascoli, U., Meijs, L. C. M. P., and Ranade, S., Public perception of "Who is a volunteer": An examination of the net-cost approach from a cross-cultural perspective. *Voluntas*, 11(1), 45-65, 2000.
- [2] Putnam, R. D. *Bowling, Alone: The Collapse and Revival of American Community*, New York, Simon & Schuster, 2000.
- [3] Musick, M. A., and Wilson, J., *Volunteers: A Social Profile*, Bloomington, Indiana University Press, 2008.

- [4] Vecina M. L., Marta E., Pozzi M., Marzana D., What do sustain our will to volunteer? A cross-cultural Italian and Spanish comparison. *Psyche*, 19(2), 19-23, 2010.
- [5] Wilson, J., Volunteering. *Annual Review of Sociology*, 26, 215-240, 2000.
- [6] Goleman, D., *Emotional Intelligence*. New York: Bantam, 1995.
- [7] Bandura, A., *Social Learning Theory*. Englewood Cliffs, Nj: Prentice-Hall, 1977.
- [8] Kohonen T., Data management by self-organizing maps. *Computational Intelligence: Research Frontiers*, Springer, 2008.
- [9] Kohonen T., *Self-organizing Maps*. Springer, Berlin, 1995.
- [10] Yan J., Thill, J.-C., Visual data mining in spatial interaction analysis with self-organizing maps. *Environment and Planning B* 36, pp. 466-486, 2009.

# On Estimating the Position and Direction of a Leader of a Group of Entities

F. Moreno\* and E. Ospina\*

(\*) Universidad Nacional de Colombia

Facultad de Minas, Bloque M8A, Carrera 80 No. 65-223

Medellín, Colombia

November 30, 2011

## 1 Introduction

Movement patterns can be identified when studying a group of moving entities such as a flock of birds, a school of fish, a convoy of vehicles, among others[4][3]. A pattern, known as *leadership* has been recognized in [2]. This pattern is characterized by a moving entity that motivates or represents the behavior of the group in order to reach a goal during a period. A leadership pattern can be defined using some temporal constraints and a geometrical arrangement between the leader entity and the other entities in the group, called *followers*. Given a group of moving entities and a time-point  $t$ , we propose a method to estimate the position and the direction where a leader should be located and headed at  $t$  in order to lead this group.

## 2 Estimating the position and direction of a leader

Our calculations and notation are based in elements of Andersson model [2].

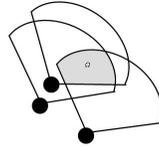


Figure 1: Intersection region  $\Omega$  (in grey)

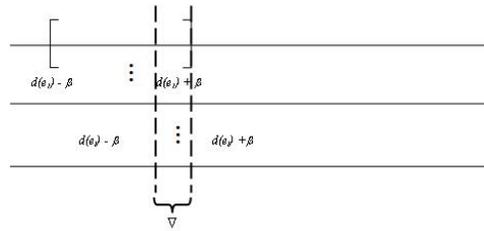


Figure 2: Intersection  $\nabla$  of  $n$  intervals

### 2.1 Estimation of the Leader Position

Let  $(xpos(e_{leader}, t), ypos(e_{leader}, t))$  be the position of a leader entity  $e_{leader}$  at time  $t$  with regard to a group of followers  $F$ . Because each entity  $e_i \in F$  must be a follower of  $e_{leader}$  then  $e_{leader} \in front(e_i) \forall e_i \in F$ . That is, the position of  $e_{leader}$  at  $t$  must be contained in the front region of each  $e_i \in F$ :  $(xpos(e_{leader}, t), ypos(e_{leader}, t)) \in front(e_i), \forall e_i \in F$ . Therefore, if  $e_{leader}$  is a leader at  $t$  for the entities in  $F$  then  $e_{leader}$  must be located in a region  $\Omega$ , *i.e.*, the intersection of the front regions of each entity in  $F$ :  $(xpos(e_{leader}, t), ypos(e_{leader}, t)) \in \Omega$ , see Figure 1.

### 2.2 Estimation of the Direction of a Leader

Let  $d(e_{leader})$  the direction of an entity  $e_{leader}$  at time  $t$ . If an entity  $e_i \in F$  is a follower of  $e_{leader}$  then  $\|d(e_i) - d(e_{leader})\| \leq \beta$ . This condition can be written as  $d(e_{leader}) \in [d(e_i) - \beta, d(e_i) + \beta]$ . That is, if  $e_{leader}$  is a leader at time  $t$  for the set of entities in  $F$  then the direction of  $e_{leader}$  must be included in the interval defined by each entity in  $F$ . Therefore,  $d(e_{leader})$  must be included in the intersection  $\nabla$  of the intervals defined by the entities in  $F$ , see Figure 2: the value of  $d(e_{leader})$  must be included in the interval delimited between the dashed lines.

Let  $Min d$  and  $Max d$  be the minimum and the maximum of the directions

of the entities in  $F$ , respectively. The intersection of these intervals is non-empty if  $Maxd \leq Mind + 2\beta$ , giving rise to an intersection interval  $\nabla$ :  $[Maxd - \beta, Mind + \beta]$ .

### 3 Experiments

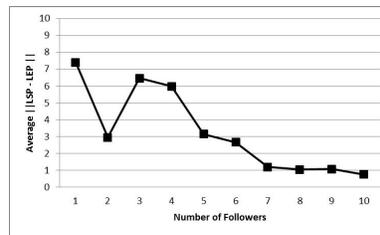
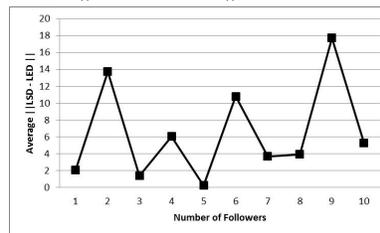
To show the expediency of our proposal, we implemented our algorithms and conducted a set of experiments on various data sets. The algorithms were implemented using Oracle Spatial features [1]. We simulate the data collection using the flocking model built in the Netlogo Library [5]. For each leadership pattern that was found, we applied algorithms to estimate the position and direction of the leader at a specific time-point; then, we compare these results with the simulated ones.

Our results showed how the number of followers affects the estimation of the position and direction of the leader entity. The results showed a better estimation of both parameters when the number of followers increases. This is reasonable because the more followers there are, the greater the number of regions to intersect and  $\Omega$  tends to be smaller. Figure 3 illustrates this aspect, where we plotted the average  $\|LSP - LEP\|$  vs.  $m$  (difference between the estimated and the simulated positions). On the other hand, the estimation of the direction, see Figure 4, did not show in some cases a good estimation due possibly to the early alignment of the entities. That is, in the first time-steps, when some estimations were done, the entities just begin to align and  $\nabla$  tends to be wider. Even so, the differences between the simulated and calculated data were small.

### 4 Conclusions and future work

Our experiments showed the expediency of our algorithms. Although more extensive experiments are needed, our results suggest that as the number of followers increase, the distance between the estimated and simulated position decreases. With regard to the estimation of the direction, we did not get an analogous behavior; even though most of the times the estimations were very close to the simulated ones.

As future work we consider the following. To propose alternative methods for estimating the position and the direction of a leader and compare them

Figure 3: Average  $\|LSP - LEP\|$  vs. Number of FollowersFigure 4: Average  $\|LSP - LEP\|$  vs. Number of Followers

with our current method. Another interesting issue is to conduct experiments to analyze the effect of parameters such as  $r$ ,  $\beta$ , and  $\alpha$  that we kept constant here. Experiments with real data are also a must as well as its comparison with Netlogo simulations.

## References

- [1] Oracle spatial users guide and reference 10g release 2, 2006.
- [2] M. Andersson, J. Gudmundsson, P. Laube, and T. Wolle. Reporting leadership patterns among trajectories. In *Proceedings of the 2007 ACM symposium on Applied computing*, pages 3–7. ACM, 2007.
- [3] M. Benkert, J. Gudmundsson, F. Hübner, and T. Wolle. Reporting flock patterns. *Algorithms-ESA 2006*, pages 660–671, 2006.
- [4] S. Dodge, R. Weibel, and A.K. Lautenschütz. Towards a taxonomy of movement patterns. *Information Visualization*, 7(3):240–252, 2008.
- [5] U. Wilensky. Netlogo flocking model (computer program). *Center for Connected Learning and Computer Based Modeling, Northwestern University*, 1998.

# Electroencephalographic indices of interpolatory type

M.V. Sebastián\* ; M.A. Navascués†

(\*) Centro Universitario de la Defensa. Academia General Militar de Zaragoza,  
Ctra. Huesca s/n, 50090 Zaragoza. Spain

(†) Departamento de Matemática Aplicada. Escuela de Ingeniería y Arquitectura.  
Universidad de Zaragoza,  
C/María de Luna 3, 50018 Zaragoza. Spain

November 30, 2011

## 1 Introduction

Hans Berger discovered in 1924 the oscillations of the electric potential of the human brain, and their trace is called electroencephalogram (EEG) since then. The analysis of computerized EEG is an important tool nowadays in the brain and behavioral sciences.

Our team has proposed numerical quantifiers of the EEG signal, coming from the methodology of fractal mathematics and the theory of approximation. We describe in this communication several one-single channel indices, along with numerical procedures for their computation. In every case, an experiment related to a disorder of attention is developed.

## 2 Fractal Complexity

We have implemented several method of data approximation by means of fractal interpolation functions (FIF), and proved the validity and convergence

---

\*e-mail: msebasti@unizar.es

of the procedures if the sampling frequency is high enough (see for instance [4] and [5]).

Given a set of data  $\{(t_n, x_n)\}_{n=1}^N$ , we consider a system of affinities ([4]) of the plane  $(w_n)_{n=1}^N$  and a transformation such that, for  $A \subseteq \mathbb{R}^2$

$$w(A) = \bigcup_{n=1}^N w_n(A).$$

This system has a global attractor  $G$ . This limit set  $G$  is the graph of a continuous function  $f$  which interpolates the data:  $x_i = f(t_i)$  for  $i = 1, \dots, N$ .

The fractal dimension of the graph of  $f$  constitutes a numerical index measuring the complexity of the signal. The parameter lies between 1 and 2.

The procedure described was applied to the study of the EEG recordings of two samples of children: a healthy control group and a set diagnosed with an Attention Deficit with Hyperactivity Disorder (ADHD). By a mere visual inspection of the EEG, no difference was observed in the patient group.

The children of control group were selected randomly by the teachers. 19 children diagnosed with ADHD were chosen, with an average age of  $9.3 \pm 1.5$ . The sample was compared with a control group of 13 children with similar age ( $9.2 \pm 1.3$ ). For every subject, the following signals were recorded: an EEG of rest with closed eyes, and an EEG during the execution of a test consisting in the recognition of a face different from the others, in series of three.

Six locations of the cortical surface were analyzed, following the 10-20 International System of Jasper: F3, F4, O1, O2, F7, F8. The equipment included filters of 0.18 Hz for low frequencies and 35 Hz for high frequencies. The sampling frequency was 128 Hz. A segment of 30 seconds was analyzed in the second minute of the record.

To compare the EEG of rest with the EEG recorded while the execution of the exercise described, the test of the sign hierarchized of Wilcoxon was used. To compare both groups during the execution of the same task the test of Mann-Whitney was performed.

The parameter undergoes a general increase on the whole cortical surface by the execution of the visual test in the ADHD group, but the difference is only significant in some locations.

In the comparison of the data obtained in the computation of the fractal dimension of the EEG during the test of faces recognition with respect to the rest EEG, in the group of children with ADHD the differences were found in

O2 to the significance level of 0.05, and in F8 at level 0.01, whereas in the control group no difference was found. These variations show the activation of the occipital zone (primary visual area) in the achievement of tasks of visual attention, and an increase in frontal area, responsible for cognitive processes.

Some differences between the group of children diagnosed with ADHD and the healthy control at rest were also found. In F8 at the significance level 0.05 and 0.01, ADHD children show a lower value of the parameter, that confirms the hypothesis of slower resting EEG activity.

### 3 Smooth Complexities

Several groups have employed Hjorth parameters for the EEG processing: Elbert et al. [2], R. F. Byring et al. [1], among others.

The Hjorth parameters are defined in terms of the spectral moments of the signal

$$\begin{aligned} A &= m_0 \\ M &= (m_2/m_0)^{1/2} \\ C &= (m_4/m_2)^{1/2}/(m_2/m_0)^{1/2} \\ C_q &= ((m_{2(q+1)}/m_{2q}) - (m_{2q}/m_{2(q-1)}))^{1/2} \quad \forall q \geq 1 \end{aligned}$$

where

$$m_{2q} = \int_{\mathbb{R}} w^{2q} S(w) dw = \frac{1}{T} \int_I (x^{(q)}(t))^2 dt,$$

$x(t)$  and  $\hat{x}(w)$  are the representation of an EEG signal in the time and frequency (Fourier Transform) domains,  $\hat{x}(w)^*$  the complex conjugate of  $\hat{x}(w)$ , and  $S(w) = \hat{x}(w)\hat{x}(w)^*$ .

For large  $q$ , the computation of  $m_{2q}$  implies a numerical differentiation of high order followed by a numerical integration, and this process may produce considerable computation errors. To avoid these problems we propose the use a method of interpolation by polynomial splines  $\sigma$  of degree  $2q - 1$  to reconstruct the signal. The computation of the quadratures for  $m_{2q}$  can be obtained by means of the following formula:

$$m_{2q} = \frac{1}{T} \int_I (\sigma^{(q)}(t))^2 dt = \frac{1}{T} \left[ (-1)^q \sum_{i=1}^N d_i x_i \right]$$

We applied the previous method for the EEG study of ADHD. The mobility  $M$  is considered an index of the mean frequency of the EEG. It is well

known that the frequency increases with age and thus we have chosen the mobility as a suitable quantifier of the brain maturity.

The parameter  $M$  was computed for the EEG at rest with closed eyes of the same samples of children of the Section 2. The channels processed were F3, F4, P3, P4, O1, O2.

It can be observed that, in general, the ADHD group presents lower values of the mobility at rest. These variations have been confirmed statistically by means of a non-parametric test (Mann-Whitney). Significant differences were found in the frontal area, locations F3 and F4, for a significance level of 0.05.

Concerning the ADHD, the results obtained by other groups indicate the existence of excessive slow activity. The experiments of Mouzé-Amady & Horwat [3] indicate that the mobility is significantly correlated with the mean frequency of the EEG, calculated by means of Fast Fourier Transform. According to these results, the children with deficient attention present lower values of the mobility, that is to say, a "slowness" in the EEG with respect to the healthy children, which confirms the described findings.

## References

- [1] Byring R.F., Salami T.K., Sainio K.O. and Örn H.P. EEG in children with spelling disabilities, *Electroencephalogr. Clin. Neurophysiol.* 79:247–255, 1991.
- [2] Elbert T., Lutzenberger W., Rockstroh B., Berg P. and Soothsayer R. Physical aspects of the EEG in schizophrenics, *Biol. Psychiatry* 32:595–606, 1992.
- [3] Mouzé-Amady M. and Horwat F. Evaluation of Hjorth parameters in forearm surface EMG analysis during an occupational repetitive task, *Electroencephalogr. Clin. Neurophysiol.* 101:181–183, 1996.
- [4] Navascués M.A., Sebastián M.V. Fitting curves by fractal interpolation: An application to the quantification of cognitive brain processes, in: *Thinking in Patterns: Fractals and Related Phenomena in Nature*, World Sci. 2004, pp. 143–154.
- [5] Navascués M.A. and Sebastián M.V. Construction of affine fractal functions close to classical interpolants, *Journal of Comp. Analysis and Appl.* 9(3):271–285, 2007.

# A computational investigation of the ignition and combustion processes of diesel-like dual-fuel sprays

J.J. López, R. Novella\*, A. García and J.F. Winklinger

CMT-Motores Térmicos, Universidad Politécnica de Valencia,

Camino de Vera s/n, 46022 Valencia, Spain

November 30, 2011

## 1 Introduction

Recent research activities in the field of diesel engines have shown the potential to reduce pollutant emissions and improve thermal efficiency by controlling fuel reactivity. However, understanding the impact of blending fuels with different physical and especially chemical properties on diesel-like spray mixing and combustion processes is still a challenge. The present work focuses on modeling a dual fuel spray in diesel-like conditions by means of CFD, comparing different diesel and gasoline blends in terms of ignition characteristics and flame structure.

## 2 Methodology

Three different cases (pure diesel and two diesel-gasoline blends with 80/20 and 60/40 mass fraction ratio) of a single spray in a constant volume vessel were calculated with the dieselFoam solver of the open source CFD code OpenFOAM<sup>®</sup> (Version 1.6). Turbulence was modeled with the standard  $k-\varepsilon$

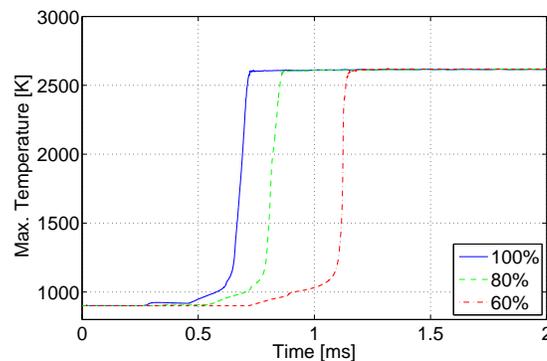
---

\*e-mail: rinoro@mot.upv.es

RANS model. N-heptane and iso-octane were used as surrogates in order to model the chemistry of diesel and gasoline, respectively. A reduced mechanism [1] was applied to describe the chemical kinetics. The numerical setup and the adopted models were adjusted and validated under inert conditions with experimental data from the Spray A case of the ECN [2].

### 3 Results

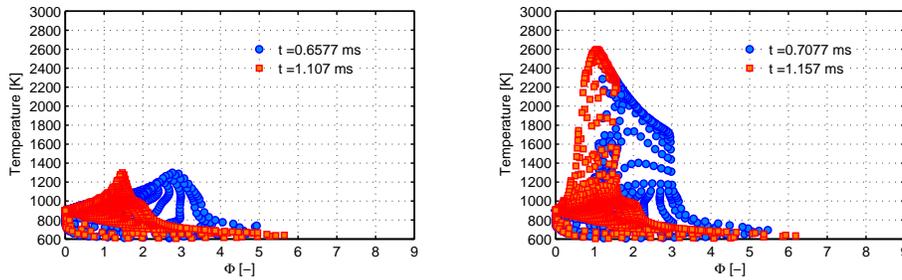
The main objective of this work was to determine the influence of modified fuel reactivity on the spray ignition and combustion processes. Therefore, the results shown in this section are chosen in order to identify the differences found in this investigation.



**Figure 1:** Ignition delay analysis. Temporal evolution of max. temperature

Fig.1 represents the temporal evolution of the maximum temperature in the vessel. Temporal differences in the temperature rise can be observed for the three different fuel blends, where the pure diesel case ("100%") shows the earliest rise of temperature. A progressive increase of ignition delay is observed with decreasing diesel proportion (percentage values in Fig.1) in the blend. This outcome goes along with the difference in ignition delay of pure n-heptane and iso-octane. Moreover, these results confirm the potential of fuel blending to control the reactivity of the spray.

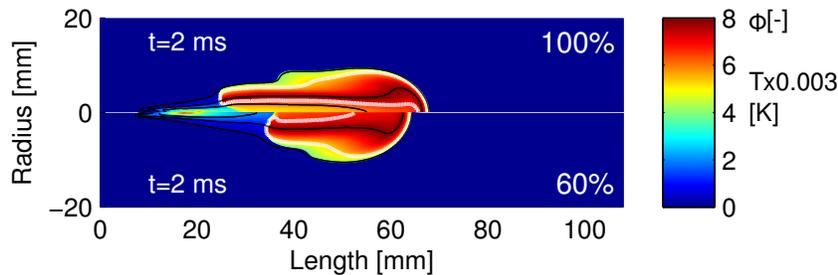
Additionally to the impact on the ignition delay, also an effect on the different levels of equivalence ratios  $\phi$  where combustion occurs can be observed. This fact is demonstrated in the  $\phi$ -T-maps in Fig.2 at start of combustion (SOC) and 0.05ms after SOC (time values stated in the figure are



**Figure 2:** Effect of the fuel proportion on the ignition process. Circles: pure diesel, squares: 60/40 diesel-gasoline blend. Left figure: at SOC, right figure: 0.05ms after SOC

milliseconds after start of injection). The ignition process of the pure diesel fuel spray starts at equivalence ratio values between 2.5 and 3, whereas the 60/40 diesel-gasoline blend ignites at an equivalence ratio of approximately 1.5.

In summary, the results evidence how controlled fuel reactivity, by adjusting blend proportions, introduces temporal and spatial changes in the auto-ignition process. The onset of combustion is shifted toward zones with lower equivalence ratio.

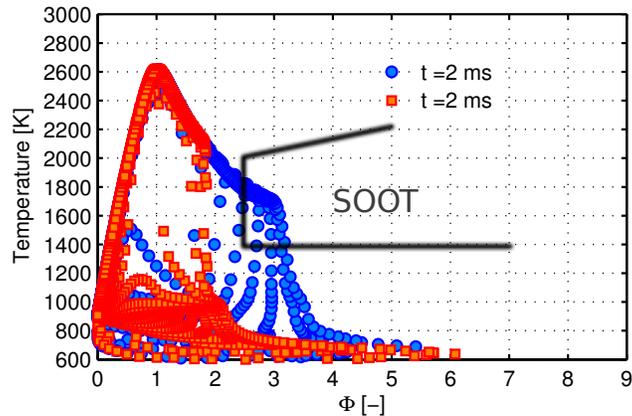


**Figure 3:** Flame structure comparison 2ms after SOI. Pure diesel (top half), 60/40 diesel-gasoline blend (bottom half)

In addition to the auto-ignition process described above, the structure of the lifted turbulent diffusion flame associated to the fuel spray combustion was investigated.

A key parameter of such flames is the lift-off length (LOL). The high interest in controlling the lift-off length arises from its relationship with the pollutant formation during combustion [3]. Several strategies exist to control the lift-off length, such as reducing the ambient temperature, decreasing the injector hole diameter or increasing the injection pressure in order to increase the lift-

off length. However, the implementation of these strategies in real engines is difficult and/or expensive. Therefore, controlling the lift-off length by adjusting the fuel reactivity is a promising approach.



**Figure 4:**  $\phi$ - $T$  map 2ms after SOI and soot region. Pure diesel (circles), 60/40 diesel-gasoline blend (squares)

Approximately 2ms after the start of injection (SOI) a quasi steady-state concerning lift-off length can be observed in all three cases. The CFD simulation results in Fig.3 compare the flame structure of pure diesel with a 60/40 diesel-gasoline blend. The color range represents the equivalence ratio and, additionally, temperature contours in zones with  $T > 3000\text{K}$  are depicted in a plane through the spray axis. The white line plotted in the figure indicates a threshold of OH mass fraction of 0.01% which has been used to determine the lift-off length of the flame as it is done in experimental investigations. A significant difference in LOL between the two depicted cases can be seen. Furthermore, it is important to point out that the reacting zone of the fuel blend does not reach rich mixture regions. On the contrary, in the pure diesel fuel spray some parts of the rich mixture (inner black line marking  $\phi=2$ ) enter the high temperature region.

The consequence of this fact is illustrated in the  $\phi$ - $T$ -map in Fig. 4, where temperature is depicted versus equivalence ratio. Additionally, the soot region is marked in the figure. Both the 60/40 and the 80/20 (not shown here) dual fuel blends show potential to avoid generating a sooting flame during combustion.

## 4 Conclusions

A theoretical investigation based on CFD modeling has been performed with the aim of evaluating the impact of fuel reactivity on the characteristics of a diesel-like reacting spray.

According to the results of this research, the expected dependence of the ignition delay was confirmed. Furthermore, flame lift-off length, which directly influences soot emission formation, was shown to be intrinsically related to fuel reactivity. The final impact on the flame structure was evidenced in terms of the equivalence ratio-temperature maps, since the reacting zones of the spray were shifted toward lower equivalence ratios, avoiding the zones of soot formation.

The suitability of combining state of the art CFD modeling tools together with an adequate chemistry model for reproducing the complex physical and chemical processes involved in dual fuel reacting sprays has been proven. Consequently, CFD modeling provides detailed qualitative spatial information about the local thermo-fluid dynamic conditions inside the spray and their time evolution. The expected trends in key parameters were well described by CFD results.

## Acknowledgement

The authors want to thank Dr. José Manuel Pastor for his support during this work and for sharing his profound knowledge and experience. Support for this research was provided by the Universidad Politécnica de Valencia inside the program named *Programas de Apoyo a la I+D+I, Primeros proyectos de investigación (reference PAID-06-11 2033)*, which is gratefully acknowledged.

## References

- [1] S. Jerzembek, N. Peters, P. Pepiot-Desjardins, H. Pitsch. Laminar burning velocities at high pressure for primary reference fuels and gasoline: Experimental and numerical investigation, *Combustion and Flame*, 156:292-301, 2009
- [2] Spray A Operating Condition, Engine Combustion Network (ECN), <http://www.sandia.gov/ecn/cvdata/sprayA.php>, September 2011

- [3] L.M. Pickett, D.L. Siebers, Soot Formation in Diesel Fuel Jets Near the Lift-off Length, *International Journal of Engine Research*, 7:103-130, 2006

# A one dimensional lumped model for calculation of heat fluxes inside turbochargers

P. Olmeda\*, V. Dolz, F.J. Arnau, and M.A. Reyes-Belmonte

(\*) P. Olmeda. CMT-Motores Térmicos, Universitat Politècnica de València

Camino de Vera s/n, 46022 Valencia, Spain.

Phone: +34 963877650 Fax: +34 963877659 e-mail: pabolgon@mot.upv.es

November 30, 2011

## 1 Introduction

In the present paper a one dimensional lumped model to calculate the heat fluxes inside a turbocharger from diesel passenger car has been presented. This lumped model will be used to determine accurately turbine and compressor efficiency from measurements of a gas stand. Traditionally turbocharger efficiency calculation has been based on the ratio of change of compressor enthalpy and isentropic turbine enthalpy [1] showed in equation (1). Nevertheless this equation provides inaccurate results, mainly at lower speeds and powers, due to the no consideration of the heat fluxes effect and the mechanical losses in the turbocharger axis.

$$\eta_{TG} = \frac{(T_{20} - T_{10}) \cdot c_p \cdot \dot{m}_C}{(T_{30} - T_{4s}) \cdot c_p \cdot \dot{m}_T} = \frac{\dot{W}_C + \dot{Q}_C}{\dot{W}_{Ts} + \dot{Q}_T} \quad (1)$$

## 2 Lumped Model

The proposed one-dimensional lumped model is shown in Figure 1. This model is based on the electrical analogy [2]. As the model shows, the turbocharger has been separated into five different regions. Each of them has

been introduced in the model as a metal node representing the surface temperature. Meanwhile turbine and compressor cases have been represented as single nodes in the lumped model ( $T$  and  $C$ ), the turbocharger housing has been divided into three different regions ( $H_1, H_2, H_3$ ) due to both its complicated geometry and the fact that oil and coolant circuits are placed inside that small element. Metal nodes in the lumped model are connected between them by means of metal conductance and connected with the working fluids by means of convective conductance. Finally, the lumped model includes capacitors in each of the metal nodes in order to represent the accumulation or release of thermal energy of those nodes during a transient evolution. Once lumped model information has been obtained using the methodology described in [3], the model will be used to determine the real power used by the turbine and the compressor.

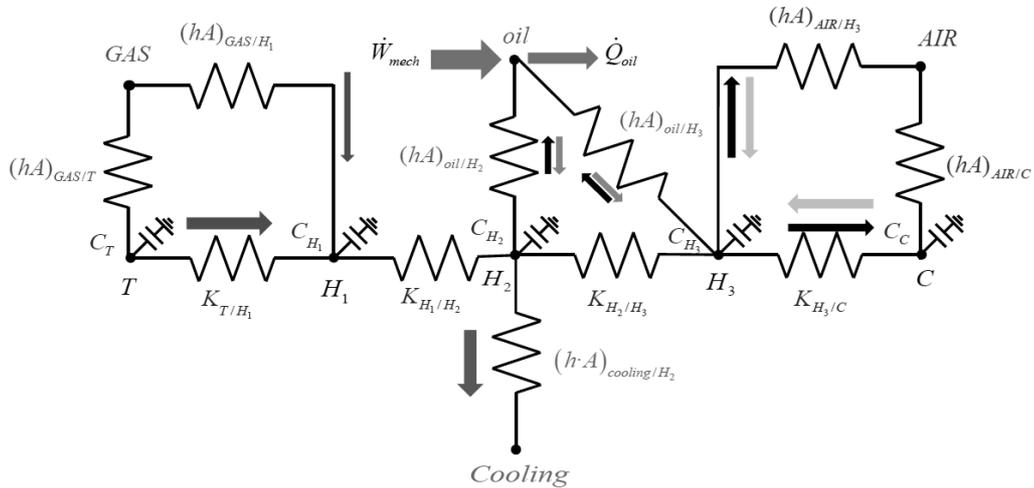


Figure 1: 1-D Lumped Model for the turbocharger. Heat fluxes are represented with arrows.

### 3 Model Application

The system of equations representing the lumped model is shown in matrix form in (2). Where  $\mathbf{K}$  is a  $9 \times 9$  matrix storing the whole information from lumped model. This matrix can be subdivided into four submatrix as indicates equation (3).

$$\mathbf{K} \cdot \mathbf{T} = \mathbf{T}_k \tag{2}$$

$$\mathbf{K} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ (\mathbf{h} \cdot \mathbf{A})_{i/\text{fluid}} & \mathbf{K}_{i/j} \end{bmatrix} \tag{3}$$

Submatrix  $\mathbf{I}$  represents the  $4 \times 4$  identity matrix,  $\mathbf{0}$  is a  $4 \times 5$  zero matrix. Columns from submatrix  $(\mathbf{hA})_{i/\text{fluid}}$  represent each of the working fluids meanwhile there is a row per each of the metal nodes. Therefore, elements from submatrix  $(\mathbf{hA})_{i/\text{fluid}}$  are zero if there is not a physical connection between node  $i$  and fluid  $j$ . Columns and rows from submatrix  $\mathbf{K}_{i/j}$  represent the metal nodes. This submatrix is tridiagonal and symmetric. Elements for the main diagonal are the sum of conductances (conductive and convective) connected with metal node  $i$  but with a minus sign, and elements from upper and lower diagonal are metal conductance between node  $i$  and  $j$ . Finally, vectors of temperatures are,

$$\mathbf{T} = [ T_{T,in} \quad T_{OI} \quad T_W \quad T_{DO} \quad T_T \quad T_{H1} \quad T_{H2} \quad T_{H3} \quad T_C ]^t \tag{4}$$

$$\mathbf{T} = [ T_{T,in} \quad T_{OI} \quad T_W \quad T_{DO} \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 ]^t \tag{5}$$

As the system of equations shows, there are five independent equations (each non-identity line in matrix (3)) meanwhile there are six unknowns (one per each node and one more for the diffuser outlet temperature). In order to solve this system of equations extra information will be obtained from turbocharger maps where turbine, compressor, oil and cooling inlet conditions are known. Besides, it has been assumed that the measured point in the map corresponds to the one defined by inlet conditions and the efficiency from map is estimated and the real mass flow is calculated with equation (6). Since the map was obtained without having in mind the heat fluxes, an iteration is performed until the diffuser outlet temperature is obtained. Then, the aerodynamic efficiency of the compressor will be known. With this efficiency, and iterative process will be followed in order to obtain the rest of temperatures.

$$m_T^{map} = \frac{N_{is,c}^{map}}{\eta_{TG}\eta_C \Delta h_{is,T}^{map}} \tag{6}$$

A study about the importance of heat fluxes for all the operative conditions of the turbocharger has been performed.

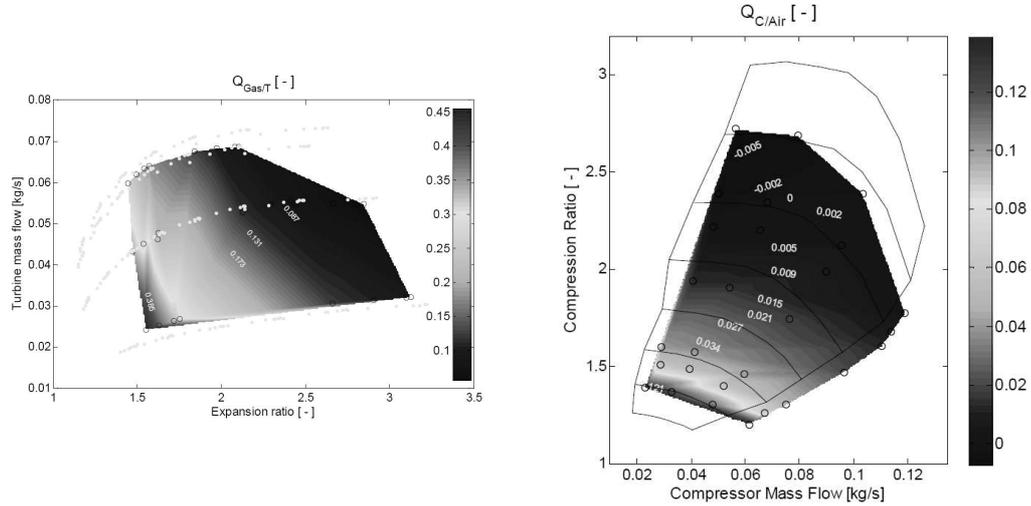


Figure 2: Heat fluxes importance in the turbocharger, a) Heat flux between exhaust gases and turbine case. b) Heat flux between air and compressor case.

This information is presented in Figure 2-(a) where the energy transferred to the turbine case and later transmitted to the rest of the turbocharger is shown. That energy is expressed in dimensionless form dividing it by the exhaust gases energy. As it is shown, an important ratio of the available energy at turbine entrance is not used during the expansion procedure and passes directly to the rest of the turbocharger as a heat flux. These ratios are near 45% for the lowest speeds and loads as it has been said traditionally [4]. Figure 2-(b) shows a similar study but performed at the compressor. As it is represented, the lower is the turbocharger speed the higher is the ratio between the energy passed from the compressor case to its air, compared with compressor power with a ratio around the 10% of the whole compressor energy.

## 4 Conclusions

A one-dimensional lumped model for heat fluxes calculation inside an automotive engine turbocharger has been presented. The whole turbocharger has

been split into different metal nodes. The application of this model into a real turbocharger has shown the importance of internal heat fluxes when the turbocharger is running at lower rotational speeds. Besides, this application allows obtaining adiabatic maps for turbine and compressor using the maps given by turbocharger manufacturers.

## References

- [1] N. Watson, S. Janota, Turbocharging the internal combustion engine, MacMillan Publishers Ltd, London, 1982.
- [2] A. J. Torregrosa, P. Olmeda, J. Martín, C. Romero, A tool for predicting the thermal performance of a diesel engine, *Heat Transfer Engineering* 32(10) (2011) 891–904.
- [3] J.R. Serrano, P. Olmeda, A. Páez, F. Vidal, An experimental procedure to determine heat transfer properties of turbochargers, *Measurement Science and Technology*, 2010 21. doi:10.1088/0957-0233/21/3/035109.
- [4] M. Jung, R. Ford, K. Glover, N. Collings, U. Christen, M. Watts, Parameterization and transient validation of a variable geometry turbocharger for mean-value modelling at low and medium speed-load points, SAE Technical Paper 2002-01-2729.

# Estimation of Diffusivity of Phycobilisomes on Thylakoid Membrane based on spatio-temporal FRAP images

Š. Papáček<sup>†</sup>\*, R. Kaňa<sup>§</sup>, and C. Matonoha<sup>‡</sup>

(†) Institute of Physical Biology, University of South Bohemia,  
Zámek 136, 373 33 Nové Hradky, Czech Republic,

(§) Institute of Microbiology, Academy of Sciences of the Czech Republic,  
Department of phototrophic microorganisms, 379 81 Třeboň, Czech Republic,

(‡) Institute of Computer Science, Academy of Sciences of the Czech Republic,  
Pod Vodárenskou věží 2, 182 07 Prague 8, Czech Republic

November 30, 2011

## 1 Introduction

The organization and dynamics of many photosynthetic pigment-protein complexes in the photosynthetic membrane is frequently studied using fluorescence confocal microscopy by means of Fluorescence Recovery After Photobleaching (FRAP) measuring technique [7]. FRAP has been used since 1970s to study lateral mobility on the cell surface. Later it has been extended to the investigation of protein dynamics within the living cells [6] including thylakoid membrane proteins [4, 3]. Based on spatio-temporal FRAP images, the mobility of photosynthetic complexes in a native intact membrane, i.e. the diffusivity or diffusion coefficient  $D$ ,<sup>1</sup> is reconstructed using either a

---

\*e-mail: papacek@greentech.cz

<sup>1</sup>I. F. Sbalzarini in [6] distinguishes between the molecular diffusion constant and the apparent diffusion constant; while the former is directly measured by single-molecule tech-

*closed form model* or *simulation based model* [5, 2]. The FRAP images are in general very noisy, with small signal to noise ratio (SNR), which requires an adequate technique assuring the reliable results.<sup>2</sup>

Our study describes the development of a method aiming to determine the phycobilisomes diffusivity on thylakoid membrane from FRAP experiments. As we know, this is usually done by experimental curve fitting to the analytical (closed form) models, see e.g. [1, 4, 3]. However, the closed form models need some unrealistic assumptions. For example, C. W. Moulineaux *et al.* [4] have exploited the rotational symmetry of the cells by bleaching a plane across the short axis of the cell. This approach allowed construction of bleaching profiles along the long axis. Supposing: (i) the infinite domain, (ii) the initial bleaching profile is Gaussian:  $y(x, t_0) = y_{0,0} \exp \frac{-2x^2}{r_0^2}$ , where  $r_0$  is the half-width of the bleach at time  $t_0 = 0$ ,  $y_{0,0}$  is the maximum depth at time  $t_0$ , i.e. the depth ( $y_{0,0} < 0$ ) of the first post-bleach signal at the center ( $x = 0$ ), and (iii) boundary conditions correspond to the complete recovery:  $y \rightarrow 0$  as  $t \rightarrow \infty$ ,  $y \rightarrow 0$  as  $x \rightarrow \infty$ ; then the solution  $y(x, t)$  of diffusion equation  $\frac{\partial y}{\partial t} = D \frac{\partial^2 y}{\partial x^2}$  and the maximum depth at time  $t$ , i.e.  $y(0, t)$  are as follows:

$$y(x, t) = \frac{y_{0,0} r_0}{\sqrt{r_0^2 + 8Dt}} \exp \frac{-2x^2}{r_0^2 + 8Dt}, \quad y(0, t) = \frac{y_{0,0} r_0}{\sqrt{r_0^2 + 8Dt}}.$$

The calculation of diffusion coefficient  $D$  then resides in the weighted linear regression: a plot of  $(\frac{y_{0,0}}{y(0,t)})^2$  against time should give a straight line with the tangent  $\frac{8D}{r_0^2}$ .

As the analytical approach has several limitations (e.g. restriction to the cell geometry, full recovery is required, bleach profile must be gaussian-like, etc.) we model the process by the Fickian diffusion equation (with realistic initial and boundary conditions) instead. The estimation of diffusivity is further formulated as an optimization problem consisting in the minimization of an objective function representing the disparity between the experimental and simulated time-varying concentration profiles.

---

niques, the latter is determined by coarse-grained methods such as FRAP, averaging over a certain observation volume.

<sup>2</sup>Let us mention that the fluorescence confocal microscope allows the selection of a thin cross-section of the sample by rejecting the information coming from the out-of-focus planes. However, the small energy level emitted by the fluorophore and the amplification performed by the photon detector introduces a measurement noise.

## 2 Problem formulation

FRAP (Fluorescence Recovery After Photobleaching) technique allows detection of diffusivity of autofluorescence compound like proteins (e.g. phycobiliproteins) and also other non-fluorescence compound that are fluorescently tagged (e.g. green fluorescence proteins - GFP). This method is based on application of short, intense laser irradiation (the so called bleach) to a small target region (Region Of Interest - ROI) of the cell that causes irreversible loss in fluorescence in this area without any damage in intracellular structures. After the "bleach" (or "bleaching"), the observed recovery in fluorescence in the "bleached area" reflects diffusion of fluorescence compounds from the area outside the bleach. For an arbitrary bleach spot and assuming (i) local homogeneity, i.e. assuring that the concentration profile of fluorescent particles is smooth, (ii) isotropy, i.e. diffusion coefficient is space-invariant, (iii) an unrestricted supply of unbleached particles outside of the target region, i.e. assuring the complete recovery,<sup>3</sup> the unbleached particle concentration  $C$  as a function of spatial coordinate  $\vec{r}$  and time  $t$  is modeled with the following diffusion-reaction equation on two-dimensional domain  $\Omega$ :

$$\frac{\partial C}{\partial t} - \nabla \cdot (D \nabla C) = R(C) , \quad (1)$$

where  $D$  is the fluorescent particle diffusivity within the domain  $\Omega$  and  $R(C)$  is a reaction term. The initial condition and time varying Dirichlet boundary conditions are:

$$C_0 = f(\vec{r}, t_0) \text{ in } \Omega, \quad C(t) = g(\vec{r}, t) \text{ in } \partial\Omega \times [t_0, T]. \quad (2)$$

The reaction term  $R(C)$  is often viewed as negligible under assumptions that diffusion of fluorescence compounds (proteins) is not restricted (e.g. by some binding to the medium) and that photobleaching of these molecules during recovery is negligible. Consequently, if  $R(C)$  is neglected, Eq. (1) becomes the Fickian diffusion equation. In contrast, under continual photobleaching during image acquisition, this reaction term could be described as a first order reaction:

$$R(C) = -k_S C , \quad (3)$$

where  $k_S$  is a rate constant describing bleaching during scanning [2].

---

<sup>3</sup>The recovery is not always complete. It is usually modelled by introducing some correction term. More consistent method resides in the special time dependent Neumann boundary condition in form of a saturation curve.

## References

- [1] D. Axelrod, D. E. Koppel, J. Schlessinger, E. Elson, W. W. Webb, Mobility measurement by analysis of fluorescence photobleaching recovery kinetics, *Biophys. J.* 16 (1976) 1055-1069.
- [2] O. N. Irrechukwu, M. E. Levenston, Improved Estimation of Solute Diffusivity Through Numerical Analysis of FRAP Experiments, *Cellular and Molecular Bioengineering* 2 (2009) 104–117.
- [3] R. Kaňa , O. Prášil, C. W. Mullineaux, Immobility of phycobilins in the thylakoid lumen of a cryptophyte suggests that protein diffusion in the lumen is very restricted, *FEBS letters*, 583 (2009) 670–674.
- [4] C. W. Mullineaux, M. J. Tobin, G. R. Jones, Mobility of photosynthetic complexes in thylakoid membranes, *Nature* 390 (1997) 421-424.
- [5] F. Mueller, D. Mazza, T. J. Stasevich, J. G. McNally, FRAP and kinetic modeling in the analysis of nuclear protein dynamics: what do we really know?, *Current Opinion in Cell Biology* 22 (2010) 1-9.
- [6] Ivo F. Sbalzarini, *Analysis, Modeling and Simulation of Diffusion Processes in Cell Biology*, VDM Verlag Dr. Muller, 2009.
- [7] J. Thomas, W.W. Webb, Fluorescence Photobleaching Recovery: A Probe of Membrane Dynamics, in *Non-Invasive Techniques in Cell Biology*, Ed(s) S. Grinstein and K. Foskett, Wiley-Liss, Inc., 129–152, 1990.

# Study of the influence of the inlet boundary conditions in a LES simulation of internal flow in a diesel injector

Raul Payri<sup>†</sup>, Jaime Gimeno<sup>†</sup>\*,  
Pedro Marti-Aldaravi<sup>†</sup>, and Gabriela Bracho<sup>‡</sup>  
(<sup>†</sup>) CMT - Motores Térmicos, Universidad Politécnica de Valencia,  
Edificio 6D, 46022, Valencia, Spain,  
(<sup>‡</sup>) GE Global Research Europe,  
Freisinger Landstrasse 50 D-85748 Garching bei Munchen, Germany,

November 30, 2011

## 1 Introduction

Realistic boundary conditions are necessary in numerical simulations in order to get validated results. In computational fluid dynamics it is widely accepted that the fluid behaviour is determined in large part by the inlet behaviour, especially in turbulent flows. For Reynolds Averaged Navier-Stokes (RANS) simulations, only mean profiles of velocity or pressure and some turbulent variables need to be defined at the inlet, but for Large-Eddy (LES) and Direct-Numerical (DNS) Simulations the turbulent inflow conditions have to be prescribed [2].

The power of LES for modelling real systems has been proved by R. Payri et al. [3]. However, no comparison of fluid behaviour in real systems with different inlet boundary conditions has been found [1]. Thus, in the present work the main idea is to simulate the internal flow of a diesel injector in order to check the accuracy of three different already available inlet conditions.

---

\*e-mail: jaigigar@mot.upv.es

## 2 Methodology

LES code [7, 15] is applied to an industrial process, such a modern common-rail nozzle injector in diesel engines [10, 11]. The pressure gradient in these systems is very high and the fluid experiences a restriction, which could affect in the flow in a way the boundary layer detaches and the flow along the hole becomes turbulent. The study is focused just to the liquid phase inside nozzle (no cavitation is considered) [12], and the main goal is to find the most suitable inlet boundary condition.

The nozzle geometry is the same that used in previous works [3, 5]. It corresponds to an axi-symmetric nozzle manufactured specially for research purposes. The orifice has an outlet diameter of 112  $\mu\text{m}$  and 1 mm of length. The shape is convergent, thus the cavitation is “avoided” [6] and the assumption that simulation involves only incompressible liquid is validated. The computational mesh used in this study is the same that was used in previous works [3, 5].

The conditions are characterized by two inlet pressures, 120 MPa and 160 MPa, correspondig to medium and high diesel injection pressures and an outlet pressure of 5 MPa. Isothermal conditions are assumed and the temperature used for the simulations was 298 K. The working fluid was winter diesel fuel with a constant density of 825.28  $\text{kg}/\text{m}^3$  and also constant viscosity of  $2.80 \cdot 10^{-3} \text{ kg}/(\text{m} \cdot \text{s})$  [9].

The Reynolds number based on the theoretical Bernouilli velocity and the exit diameter is 17400 for medium injection pressure case and 20200 for high injection pressure case. Both values are high enough to expect the flow to be turbulent. Before the calculation it was unknown if the turbulence is fully developed because the hole could not be long enough.

The boundary conditions used in the simulations are described below. Inlet conditions are detailed in the next subsection.

- Walls: non-slip velocity condition.
- Outlet: uniform pressure condition was used, and zero gradient velocity setup [5].

### 2.1 Inlet conditions

Three different inlet conditions were tested. According to [1], the inlet boudary should:

- be stochastically varying;
- ... on scales down to the filter scale (spatially and temporally);
- be compatible with Navier-Stokes equations;
- “look” like turbulence;
- allow the easy specification of turbulent properties;
- and be easy to implement and to adjust to new inlet conditions.

However, the chosen inlet conditions do not meet all of these requirements. They are:

- Constant pressure: uniform pressure condition and zero gradient velocity setup.
- Turbulent inlet: similar to the previous condition, but a random Gaussian noise is added over pressure variable (with an amplitude of 5% [16, 8]). It does not reproduce the coherent structures and the oscillations could be quickly damped.
- Mapped inlet (or direct mapped): internal mapping condition in which a precursor domain is integrated in the main domain (leading to higher computational cost).

## 2.2 Procedure

The governing equations (filtered Navier-Stokes equations) are solved using the finite volume CFD code OpenFOAM 1.5 [8], which employs temporal and spatial discretisation schemes that are bounded and preserve the proper physical limits on the fluid-dynamics variables. The solution procedure employs the *Oodles* solver, which is a generic single-phase incompressible LES solver.

### 3 Results and discussion

#### 3.1 Time evolution

A comparison of the instantaneous velocity trend is made. Only the stabilized time will be shown, defined as the simulated time at which the oscillations have appeared and the flow is stabilized in time. The previous trend is always the same. Table 1 summarizes this time for every simulation. It can be seen, in both cases, that the turbulent inlet aids the oscillations to appear. Direct mapped inlet stabilizes the flow even in less time, but this does not mean mapped inlet is the fastest option, the domain is more than double so the computing time will be larger (more than double).

Case name	Stabilization time [ms]
1200_simple	60.6395
1200_turbulent	43.969
1200_mapped	23.865
1600_simple	64.9329
1600_turbulent	24.978
1600_mapped	27.565

Table 1: Stabilization time for every simulation made.

Figure 1 shows the velocity time evolution for 120 MPa injection pressure case. Figure 1a shows the expected result, the turbulence is generated at the restriction (entrance of the hole) and developed along the hole. So, there are no oscillations at the entry section, some at the middle and more at the exit. The behaviour with the turbulent inlet is the same, shown in Figure 1b, but some oscillations can be observed at the entry. This is because of the artificial oscillations that the boundary is imposing close to the entry section. Nevertheless, no oscillations at all can be seen for the mapped inlet simulation. No turbulence is generated with this boundary. The same fluid behaviour is observed for the high injection pressure case.

#### 3.2 Validation using experimental results

The simulation results are compared against equivalent experimental data performed in [6]. The parameter used for comparison purposes is  $C_v$ , as de-

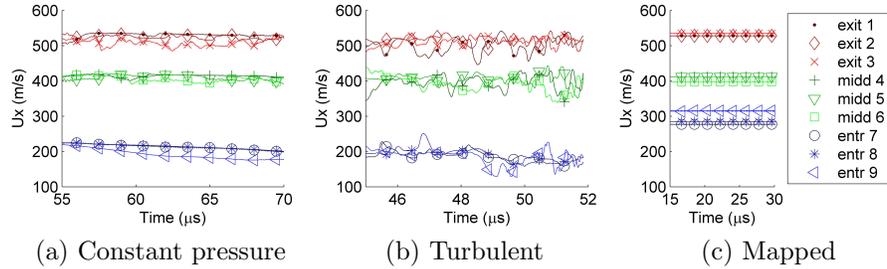


Figure 1: Velocity time evolution for stabilized flow in three different points of three different sections for 120 MPa injection pressure case.

finied by [14]. The comparison is summarized in Table 2. As can be seen, constant pressure inlet estimates the experimental value very well meanwhile turbulent boundary overestimates and direct mapped underestimates it slightly. With this result any inlet condition can be dismissed, the three of them are in a range narrow enough to say that they are right, especially considering the error in the experimental measurement.

Case name	$C_v$ [-]	$C_v _{exp}$ [-]	Deviation [%]
1200_simple	0.902	0.907	-0.569
1200_turbulent	0.912	0.907	0.619
1200_mapped	0.888	0.907	-2.034
1600_simple	0.908	0.909	-0.166
1600_turbulent	0.964	0.909	6.025
1600_mapped	0.909	0.909	-0.002

Table 2: Simulated and experimental velocity coefficient values and deviation between them.

## 4 Conclusions

There are big differences in the turbulent fluid behaviour, but none of the inlet has been better for internal flow in diesel injectors:

- by using a constant pressure inlet, velocity and oscillations similar to the experimentally obtained are calculated, but the required time is high and the inlet does not have the structure of the turbulence;
- by using a turbulent inlet condition, the required time decreases significantly, though the calculus is more unstable and the fluctuations do not have the coherent structure of the turbulence neither;
- and by using the mapped inlet condition the calculus is stable and less simulated time is required to get stabilized flow, but no turbulence is generated in a reasonable time.

However, considering the presented results the next solution to improve the simulation is purposed:

1. using the constant pressure inlet until the exit velocity reaches values close to the real ones,
2. then change the inlet condition to turbulent inlet to generate oscillations faster,
3. and finally switch to a direct mapped to stabilize the flow, obtain real turbulent structures and simulate more time.

This combination will be case dependent.

## References

- [1] G. R. Tabor and M. H. Baba-Ahmadi. Inlet conditions for large eddy simulation: a review *Computer & Fluids*, 39:553–567, 2009.
- [2] N. Jarrin and S. Benhamadouche and D. Laurence and R. Prosser. A synthetic-eddy-method for generating inflow conditions for large-eddy simulations *International Journal of Heat and Fluid Flow*, 27:585–593, 2006.
- [3] R. Payri and B. Tormos and J. Gimeno and G. Bracho. The potential of Large-Eddy Simulation (LES) code for the modeling of flow in diesel injectors *Mathematical and Computer Modelling*, 52:1151–1160, 2010.

- [4] C. Fureby and G. Tabor and H. G. Weller and A. D. Gosman. A comparative study of subgrid scale models in homogeneous isotropic turbulence *Physics of Fluids*, 9:1151–1160, 1997.
- [5] R. Payri and B. Tormos and J. Gimeno and G. Bracho. Large Eddy Simulation for high pressure flows: model extension for compressible liquids *Mathematical and Computer Modelling*, 54:1725–1731, 2011.
- [6] R. Payri and F. J. Salvador and J. Gimeno and A. García. Flow regime effects over non-cavitating diesel injection nozzles *Proceedings of the Institution of Mechanical Engineers, Part D, Journal of Automobile Engineering*, In Press. Corrected Proof, doi:10.1177/0954407011413056, 2011.
- [7] U. Piomelli. Large-eddy simulation: achievements and challenges *Progress in Aerospace Sciences*, 332–362, 1999.
- [8] OpenFOAM Ltd.. Title of the paper *FOAM - The complete guide*, <http://www.opencfd.co.uk>.
- [9] R. Payri and F. J. Salvador and J. Gimeno and G. Bracho. The effect of temperature and pressure on thermodynamic properties of diesel and biodiesel fuels *Fuel*, 90:1172–1180, 2011.
- [10] R. Payri and F. J. Salvador and J. Gimeno and J. de la Morena. Influence of injector technology on injection and combustion development - Part 1: Hydraulic characterization *Applied Energy*, 88:1064–1074, 2010.
- [11] J. Benajes and R. Payri and S. Molina and A. H. Plazas. Investigation of the influence of injection rate shaping on the spray characteristics in a diesel common rail system equipped with a piston amplifier *Journal of Fluids Engineering-Transactions of the ASME*, 127:1102–1110, 2005.
- [12] R. Payri and F. J. Salvador and J. Gimeno and J. de la Morena. Study of cavitation phenomena based on a technique for visualizing bubbles in a liquid pressurized chamber *International Journal of Heat and Fluid Flow*, 30:768–777, 2009.
- [13] F. J. Salvador and J. V. Romero and M. D. Roselló and J. Martínez-López. Validation of a code for modeling cavitation phenomena in Diesel

injector nozzles *Mathematical and Computer Modelling*, 52:1123–1132, 2010.

- [14] R. Payri and J. M. García and F. J. Salvador and J. Gimeno. Using spray momentum flux measurements to understand the influence of diesel nozzle geometry on spray characteristics *Fuel*, 84:551-561, 2005.
- [15] P. Sagaut, Large Eddy Simulation for Incompressible Flows. Berlin, Springer, 2001.
- [16] Eugene de Villiers, The potential of Large Eddy Simulation for the Modeling of Wall Bounded Flows. Department of Mechanical Engineering, Imperial College of Science, Technology and Medicine, 2006.

# Leadership groups on Social Network Sites based on personalized PageRank

F. Pedroche\*, F. Moreno†, A. González†, and A. Valencia†

(\*) Institut de Matemàtica Multidisciplinària Universitat Politècnica de València

Camí de Vera s/n. 46022 València. Espanya.

(†) Escuela de Sistemas, Universidad Nacional de Colombia, Sede Medellín

Carrera 80 No 65-223. Medellín. Colombia.

November 30, 2011

## 1 Introduction

Some usual centrality measures (degree, betweenness, etc.) can be used to assign importance to users in a Social Network Site (SNS). We use the concept of PageRank [1] since it has proved to be of utility in some fields, apart of being in the core of the searcher Google.

The framework presented is based on the concept of *Leadership group* recently introduced by one of the authors [2]. In particular, we show how to analyze the structure of the *Leadership group* as a function of a single parameter.

We classify the users of an SNS based on the Personalized PageRank (PPR) vector. PPR is PR when using some prescribed *personalization vector*. PPR was originally introduced to bias PR to some nodes [3]. See, e.g., [4] for an analytical formulation. An example of how to use topics of the queries to bias the PageRank can be found in [5].

We are interesting in classifying the nodes of a network considering the direct graph of the network and the features of the nodes. PPR is used to

---

\*e-mail:pedroche@imm.upv.es. This work is supported by Spanish DGI grant MTM2010-18674.

include some features of the nodes. The method presented allows to give an extra of PR to some nodes in a controlled way. In this paper we call a *leader* a node that has higher PR than the others.

## 2 Preliminaries

Let  $\mathcal{G} = (\mathcal{N}, \mathcal{E})$  be the directed graph representing a Social Network Site. Users are represented by the set of nodes  $\mathcal{N} = \{1, 2, \dots, n\}$  and the set of directed links is  $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$ . The link represented by the pair  $(i, j)$  belongs to the set  $\mathcal{E}$  if and only if there exists a link pointing from node  $i$  to node  $j$ . In this paper we assume that each node has at least one outlink.

Let  $0 < \alpha < 1$  be the damping factor (that we use as  $\alpha = 0.85$ ). Let  $\mathbf{e} \in \mathbb{R}^{n \times 1}$  be the vector of all ones and let  $\mathbf{v}$  be the personalization (or teleportation) vector, i.e.,  $\mathbf{v} = (v_i) \in \mathbb{R}^{n \times 1} : v_i > 0$  for all  $i \in \mathcal{N}$  and  $\mathbf{v}^T \mathbf{e} = 1$ . The Google matrix is defined as  $G = \alpha P + (1 - \alpha) \mathbf{e} \mathbf{v}^T$ , and is an stochastic and primitive (irreducible and aperiodic) matrix [3]. The PageRank vector is defined as the unique left Perron vector of  $G$ , that is:  $\pi^T = \pi^T G$ , with  $\pi^T \mathbf{e} = 1$ . Denoting  $\mathbf{e}_i$  the  $i$ th column of the identity matrix of order  $n$ , the PageRank of a node  $i$  is  $\pi_i = \pi^T \mathbf{e}_i$ . We call basic PageRank, and denote it by *basic PR* to the vector  $\pi(\mathbf{e}/n)$ . We call *basic leader* a node that is in the top of the *basic PR*.

## 3 Leadership group

The fundamentals of the model were presented in [2], where the concept of *Leadership group*,  $\mathcal{L}$ , was introduced. The following two definitions constitute the framework that allows classify users using PPR.

**Definition 1.** Given a directed graph  $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ , let  $0 < \epsilon \leq \frac{n-1}{n}$  and let  $\mathbf{v}_i(\epsilon) = [v_{ij}] \in \mathbb{R}^{n \times 1} : v_{ii} = 1 - \epsilon, v_{ij} = \epsilon/(n-1)$  if  $i \neq j$ . For each  $i \in \mathcal{N}$ , let  $PR_i = \pi(\mathbf{v}_i)$ . and we denote as  $(PR_i)_j$  the  $j$ th entry of  $PR_i$ .

Note that  $(PR_i)_j$  represents the value of the PR corresponding to node  $j$  when using the personalization vector  $\mathbf{v}_i(\epsilon)$ . This is the centrality measure that we use.

**Definition 2.** Given a directed graph  $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ , and  $0 < \epsilon \leq \frac{n-1}{n}$ , the Leadership group,  $\mathcal{L} \subseteq \mathcal{N}$  is defined as follows:  $j \in \mathcal{L}$  if, for some  $i \in \mathcal{N}$  it holds that

$$(PR_i)_j \geq (PR_i)_k \text{ for all } k \neq j. \tag{1}$$

i.e. for some personalization vector  $\mathbf{v}_i(\epsilon)$ , node  $j$  has the greatest PageRank. The number of different indices  $i \in \mathcal{N}$  for which (1) occurs is called the frequency of node  $j$  in  $\mathcal{L}$ , and we denote it as  $\nu_{\mathcal{L}}(j)$ .

### 3.1 Example

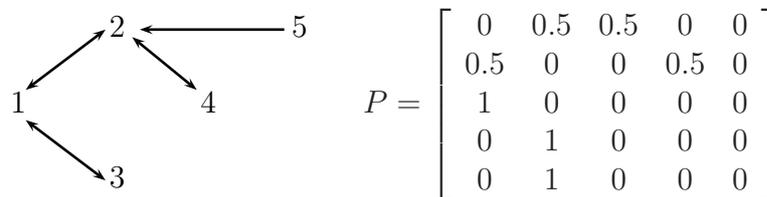


Figure 1: A graph and its corresponding row stochastic matrix

Node $j$	$PR_1$	$PR_2$	$PR_3$	$PR_4$	$PR_5$	$\nu_{\mathcal{L}}(j)$
1	<b>0.38</b>	0.29	<b>0.34</b>	0.26	0.26	2
2	0.30	<b>0.39</b>	0.27	<b>0.35</b>	<b>0.35</b>	3
3	0.17	0.13	0.25	0.12	0.12	0
4	0.14	0.18	0.13	0.25	0.16	0
5	0.01	0.01	0.01	0.01	0.11	0

Table 1:  $PR_i$ , and  $\nu_{\mathcal{L}}(j)$  for the graph of Fig. 1, with  $\epsilon = 0.3$ .

In Table 1 we show the elements of the vector  $PR_i$ , for  $\epsilon = 0.3$ , for the graph shown in Fig. 1. Note that in this example changing  $\mathbf{v}_i(\epsilon)$  we obtain a different ranking. That is why we claim that we use PPR depending on two parameters ( $i$  and  $\epsilon$ ) as a centrality measure. In this example we have  $\mathcal{L} = \{1, 2\}$  and  $\nu_{\mathcal{L}}(1) = 2$  and  $\nu_{\mathcal{L}}(2) = 3$ . Note also that, giving a node  $i$ , the maximum value of the PPR for that node is obtained when computing  $PR_i$ ; Note also that there are nodes such as nodes 3, 4 or 5 that do not win even though we bias the PPR to them, for the  $\epsilon$  considered. Computing the same experiment for some values of  $\epsilon$  we find that for  $\epsilon \leq 0.68$  we have that

$\mathcal{L} = \{1, 2\}$  while for  $\epsilon \geq 0.69$  we have that  $\mathcal{L} = \{1\}$ . Therefore, in this graph the ranking, the structure of  $\mathcal{L}$  and  $\nu_{\mathcal{L}}(j)$  depend on  $\epsilon$ .

In this framework we can define the probability to be a leader in the following form.

**Definition 3.** Given a directed graph  $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ , and  $0 < \epsilon \leq \frac{n-1}{n}$ , let  $\mathcal{L}$  and  $\nu_{\mathcal{L}}(j)$  given by Definition 2. We define the probability of node  $i \in \mathcal{N}$  to be a leader as  $P(i, \epsilon) = \frac{\nu_{\mathcal{L}}(i, \epsilon)}{\sum_{j \in \mathcal{N}} \nu_{\mathcal{L}}(j, \epsilon)}$ .

## 4 Conclusions

We have presented a theoretical framework for classifying users in SNSs. We have shown that the Personalized PageRank can be used as a centrality measure depending on two parameters. We have shown how to analyze the Leadership group using a single parameter  $\epsilon$ . We have introduced the measure  $\nu_{\mathcal{L}}(j)$  that can be also used as a centrality measure. We have introduced the probability to be a leader in this framework.

## References

- [1] L. Page, S. Brin, R. Motwani, T. Winograd, The PageRank Citation Ranking: Bringing Order to the Web, Stanford Digital Library Technologies Project, 1999.
- [2] F. Pedroche, Competitivity Groups on Social Network Sites, Mathematical and Computer Modelling, 52 (2010), p. 1052-1057.
- [3] A. N. Langville, C. D. Meyer. Google's Pagerank and Beyond: The Science of Search Engine Rankings, Princeton University Press, 2006.
- [4] T. Haveliwala, S. Kamvar, G. Jeh. An Analytical Comparison of Approaches to Personalizing PageRank, Technical Report. Stanford. 2003.
- [5] T. H. Haveliwala, Topic-sensitive PageRank: A context-sensitive ranking algorithm for web search, IEEE Transactions on knowledge and data engineering, vol. 15, No. 4. 2003.

# On the computation of output bounds on parallel inputs pharmacokinetic models with parametric uncertainty

Diego de Pereda<sup>\*,†</sup>, Sergio Romero-Vivo<sup>‡</sup>, and Jorge Bondia<sup>†</sup>

(†) Institut Universitari d'Automàtica i Informàtica Industrial,

(‡) Institut Universitari de Matemàtica Multidisciplinar,

Universitat Politècnica de València, Camino de Vera S/N - 46022 Valencia.

November 30, 2011

## 1 Introduction

Compartmental modelling is a common approach to simulate biological processes. Furthermore, these models have been used in many diverse areas such as economics, engineering, medicine, or human sciences. In particular, many models have been developed to study pharmacokinetic processes, such as the examples analysed in this paper.

This work is focused on the parallel inputs model [1], which is based on a suggestion given by *Jacquez*, who considered that the single-peak concentration-time response that usually follows the oral administration of a drug could be modelled using a single linear chain of identical compartments connected together in series. A similar approach is based on two or more parallel linear chains connected to the output compartment with elimination rate  $k_e$ , as shown in Figure 1. Each chain is formed by a number (that can vary for each chain) of identical compartments. Due to the possible different pathways to reach the output, this type of model is usually used to analyse pharma-

---

\*e-mail: dpereda@upvnet.upv.es

cokinetic processes in which inputs show effects on the output with different delays in time.

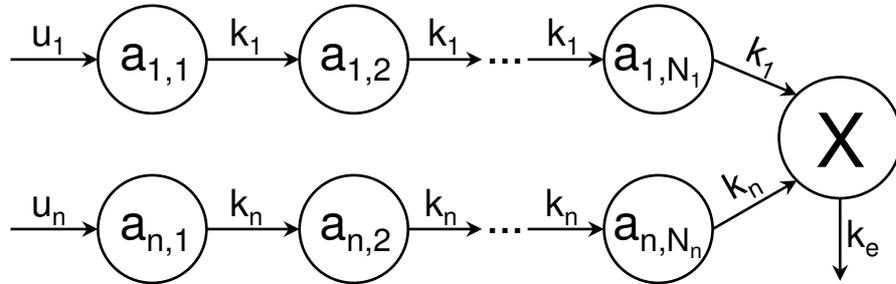


Figure 1: *Diagram of the parallel inputs model.*

When studying a biological process with a compartmental model there is always some mismatch between the model and real life, caused because mathematical models are a simplified version of the actual process. This mismatch yields non-modelled dynamics. Furthermore, a common characteristic of biological processes is variability, leading to parametric uncertainty. The exact values of the model parameters and initial conditions are unknown, but they can be bounded by intervals. There is just one solution for constant parameters, but parametric uncertainty yields a set of different possible solutions. The computation of an output envelope must guarantee the inclusion of all the possible solutions for the model.

The aim of this work is to propose a new method to minimise the over-estimation of output bounds on the parallel inputs model.

## 2 The computation of output bounds

Different approaches have been applied to compute the set of possible solutions when interval parametric uncertainty is considered. These approaches have been compared in two cases: subcutaneous insulin absorption for artificial pancreas research, and the study of the double-peak phenomenon observed for certain drugs [2]. The results for this second example are shown in Figure 2.

Montecarlo methods have traditionally been used to deal with uncertainty. However, they are not applicable for computing output bounds be-

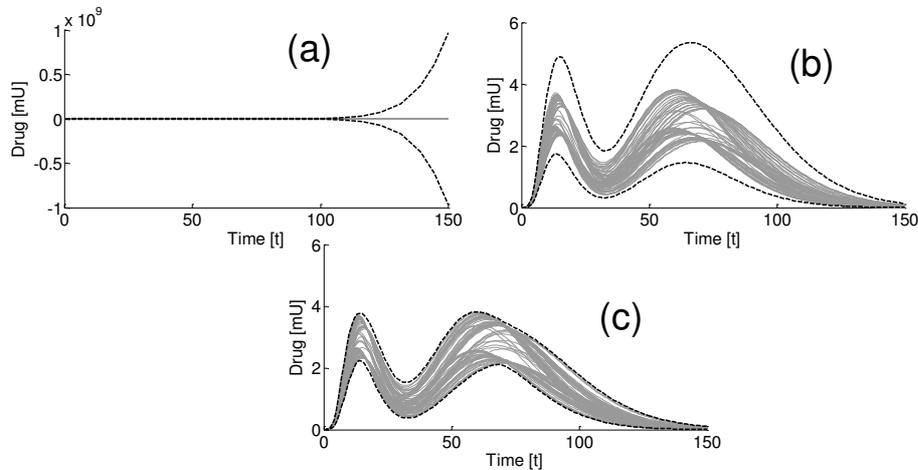


Figure 2: *Improvements computing double-peak phenomenon bounds. (a) VNODE-LP computation. (b) Monotonicity analysis of the parallel inputs model. (c) Monotonicity analysis of the equivalent one-compartment model.*

cause, independently of the number of simulations executed, the bounds obtained do not guarantee the inclusion of all the possible solutions. This inclusion guarantee is a priority when a model is used as part of a medical decision system, as is the case here, since the output bounds must ensure that all the possible responses by the patient are outside risk levels. Other areas where this is needed include error-bounded parametric identification and constraint-satisfaction problems. Furthermore, the computational cost of Montecarlo methods increases proportionally to the number of simulations performed to cover the uncertain input space sufficiently. For these reasons, Montecarlo simulations have just been used to measure the overestimation produced by the different methods.

Region-based approaches have been considered for the output bounds computation, using VNODE-LP [3] software. This C++ solver computes guaranteed output bounds, but the error obtained produces a large overestimation, as seen in Figure 2a.

The most common method in the literature to reduce overestimation is to perform a monotonicity analysis of a trajectory-based approach [4]. After this analysis is performed, only non-monotone compartments or parameters produce an overestimation in the output bounds computation. In the parallel

inputs model, the only non-monotone parameters are  $k_i$ . An overestimation is produced due to these parameters, as seen in Figure 2b, but this is much smaller than the error obtained with the VNODE-LP method.

Finally, our proposal consists in obtaining an equivalent model combining analytical solutions of the input chains with a monotonicity analysis. This approach allows computing critical points of the non-monotone parameter  $k_i$ , which helps to compute tighter output bounds, as seen in Figure 2c. Furthermore, the computational cost is just 0.02 seconds to obtain the solution bounds.

Our proposed method outperforms previous approaches for the computation of output bounds on the parallel inputs model, as it minimises the overestimation produced, and also because of its low computational cost.

## Acknowledgements

This work was partially supported by the Spanish Ministerio de Ciencia e Innovación through Grant DPI-2010-20764-C02, and by the Universitat Politècnica de València through Grant PAID-05-09-4334.

## References

- [1] J. Jacquez, Compartmental analysis in biology and medicine. BioMedware, 1996.
- [2] K. Godfrey, P. Arundel, Z. Dong, and R. Bryant. Modelling the Double Peak Phenomenon in pharmacokinetics. *Computer Methods and Programs in Biomedicine*, 104(2):62-69, 2011.
- [3] N. Nedialkov. VNODE-LP: A validated solver for initial value problems in ordinary differential equations. *Technical Report CAS-06-06-NN*, Department of Computing and Software, McMaster University, Hamilton, Ontario, Canada, L8S 4K1, 2006.
- [4] E. Sontag. Monotone and near-monotone biochemical networks. *Systems and Synthetic Biology*, 1(2):59-87, 2007.

# A chemical species transport model with variable gas properties for gas dynamic calculation of internal combustion engines

J. R. Serrano, H. Climent, P. Piqueras\* and O. García-Afonso  
CMT-Motores Térmicos. Universitat Politècnica de València.

November 30, 2011

## 1 Introduction

The use of aftertreatment systems in internal combustion engines has become a standard in recent years [1]. The inclusion of these systems in 1D gas dynamics codes for internal combustion engine modelling [2] demands the development of a chemical species transport model able to consider the influence of chemical species conversion along the engine flow-path on emissions and flow thermal properties.

This work presents a chemical species transport model to account for variable composition and gas properties along the engine flow path. The performance for chemical species transport in 1D elements of shock-capturing methods, such as the two-step Lax&Wendroff method and the Sweby's TVD scheme, is analysed by means of shock-tube tests. The influence of the fluid modelling as perfect or non-perfect gas on the numerical methods features and the flow characteristics on shock-tube results is analysed.

---

\*Camino de Vera s/n, 46022 Valencia, Spain. e-mail: pedpicab@mot.upv.es

## 2 Chemical species transport model

The governing equations for one-dimensional unsteady compressible non-homentropic flow, i.e. the mass, momentum and energy conservation equations, form a hyperbolic system of partial differential equations [3]. Additionally, the solution of the chemical species transport along the 1D elements requires  $n - 1$  equations of chemical species conservation in the governing equations system, where  $n$  is the number of chemical species to be transported. The chemical species conservation equation in vector form is

$$\frac{\partial(\rho \mathbf{Y} F)}{\partial t} + \frac{\partial(\rho u \mathbf{Y} F)}{\partial x} = \rho F \dot{\mathbf{Y}} \quad (1)$$

where  $\mathbf{Y}$  is a vector including the mass fraction of  $n - 1$  different chemical species. The mass fraction of the chemical species  $n$  is given by the following compatibility equation:

$$Y_n = 1 - \sum_{j=1}^{n-1} Y_j \quad (2)$$

The transport of chemical species in 1D elements can be performed solving the governing equations systems by means of finite difference numerical methods, which need to be adapted. In the case of the finite difference numerical methods applied in this work, the following considerations have been taken into account:

- In the case of the two-steps Lax&Wendroff method [4], the  $n - 1$  equations of chemical species transport are discretized in the same way than the mass conservation equation. The only difference is that both the solution and the flux terms have to be multiplied by the mass fraction of the chemical species  $i$ ,  $Y_i$ .
- In the case of the Sweby's TVD scheme [5], the transport of chemical species requires the use of the vector formulation of the Sweby's TVD scheme to solve the mass, momentum and energy conservation equations, whereas every chemical species conservation equation is solved applying the scalar formulation.

### 3 Results and discussion

This section is focused on the evaluation of the two-step Lax&Wendroff method and the Sweby's TVD scheme to solve the chemical species transport. These methods are applied to shock-tube tests derived from the Riemann problem.

#### Case I: Thermal and composition contact discontinuity

Case I is a shock-tube test consisting of a duct with a diaphragm separating two regions where the flow has different conditions in pressure, temperature and composition. When the diaphragm is removed, the initial conditions generate a shock wave and a strong thermal and composition contact discontinuity.

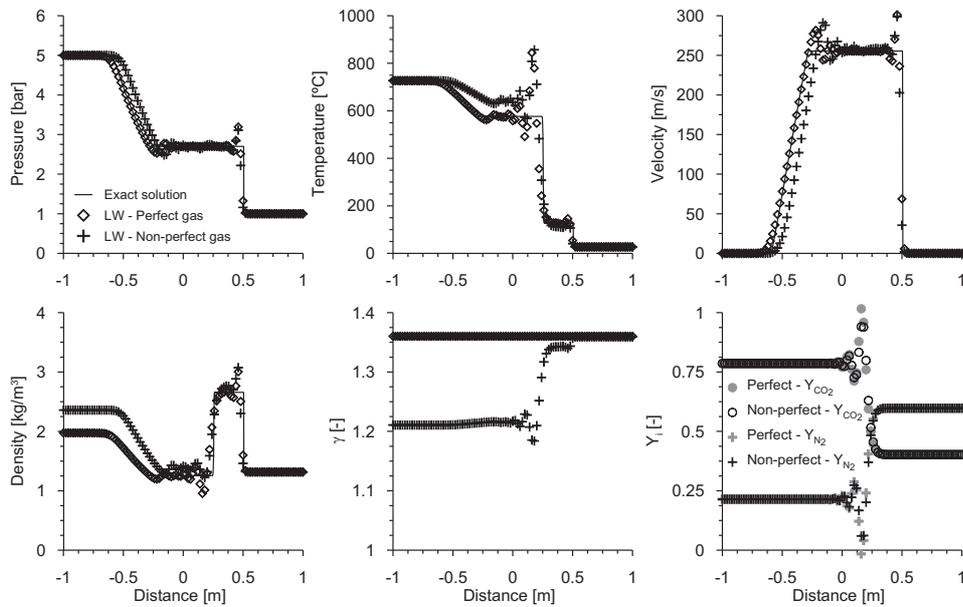


Figure 1: Case I. Results of the two-step Lax&Wendroff method with chemical species transport. Time  $t = 0.001s$ .

Figure 1 shows the results corresponding to the two-step Lax&Wendroff method. As common in symmetric shock-capturing methods of second or

higher order, the solution is dispersive and spurious around the discontinuities. This behavior appears also in the transport of chemical species.

Regarding the numerical solution considering non-perfect gas, it is worth to note that the main differences with respect to the exact solution appear in the left side, since the exact solution has been obtained with the fluid properties of the right side. The differences affect mainly the propagation velocity of the rarefaction wave, since it is the discontinuity more exposed to fluid properties differences with respect to the exact solution. A slight influence appears also in the region between the contact discontinuity and the shock wave because of the temperature effect.

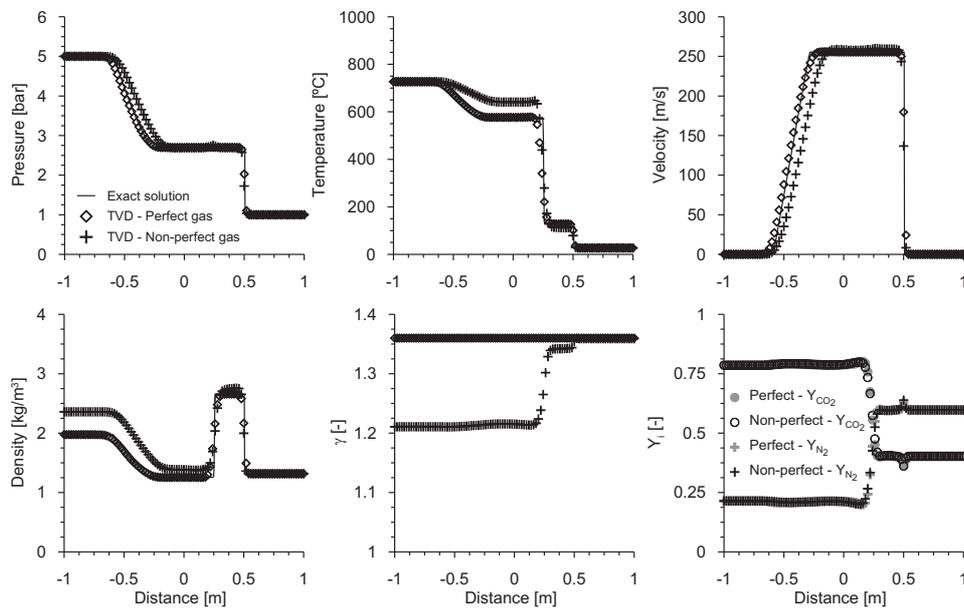


Figure 2: Case I. Results of the Sweby's TVD scheme with chemical species transport. Time  $t = 0.001s$ .

The same simulations have been performed with the Sweby's TVD scheme. As shown in Figure 2, the use of this method allows removing the spurious oscillations around the discontinuities, independently of the perfect or non-perfect gas assumption. However, this approach involves numerical diffusion around the discontinuities. It affects the mass fraction solution mainly in the region of the contact discontinuity but influences also in the proximity of the shock wave and in the end of the rarefaction wave.

### Case II: Contact discontinuity in ICE

This case is devoted to evaluate the magnitude of the influence that the non-perfect gas assumption has on the solution of a contact discontinuity between a wave of exhaust gas and fresh air in internal combustion engines.

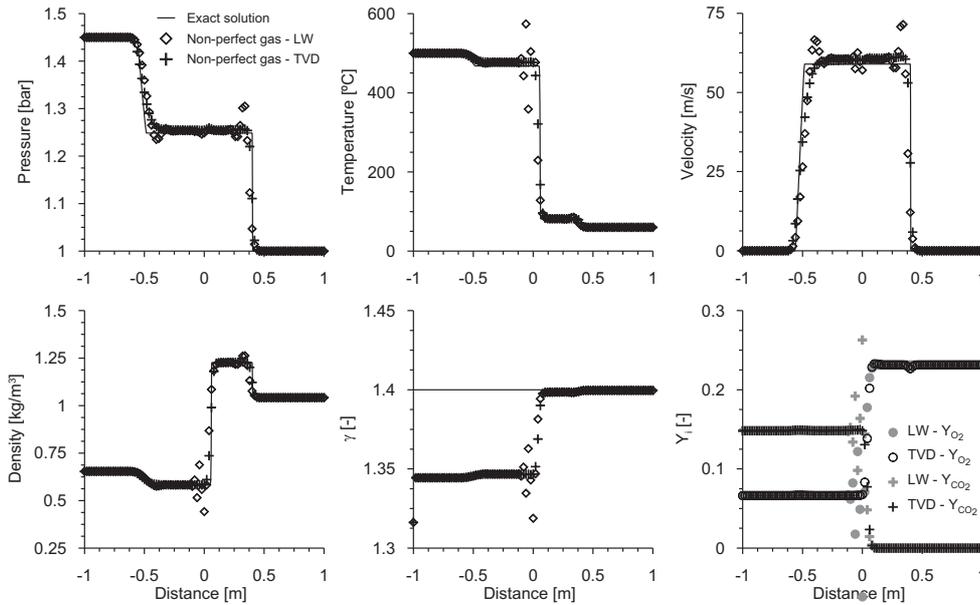


Figure 3: Case II. Chemical species transport considering variable gas properties. Time  $t = 0.001$  s.

The results shown in Figure 3 evidence that the variable gas properties calculation brings noticeable differences in temperature, pressure and velocity prediction in thermal and composition contact discontinuities typical from ICE. The variations affects the propagation velocity of shock and rarefaction waves and the position of the contact discontinuity.

## 4 Summary and conclusions

The paper presents a chemical species transport model for gas dynamic calculation and an analysis of the influence on the solution of representative shock capturing methods. The two-step Lax&Wendroff method has shown typical

spurious oscillations of second order symmetric schemes. On the other hand, the Sweby's TVD scheme mitigates this effect but introduces numerical diffusion around the discontinuities. Influence of the rarefaction and shock waves has been also identified on the chemical species transport.

The transport of the chemical species has additionally allowed the solution of the governing equations with the assumption of non-perfect gas. The shock-tube tests results has evidenced that variable gas properties as function of temperature and composition must considered for a suitable prediction of wave propagation velocity at typical ICE operating conditions.

## Acknowledgements

This work has been partially supported by the spanish Ministerio de Ciencia e Innovación through grant number DPI2010-20891-C02-02.

## References

- [1] V. Bermúdez, J. Serrano, P. Piqueras and O. García-Afonso. Assessment by means of gas dynamic modelling of a pre-turbo diesel particulate filter configuration in a turbocharged hsd diesel engine under full-load transient operation, *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering* 225(9):1134–1155.
- [2] A. Torregrosa, J. Serrano, F. Arnau and P. Piqueras. A fluid dynamic model for unsteady compressible flow in wall-flow diesel particulate filters, *Energy* 36(1):671–684, 2011.
- [3] J. Serrano, F. Arnau, P. Piqueras, A. Onorati and G. Montenegro. 1D gas dynamic modelling of mass conservation in engine duct systems with thermal contact discontinuities, *Mathematical and Computer Modelling* 49(5-6):1078–1088, 2009.
- [4] P. Lax and B. Wendroff. Systems of conservation laws, *Communications on Pure and Applied Mathematics*, 17:381–398, 1964.
- [5] P. Sweby. High resolution schemes using flux limiters for hyperbolic conservation laws, *SIAM Journal of Numerical Analysis* 21(5):995–1011, 1984.

# Optimal control of long-term sentimental dynamics

José-Manuel Rey\*

Universidad Complutense de Madrid

Departamento de Análisis Económico

Campus de Somosaguas

28223 Madrid, Spain.,

November 30, 2011

## 1 Introduction

This contribution is concerned with love dynamics. In particular, we consider committed relationships intended to endure. Given the current grim statistics about divorce, a fundamental problem consist of understanding what makes some couples successful while others are miserable. This is an old question already posed by psychologist Louis Terman in the 1930s [1].

A further contemporary problem arises from sociological evidence. On one hand, there is an epidemic of failure among couple relationships in the western world. This is evidenced in many surveys and studies, e.g. [?]. Also, there is evidence that most people consider a long-term relationship as a main ingredient of happiness and, furthermore, most of them would consider their own relationship stable [4, 5]. These two facts together pose what can be called the *failure paradox*: how is it that, while most couples are ready to commit to a lasting happy relationship, a massive proportion of them will fail?

---

\*e-mail: j-man@ccee.ucm.es

This note seeks to contribute to understand the problem of long-term sentimental relationships via a mathematical model based in optimal control theory.

## 2 The (Adam and Eve) model

The model makes three simple assumptions, that may be considered plausible in the case of Adam and Eve's love story.

The first assumption is *homogamy*: the couple itself is the planning unit. In particular, no interaction dynamics within the couple is considered. This strong assumption amounts to considering the partners as similar. This seems valid in the West, where people choose their long-term partners among those who are similar to them (see e.g. [6]). A weaker form of similarity is needed here: partners share the parameters and functional structure defined below.

The second assumption is the *second law of thermodynamics of marital relationship*: there is a natural decay of the sentimental dynamics, that can be counteracted with effort. This fact is widely recognized by scholars and therapists and it was formulated as a law by Gottman et al [1]. The second law can be written as a linear differential equation that describes the decay of the variable  $x(t) \geq 0$  describing the state of the relationship (the *feeling*) plus a recovery term that comes from the effort making, whose level is given by the variable  $c(t) \geq 0$  (the *effort*). The equation reads

$$\frac{dx}{dt} = -rx + ac, \quad (1)$$

where  $r > 0$  measures the strength of the feeling fading inertia and  $a > 0$  represents effort efficiency.

The third assumption concerns *well-being valuation*: couples evaluate their relationship through a cost-benefit approach. Feeling is something good that produces satisfaction but in a decreasing fashion. So feeling satisfaction is described by some increasing and concave function  $U(x)$  that goes flat when feeling becomes large, i.e.  $U'(x) > 0$ ,  $U''(x) < 0$ , and  $U'(x) \rightarrow 0$  as  $x \rightarrow +\infty$ . Bad things, however, escalate. Since high levels of effort are considered a bad thing, effort discomfort  $D(c)$  is assumed to be increasing and convex beyond certain level  $c^* \geq 0$ . However, effort levels below  $c^*$  may have a positive effect on well-being, so that partners enjoy making the effort provided it is not excessive. It is thus assumed that  $D''(c) > 0$ ,  $D'(c^*) = 0$ , and  $D'(c) \rightarrow +\infty$  as  $c \rightarrow +\infty$ .

The assumptions above allows us to formulate the couple's plan for their relationship as an optimal control problem: the couple determines the effort path that makes them as happy as possible and such that the entailed feeling-effort dynamics is sustainable in the long-term. Total happiness is expressed as the aggregate net well-being during the non-terminating life of the relationship, i.e.

$$\int_0^{\infty} \exp(-\rho t) (U(x) - D(c)) dt.$$

The instantaneous satisfaction is balanced with an exponential term that represents temporal preference in a standard way.

### 3 Results

Optimal control theory [7] produces the fundamental equation for the optimal effort dynamics:

$$\frac{dc}{dt} = \frac{1}{D''(c)} ((\rho + r)D'(c) - aU'(x)). \quad (2)$$

It is apparent from above that the favorite effort level  $c^*$  is never a solution of the couple's problem.

- *The back room analysis.* The analysis of the coupled feeling-effort dynamics (1)-(2) produces significant features. First, there is a unique sentimental equilibrium  $E = (x^\heartsuit, c^\heartsuit)$ .  $E$  is admissible if  $x^\heartsuit > x_{\min}$ , where  $x_{\min}$  is a certain threshold value below which the relationship is not considered satisfactory.

A key finding is that the required effort at equilibrium  $c^\heartsuit$  is always bigger than the effort level preferred a priori  $c^*$ . The existence of this effort gap represents an unexpected difficulty for the viability of the relationship: even if the feeling is rewarding,  $E$  may not be viable if the size of the gap is important.

The global dynamics determines the evolution of the sentimental dynamics, which is typically expected to start at a high feeling level  $x_0 \gg x_{\min}$ . The phase space turns out to be that of a nonlinear saddle. In particular, the feeling-effort dynamics is unstable. Thus, a

long-term sustainable dynamics is only possible if the initial effort  $c(0)$  is fine-tuned to get onto the stable manifold  $W^-$  –say at a point  $A$ – and then the trajectory of  $A$ ,  $\gamma_A \subset W^-$ , is followed until  $E$  is reached. Other trajectories are either not admissible or lead to feeling extinction. Again there is an effort gap along  $\gamma_A$ : the required effort level to make the transition from  $A$  to  $E$  is always above  $c^*$ . Thus, lasting relationships are possible: if  $E$  is rewarding enough, if the effort gap is tolerable, if instability is resisted along the way to  $E$ , if the couple is alert to watch the correct effort permanently.

- *The front room analysis.* Things look different from the front room, where real relationships take place. While lasting relationships are possible they are somehow exceptional. A successful long-term story seems very demanding, namely: equilibrium must be found and the problem solved to find  $\gamma_A \subset W^-$  –and there is only one such successful trajectory–; the equilibrium  $E$  must be viable, i.e.  $x^\heartsuit > x_{\min}$ ; effort stress must be tolerable along the path  $\gamma_A$ ; effort instability must be resisted along  $\gamma_A$ ; and the effort gap that still persists if  $E$  is reached must remain monitored.
- As a consequence, there are apparent reasons not to be optimistic about the future of a relationship in the front room.

## 4 Conclusions

Some key sources of instability revealed by optimal control theory explain the endangered future of a committed relationship. The analysis suggests a general robust mechanism that may be acting behind the breakdown of many couple relationships in realistic scenarios. It is the combination of two effects –the effort gap plus the instability of the dynamics–, provoking that when effort is relaxed, the feeling is driven far from the successful path towards non-admissible levels.

The theory also offers a possible explanation for one mystery in the field of marital therapy [1]: the fact that couples that went through therapy relapse after some time. The model suggests that, if therapy restores the state of things back to the stable path, instability and the effort stress may operate together to take the relationship out of the successful track again after some time.

The model study also produces a mechanistic explanation of the failure paradox, as a consequence of the couple's planning behavior and the second law. The theory so gives meaning to a paradoxical statement: Our love was to last forever. Hence we are now breaking up.

## References

- [1] J.M. Gottman, J.D. Murray, C.C Swanson, R. Tyson, K.R. Swanson, The mathematics of marriage-Dynamic nonlinear models, The M.I.T. Press, 2002.
- [2] Martin T.C., Bumpass L., Recent trends in marital disruption, *Demography*, 26: 37-51, 1989.
- [3] Bramlett M.D. and Mosher W.D. , Cohabitation, marriage, divorce, and remarriage in the United States, *National Center for Health Statistics, Vital Health Stat.* 23(22), 2002.
- [4] Sweet J.A. and Bumpass L.L. The National Survey of Families and Households -Waves 1,2, and 3: Data Description and Documentation. Center for Demography and Ecology, University of Wisconsin-Madison, 2002.
- [5] Pettijohn T.F. II and Pettijohn T.F. Perceived happiness of college students measured by Maslow's hierarchy of needs. *Psychological Reports*, 79: 759-62, 1996.
- [6] Buston P.M. and Emlen S.T. Cognitive processes underlying human mate choice: the relationship between self-perception and mate preference in Western society, *Proc. Natl. Acad. Sci. USA*, 100(15): 8805-8810, 2003.
- [7] A. Seierstad and K. Sydsæter, Optimal Control Theory with Economic Applications, (2nd reprint), North-Holland, 1999.

# Applying a simulation model to manage waiting lists for hospital inpatient activity in a EU region

Francisco Reyes-Santías <sup>\*†</sup>      Carmen Cadarso-Suárez <sup>‡</sup>

Adela Martínez-Calvo <sup>§</sup>

November 30, 2011

## 1 Introduction

The goal of this manuscript is to simulate both the daily inpatient activity and the length of the stay patterns in the Galician NHS hospitals. To achieve this aim, the main idea is to fit a distribution to each variable which allows us to generate new values for them by means of a Monte-Carlo procedure. Once we produce these new observations, we can also create the corresponding simulated waiting list and the occupancy rate (the daily percentage of occupied beds).

We also analyze how the number of hospital beds affects inpatient activity, length of stay and, consequently, waiting list. For that, the simulation was repeated for several increased percentages in the number of available hospital beds for two cases: when number of beds does not affect to inpatient activity (Case 1), and when the number of beds affects to inpatient activity (Case 2).

## 2 Methodology

Our working data are the daily inpatient activity and the length of stays of patients admitted to 14 NHS Galician Hospitals in 2009. During this period, hospitals range a number between 80 and 1300 beds. Due to differences between the inpatient activity in working days and holidays, this dataset was split into

---

<sup>\*</sup>Corresponding author at: Instituto Universitario de Ciencias Neurológicas (IUCINE), University of Santiago de Compostela, Spain. Tel.: +34 678487167. E-mail address: Francisco.Reyes.Santias@sergas.es (F.R. Santías).

<sup>†</sup>Departamento de Organización de Empresas, University of Vigo, Spain. Instituto Universitario de Ciencias Neurológicas (IUCINE), University of Santiago de Compostela; Unidad de Epidemiología Clínica, Hospital Clínico de Santiago de Compostela.

<sup>‡</sup>Unit of Biostatistics, Department of Statistics and Operations Research, University of Santiago de Compostela; Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS).

<sup>§</sup>Department of Statistics and Operations Research, University of Santiago de Compostela; Centre of Supercomputing of Galicia - CESGA.

two groups: Monday to Friday, except July/August/December days (working days); and Saturdays, Sundays, and July/August/December days (holidays).

**Simulating the daily inpatient activity.** Assuming that data are independent observations of a Normal variable, one can estimate the mean  $\mu_{wd}$  and the standard deviation  $\sigma_{wd}$  for working days ( $\mu_{hd}$  and  $\sigma_{hd}$  for holidays) by means of the function *fitdistr* included in the R package *MASS* (see [1]). Thus, new observations can be drawn from  $\mathcal{N}(\hat{\mu}_{wd}, \hat{\sigma}_{wd})$  or  $\mathcal{N}(\hat{\mu}_{hd}, \hat{\sigma}_{hd})$  by means of the R function *rnorm* included in the *stats* package (see [1]).

**Simulating length of stay.** We have assumed that the variable is continuous, and we have chosen an approach based on kernel density estimation (see [3]). If  $x_1, x_2, \dots, x_n$  are the observed length of stay, the density  $f$  of the variable can be estimated by  $\hat{f}_h(x) = (nh)^{-1} \sum_{i=1}^n K(h^{-1}(x - x_i))$ , where  $K$  is a kernel function, and  $h$  is a positive parameter called bandwidth. Specifically, we have considered as  $K$  the standard normal density function, and  $h$  is estimated using the function *density* in the R *stats* package (see [1]). Then, a new value  $x^*$  can be generated as follows: (1) select randomly  $x \in \{x_1, \dots, x_n\}$ ; (2) generate  $z \sim N(0, 1)$ ; (3) build  $x^* = x + \hat{h} * z$ , being  $\hat{h}$  the estimated bandwidth for  $\hat{f}_h$ .

**Simulating hospital activity.** Now we are ready to simulate hospital activity during a given period as follows: (1) generate an inpatient activity value, taking into account both new patients and waiting list patients; (2) detect the number of free beds and occupy them with patients for which generate the length of their stays; and (3) if the daily inpatient activity exceeds the available number of beds, include the remaining patients in the waiting list.

**‘Elasticity’ term: the beds effect.** We want to analyze what happens if the number of beds is modified in two cases. In Case 1, the daily inpatient activity is simulated according to the procedure seen before, with parameters estimated regardless of hospital beds. However, for Case 2, we have fitted a linear regression model where the number of beds is the covariate and the median (or median absolute deviation) of daily inpatient activity is the response. Hence, we have obtained that  $\hat{\mu}_{wd} = 3.9619 + 0.0972 \times beds$ ,  $\hat{\sigma}_{wd} = 2.3747 + 0.0165 \times beds$ ,  $\hat{\mu}_{hd} = 4.2362 + 0.0643 \times beds$ , and  $\hat{\sigma}_{hd} = 2.2984 + 0.0181 \times beds$  for holidays. Thus, the number of beds has a ‘linear’ effect in the inpatient activity pattern.

### 3 Main results

For a year, we have simulated the variables of interest for three hospitals (C1, C2 and C3), each of them belong to one of the three clusters presented by Reyes [2] which group hospitals depending on the number of specialities and the type of services they may offer (high, average and low case-mix). Regarding the hospital beds, four situations were considered: the real number of beds (0%), and three

Case	Hospital	0%	5%	10%	20%
No beds effect	C1	93.50	89.10	84.93	77.89
	C2	98.22	93.64	89.33	81.90
	C3	90.78	86.14	82.98	75.56
Beds effect	C1	93.69	93.63	93.43	93.12
	C2	92.21	91.83	91.70	91.40
	C3	100.00	99.99	99.98	99.75

Table 1: Cases 1 & 2. Mean of simulated daily occupancy rates for C1, C2 and C3 for a year period (from January to December).

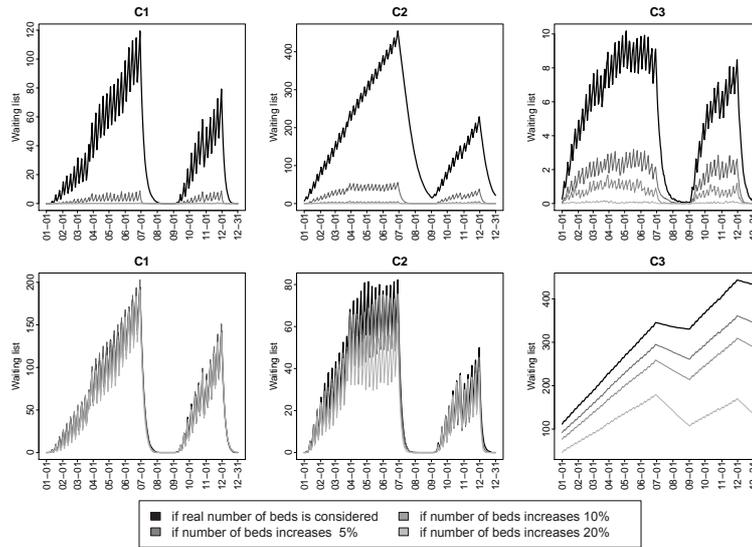


Figure 1: Cases 1 & 2 (first & second row, respectively). Simulated daily waiting list for hospital C1, C2 and C3 for a year period (from January to December).

different increases (5%, 10% and 20%). The study was repeated 500 times and a summary is shown in Figure 1 and Table 1. For Case 1, for all clusters, waiting lists disappear and the daily occupancy rate is drastically reduced when the number of beds rises. Taking into account the effect of the number of beds in the daily inpatient activity (Case 2), for C1, there are no significant differences in terms of waiting lists or occupancy rate when the number of beds increases. However, the increment in the number of beds reduces the waiting list and increases the percentage of days without waiting lists for C2, whereas an increase in the hospital beds leads to an increase in waiting list in C3.

As [4] in their study of the effect of hospital bed reduction in bed use 10 European countries found that admission rates appear to be sensitive to bed supply, with a positive elasticity of 1.44, the same has been found in the present study. It seems that there are no significant differences in terms of waiting lists

and occupancy rate when the number of beds of teaching hospitals increased. In that sense, [5] suggested that an increase in the supply of hospital beds tends to generate additional demand, either in the form of more patients admitted or patients treated for longer periods of time or some combination of the two. Supply-side policies may also be disappointing in their effects on waiting times for small rural hospitals. The same results have been found by [6] who show that a common experience is to take measures aimed at reducing waiting times by increasing activity, only to find that after a brief period demand has increased and waiting times have reverted to levels similar to those before the introduction of the measures, since demand responds positively to reductions in waiting times. On the other side, our study shows that higher capacity in terms of beds is associated with lower waiting times for medium size hospitals. The same results were found by [7] showing that econometric evidence suggests that higher capacity (more beds and physicians), is associated with lower waiting times.

## 4 Conclusions

The study shows the usefulness of simulation techniques to examine a hospital system (in particular, daily inpatient activity and length of stay), and to analyze the drop in waiting list numbers due to clustering under new-bed allocation.

**Acknowledgements.** The first and second authors would like to express their gratitude for support received from the Galician Regional Authority, Xunta de Galicia (INCITE08PXIB208113PR and INCITE08CSA0311918PR). The work of the third author was supported by Ministerio de Ciencia e Innovación (MTM2008-03010), and by Consellería de Innovación e Industria (PGIDIT07PXIB207031PR), and Consellería de Economía e Industria (10MDS207015PR), Xunta de Galicia.

## References

- [1] R Development Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2010. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- [2] F. Reyes, Adopción, difusión y utilización de la Alta Tecnología Médica en Galicia, in: Tomografía Computerizada y Resonancia Magnética, Universidade de A Coruña, Servizo de Publicacións, A Coruña, 2009.
- [3] M.P. Wand, M.C. Jones, Kernel Smoothing, Chapman & Hall, London, 1995.
- [4] M. Kroneman, J.J. Siegers, The effect of hospital bed reduction on the use of beds: a comparative study of 10 European countries, *Social Science & Medicine* 59 (2004) 1731-1740.
- [5] H. Zeraati, F. Zayeri, G. Babae, N. Khanafshar, F. Ramezanzadeh, Required hospital beds estimation: A simulation study, *Journal of Applied Sciences* 5 (2005) 1189-1191.

- [6] L. Siciliani, J. Hurst, Explaining Waiting Times Variations for Elective Surgery across OECD Countries, OECD Health Working Papers 7, 2003.
- [7] S. Martin, P.C. Smith, Rationing by waiting lists: an empirical investigation, *Journal of Public Economics* 71 (1999) 141-164.

# Modelling big game populations when hunting is age and sex selective

R. Cantó, B. Ricarte and A.M. Urbano \*

Institut de Matemàtica Multidisciplinar,  
Universitat Politècnica de València, E-46022, Valencia

November 30, 2011

## 1 Introduction

Quasi-positive linear control systems can be used to manage the big game population of a natural reserve where environmentally sustainable development must be guaranteed. These models allow us to estimate the annual measures (improvements or hunting) that must be implemented to take and maintain the studied population around the carrying capacity of the reserve, with an individual distribution that leads to maximum efficiency.

In this work we want to analyze the behaviour of a population that we have structured in different stage classes according to age and sex. For that on the basis of the quasi-positive systems, we construct a sex-specific age-at-harvest mathematical model that can analyzes the dynamic evolution of a population as a result of its intrinsic attributes and as well as external effects. Consequently, the model could be used in the developing of the cynegetic plan of a hunting natural reserve.

---

\*e-mail: {rcanto, bearibe, amurbano}@mat.upv.es. Supported by the Spanish DGI grant MTM2010-18228 and by the UPV under its research program PAID-06-10

## 2 Mathematical model

From the Lefkovich model [2, 3] and the quasi-positive lineal systems [1], we develop a new model that takes into account both age and sex of the individuals. Except for the youngest ones, which constitute the first stage class and where we assume the same percentage of males and females, we divide the life cycle of females into  $n_f$  stages and the life cycle of males into  $n_m$  stages. Then, the dynamic evolution of a population is given by:

$$N(t + 1) = AN(t) + U(t), \quad t \in \mathbb{Z}$$

where  $N(t) \in \mathbb{R}_+^n$  represents the individuals in each stage class at instant  $t$ , with  $n = 1 + n_f + n_m$  and  $U(t)$  shows the annual measures (improvements or hunting) that must be implemented to take the population around the carrying capacity of the studied area with a desired individual distribution.

Matrix  $A$  is given by:

$$A = \left[ \begin{array}{ccccc|cccc} p_0 & b_1 & b_2 & \cdots & b_{n_f} & 0 & 0 & \cdots & 0 \\ g_0/2 & p_1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & g_1 & p_2 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & p_{n_f} & 0 & 0 & \cdots & 0 \\ \hline g_0/2 & 0 & 0 & \cdots & 0 & p_{n_f+1} & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & g_{n_f+1} & p_{n_f+2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & p_{n-1} \end{array} \right]$$

with  $b_i \geq 0$ ,  $1 > p_i \geq 0$  and  $1 \geq g_i > 0$ .

**Proposition 1.** *Matrix  $A$  has a strictly dominant eigenvalue  $\lambda_1$  if the older female stage class is fertile, the newborn individuals or some females always survive and remain in the same stage and the maximum probability of surviving and remaining in the same state in females is greater than or equal to the probability in males.*

Without any external effect ( $U(t) = 0$ ), the asymptotic population growth only depends on  $\lambda_1$ , whereas the distribution of individuals among the different stage classes tends to stabilize independently of population growth proportionally to the associated eigenvector  $u_1$ .

If for economic reasons it is convenient to change this distribution and achieve another desired distribution  $w = [w_1 \ w_2 \ \dots \ w_n]^T$  we assume that  $U(t) = \text{diag}(\varepsilon_1(t), \varepsilon_2(t), \dots, \varepsilon_n(t))N(t) = M_\varepsilon(t)N(t)$ . Hence,

$$N(t + 1) = (A + M_\varepsilon(t))N(t) = A(t)N(t)$$

Matrices  $A(t)$  are obtained by the next algorithm where they are denoted by  $At$ ,  $D$  denotes the corresponding  $\lambda_1(t)$ ,  $t = 0, 1, \dots, k - 1$ , and  $Df$  the desired eigenvalue  $\lambda_{1k}$  the unitary eigenvector of which is  $w$ .

**Algorithm 1.** *Introduce  $A$ ,  $n_f$ ,  $n$ ,  $w$  and  $tol$ .*

1. *Let  $Df = A(1, 1) + A(1, 2 : n_f + 1)w(2 : n_f + 1)/w(1, 1)$*
2. *Obtain  $D$*
3. *Let  $At = A$*
4. *While  $|D - Df|/Df > tol$*

*For  $i = 2 : n_f + 1$ ,  $At(i, i) = D - At(i, i - 1)w(i - 1)/w(i)$ , end*  
 *$At(n_f + 2, n_f + 2) = D - At(n_f + 1, 1)w(1)/w(n_f + 1)$*   
*For  $j = n_f + 3 : n$ ,  $At(j, j) = D - At(j, j - 1)w(j - 1)/w(j)$ , end*  
*Obtain  $D$*

*end*

But we must also check that the desired population density  $P_f$  has not been rising. In any other case, we decrease the population growth applying a new negative control  $u(t) = \alpha \sum_{i=1}^n N_i(t) = \alpha [1 \ 1 \ \dots \ 1]N(t) = \alpha e^T N(t)$  distributed according to  $w$ . Therefore,

$$N(t + 1) = AN(t) + U(t) = AN(t) + (M_\varepsilon(t) - \alpha w e^T)N(t) \tag{1}$$

Finally, once  $P_f$  with the individual distribution  $w$  is achieved we apply a negative control  $u(t)$  to balance the population, which is estimated from the equation,

$$N(t + 1) = A_k N(t) - w u(t) \tag{2}$$

solving by less squares approximation so that  $u(t)$  is given by:

$$u(t) = (w^T w)^{-1} w (A_t N(t) - w P_f). \tag{3}$$

### 3 Example

The Spanish wild goat population in the Valencian natural reserve of “Muela de Cortes” has been selected to illustrate the model. This population is divided in nine stage classes: newborn individuals, four stage classes of females and four stage classes of males. Without any external influence, the population tends to grow in an indefinite manner, reaching a stable distribution that differs from the optimal distribution by 12%. To modify it we apply the quasi-positive model (1). Now, we get the desired population density and distribution in six years if a hunting strain of 8% and a tolerance of 5% are considered. From time  $t \geq 6$ , we consider  $A(t) = A(6)$  and apply equations (2) and (3) to keep the population in this conditions. All these results are summarized in Figure 1. After 5 years the population distribution differs with the optimal by less than 5%.

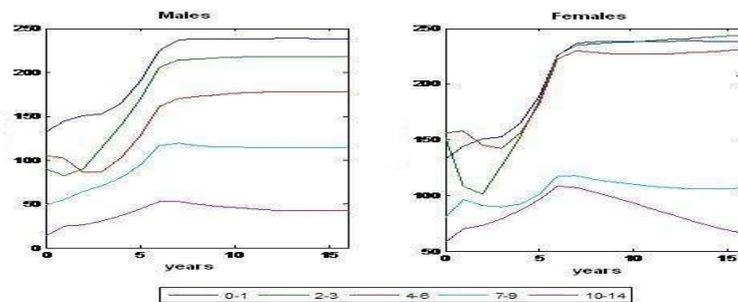


Figure 1: Population evolution estimated by the model (1).

### References

- [1] Cantó R., Ricarte B. and Urbano A. M. Definición y aplicación de los sistemas de control lineales quasi-positivos. *VI Congreso La investigación ante la sociedad del conocimiento: 75–78*, SPUPV, Alcoi, Alicante, 2009.
- [2] Crouse D. T., Crowder L. B. and Caswell H. A stage-based population model for loggerhead sea turtles and implications for conservation, *Ecology*, 68(5): 1412–1423, 1987.
- [3] Lefkovich L. P. The study of population growth in organisms grouped by stages, *Biometrics* 21: 1–18, 1965.

# Stability and convergence analysis of several approaches for a mathematical description of a small animal computed tomographic system. \*

M-J Rodríguez-Alvarez <sup>†</sup>, F. Sánchez <sup>†</sup>, A. Soriano <sup>†</sup>,  
A. Iborra <sup>‡</sup> and J.C. Cortés <sup>‡</sup>

(<sup>†</sup>) IBM, Centro mixto CSIC-Universitat Politècnica de València-CIEMAT  
Ciudad Politécnica de la Innovación, Camino de Vera s/n E-46022 Valencia (Spain)

(<sup>‡</sup>) Instituto de Matemática Multidisciplinar-Universitat Politècnica de Valencia  
Ciudad Politécnica de la Innovación, Camino de Vera s/n E-46022 Valencia (Spain)

November 30, 2011

## Abstract

The high computational cost for iterative image reconstruction algorithms in Computed Tomography (CT) can be partially avoided by using polar pixel techniques [1]. To deal with the 3D nature of the problem, exhaustive comparison among three different approaches was carried out in a 2D based image reconstruction procedure. From these considered 2D approaches, the one showing best performances was finally implemented in a 3D based image reconstruction model. We also present a comparison between 3D images reconstructed with our procedure and the well known FBP algorithm.

## 1 Introduction

Image reconstruction in computed tomography (CT) has been dominated by analytical methods like Filtered BackProjection (FBP) [2] because these methods produce images with a reasonable quality with low cost in terms of computing time. Recent advances in computer science enable the use of reconstruction techniques based on iterative methods [3, 4, 5, 6] as an alternative to FBP, without great penalty on the time needed for image reconstruction. Iterative methods allow a detailed mathematical description of the tomographic systems.

---

\*This work was supported by the Spanish Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica (I+D+I) under grants nos. MTM2009-08587, FIS2010-21216-C02-01, and by the Valencian Government under grants nos. PROMETEO 2008/114 and APOSTD/2010/012.

<sup>†</sup>mjrodri@mat.upv.es

As a consequence of the improvement of the image quality, less patient dose is needed during CT examination. Promising results obtained at the Mayo Clinic in Arizona reported a 63 % reduction in dose received by the patient when iterative reconstruction methods were considered instead of FBP [7].

The mathematical description of the tomographic system in iterative reconstruction process is done through the so-called weight matrix. Three different ways to obtain the weight matrix are presented and compared for a 2D-CT scanner. The best of which is extended from 2D to 3D. Performance evaluation of the different approaches proposed in this work is conducted with real data from a small animal micro CT. The FBP algorithm has been implemented to compare 3D iterative images against the usual algorithm for CT image reconstruction.

## 2 Iterative reconstruction

The iterative image reconstruction can be regarded as a standard statistical estimation problem. Maximum Likelihood Expectation Maximization (MLEM) algorithm [3] has attracted considerable interest in the area of image reconstruction, as it produces images with better quality than FBP [2]. In order to speed up the reconstruction process the Ordered Subsets Expectation Maximization (OSEM) algorithm was proposed [4]. It is a modified MLEM that groups projections into subsets of angles. Within each iteration the image is updated as many times as the number of subsets, proportionally accelerating the convergence.

In order to take full advantage from the symmetry of the scanner geometry, we propose to use polar discretization to arrange the pixels in the field of view (FOV). So that the weight matrix elements are equal for all the projections. Since more than one hundred projections are usually required to obtain a CT image, the polar description cuts down the time and storage requirements of the weight matrix. Recent studies [1] reveal that polar pixel-based weight matrices allow an efficient implementation of 3D-MLEM and 3D-OSEM, with an image quality equal to the equivalent cartesian ones [8].

The elements  $w_{ij}$  of the weight matrix represent the contribution of each voxel  $j$  to the attenuation of beam  $i$ . The way the  $w_{ij}$  for each pixel (voxel in 3D) are defined, strongly influences the quality of the reconstructed image obtained after the iterative process. Three different ways to calculate the weight matrix have been considered in this work.

**Nearest neighbor** Each pixel only contributes to its closest beam.

**Joseph's Method** [9], Assumes that each pixel contributes to its two surrounding beams. The weighting factor decreases with the relative distance of the pixel respect to each beam.

**Area intersection** The weights  $w_{ij}$  are defined as the area of the pixel  $j$  which intersects a beam  $i$ .

The performance of each approach has been assessed with a 50 mm height and 55 mm in diameter cylindrical phantom of plexiglass. Five holes of 8 mm in

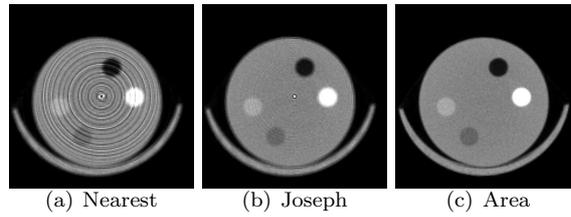


Figure 1: 2D reconstructed images.

diameter were axially drilled at 16 mm off the axis. Four of the holes were filled with 8 mm in diameter inserts of PolyEthylene, plexiglass, Delrin and Teflon which model adipose tissue, water, organs tissue and soft bone respectively. The remaining hole was left empty to model an air region inside the body. In order to assess the performance of each approach, the average and standard deviation values in each material during the iterative process have been considered.

### 3 Results

Although mean values nearly converge to the same value in all three approaches considered, strong differences are appreciated in the noise obtained with the different approaches considered. Transversal slices of the reconstructed object are shown in Fig. 1. It permits to identify the ring artifacts which lead to differences in the noise measured in each approach. Strong artifacts are obtained when the nearest neighbor approach is considered. Artifacts diminish considerably if the two surrounding LORs are considered, although little artifacts are still appreciated. No evidences of ringing artifacts are seen in Fig. 1 when the weights matrix is calculated as the area intersection between LOR and voxel.

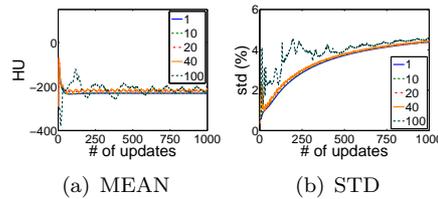


Figure 2: Results obtained for 3D sector matrix as a function of update number for 1, 10, 20, 40 and 100 subsets.

The performance evaluation of the iterative process when 1, 10, 20, 40 and 100 subsets are considered is shown in Fig. 2 for the 3D case and the "Area intersection" approach. After about 50 updates, average values measured in each RoI converge to the same value when less than 20 subsets are considered. As the number of subsets increases an oscillation of the average value, whose period

coincides with the number of subsets, becomes more and more appreciable.

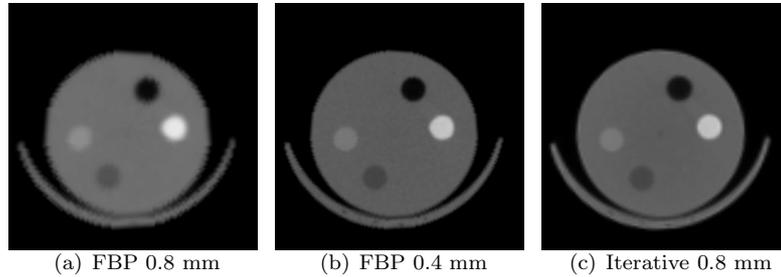


Figure 3: Comparison of transversal 3D images reconstructed with OSEM and FBP with several voxel sizes.

The comparison of reconstructed images obtained with iterative and FBP reconstructions is shown in Fig. 3. The iterative reconstruction procedure consisted of 5 iterations of OSEM algorithm with 20 subsets. When comparing reconstructed images with the same voxel size of 0.8 mm, iterative method OSEM exhibits a superior quality. If the voxel size in FBP reconstruction is reduced to 0.4 mm the image quality improves, being comparable to that in Fig. 3(c). Results in Fig. 3 confirm that useful images for medical applications in small animals using voxel sizes as big as 0.8 mm are possible when iterative methods are considered. FBP reconstruction requires voxel sizes of about 0.4 mm to provide useful images therefore increasing image sizes.

## 4 Conclusions

Research presented in this work shows up the relevance of an accurate mathematical description of the scanner in computed tomography image reconstruction. Real data from a small animal micro CT has been considered in this work. In order to take advantage from the symmetry of the scanner geometry we have used polar discretization of the pixels within the FOV. The proposed polar discretization reduces the storage requirements of the system matrix on the computer by a factor that equals the number of projections used in the tomographic system, usually 200. Although all weight matrices considered nearly converge to the the same value in average, noise in the image exhibited strong differences. The comparison of transversal slices reflects that noise differences are due to strong “ring” artifacts present in less accurate approaches. From the three approaches evaluated, that considering the weights as the area intersection between LOR and pixel had the best performance which has been extended to 3D. Results obtained with iterative reconstruction were finally compared against FBP reconstructed images. Superior image quality is appreciated in iterative reconstruction when the same voxel size is used.

## References

- [1] M. J. Rodríguez-Alvarez, F. Sánchez, A. Soriano, A. Iborra, C. Mora, Exploiting symmetries for weight matrix design in CT imaging, *Mathematical and Computer Modelling* 54 (7-8) (2011) 1655–1664.
- [2] L. Feldkamp, L. Davis, J. Kress, Practical cone-beam algorithm, *Journal of the Optical Society of America* 1 (1984) 612–619.
- [3] L. Sheep, Y. Vardi, Maximum likelihood reconstruction for emission tomography, *IEEE Transactions on Medical Imaging* 1 (1982) 113–122.
- [4] H. Hudson, R. Larkin, Accelerated image reconstruction using ordered subsets of projection data, *IEEE Transactions on Medical Imaging* 13 (1994) 601–609.
- [5] E. Y. Sidky, X. Pan, Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization., *Physics in Medicine and Biology* 53 (2008) 4777–4807.
- [6] J. Bian, J. H. Siewerdsen, X. Han, E. Y. Sidky, J. L. Prince, C. A. Pelizzari, X. Pal, Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam ct, *Physics in Medicine and Biology* 55 (2010) 6575–6599.
- [7] A. K. Hara, R. G. Paden, A. C. Silva, J. L. Kujak, H. J. Lawder, W. Pavlicek, Iterative reconstruction technique for reducing body radiation dose at CT: Feasibility study., *American Journal of Roentgenology* 193 (2009) 764–771.
- [8] C. Mora, M. J. Rodríguez-Alvarez, J. V. Romero, New pixellation scheme for CT algebraic reconstruction to exploit matrix symmetries, *Computers & Mathematics with Applications* 56 (3) (2008) 715–726.
- [9] P. M. Joseph, An improved algorithm for reprojecting rays through pixel images, *IEEE Transactions on Medical Imaging* MI-1 (1982) 192–196.

# Study of the turbulence induced by cavitation phenomenon in diesel injector nozzles by Large Eddy Simulation (LES)

F.J. Salvador<sup>†\*</sup>, J. Martínez-López<sup>†</sup>, J.-V. Romero<sup>‡</sup>  
and M.-D. Roselló<sup>‡</sup>

(†) CMT- Motores Térmicos, Universitat Politècnica de València,  
Camino de Vera s/n, Edificio 6D, 46022, Valencia, Spain,

(‡) Instituto Universitario de Matemática Multidisciplinar,  
Universitat Politècnica de València,  
Camino de Vera s/n, Edificio 8G, 2º, 46022, Valencia, Spain.

November 30, 2011

## 1 Introduction

Most of the improvements achieved to reduce the emissions of diesel engines and improve their efficiency are due thanks to the advances in fuel injection systems. One of the most widely used strategies in recent years has been the increase of injection pressure, which has forced that cavitation and turbulence play a fundamental role on the internal flow and spray development [1].

The experimental study of the turbulence developed within diesel injector nozzles and its interaction with cavitation phenomenon present huge difficulties due to extremely small size of the holes, the existence of a multiphase flow and the high velocities found in the discharge orifices, so the use of Large Eddy Simulation (LES) has become a great alternative to advance in the knowledge of the internal flow [2, 3].

---

\*e-mail: fsalvado@mot.upv.es

## 2 Description of the geometry simulated

The nozzle studied is a cylindrical six-hole microsac nozzle with a diameter of  $170\ \mu\text{m}$ , 1 mm of length and a curvature inlet radius of  $13\ \mu\text{m}$ , simulating only the domain belonged to one orifice of the whole nozzle (see Figure 1) in order to reduce the high computational cost typical of LES techniques, which capture the large scale motions of the flow and models the small scale motions that occur on length scales smaller than the mesh spacing. The geometry has been discretised in a hybrid mesh of 1.836.964 cells, with cell size ranging in the orifice from  $0.6\ \mu\text{m}$  in the orifice wall to  $3\ \mu\text{m}$  in the orifice core ( $D/56.7$ ) and  $8\ \mu\text{m}$  in the rest of the nozzle.

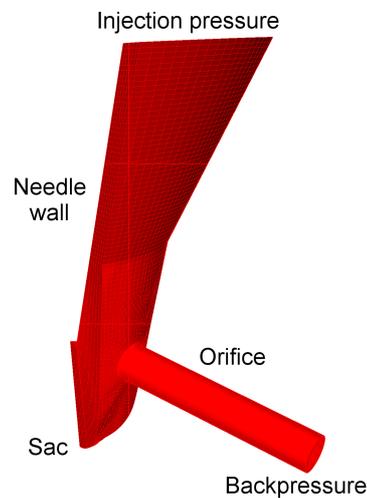


Figure 1: Multi-hole nozzle simulated.

## 3 Validation of the code

For validation purposes and subsequent study of the cavitation and turbulence interaction, four simulations have been calculated with OpenFOAM fixing the injection pressure to 160 MPa and varying the backpressure to 3, 5, 7 and 9 MPa. Each simulation has been run in parallel using 30 processes, needing 90 days to complete  $100\ \mu\text{s}$ .

Figure 2 shows the comparison of numerical results for the mass flow, momentum flux and effective velocity against experimental data obtained in CMT-Motores Térmicos laboratories. As can be observed, for all the parameters and pressure conditions compared, CFD calculations slightly overestimate the experimental results, following always the tendencies of the experimental data.

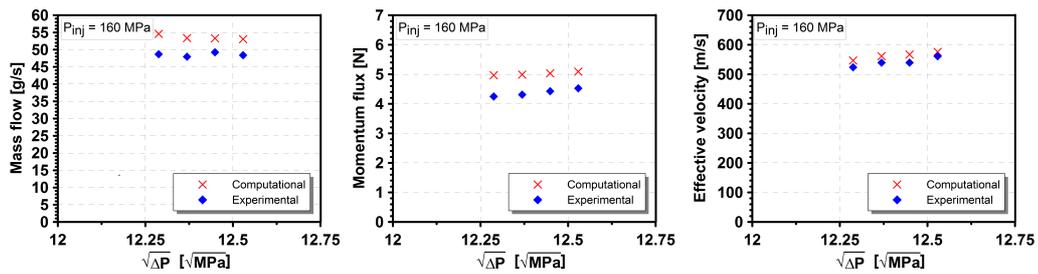


Figure 2: Comparison of experimental and numerical results.

Attending to the mass flow injected by the nozzle, represented in the left plot, it can be seen that the mass flow is collapsed in both cases, keeping constant for all the pressure conditions. This indicates that the flow is at cavitating conditions [4]. The deviation between experimental and computational values is around 10%. However, due to the measurements uncertainties in the determination of the mass flow, momentum flux, and the internal geometry, and the complexity of modeling cavitating flows it can be considered that the code is able to predict the behavior of the flow with enough degree of confidence.

Opposite to the mass flow behavior, the momentum flux represented in the plot placed in the middle of Figure 2 is not choked, so an increment of effective velocity is expected since it is calculated dividing the momentum flux values by the mass flow. Effectively, as can be seen on the right of Figure 2, the effective velocity shows a linear increase with the root of pressure drop, reaching values of almost 600 m/s, with a deviation between experimental and computational results for the velocity always lower than 5%. This velocity rise is justified by the fact that the presence of vapour near the orifice wall due to cavitation phenomenon reduces the friction losses (slip condition) and the effective diameter for the liquid phase [1].

## 4 Vapour phase distribution and turbulence development

Attending to the mass flow results seen in Figure 2, it is expected to see vapour bubbles into the discharge orifices for any pressure condition simulated, since the mass flow remains collapsed due to cavitation phenomenon.

In order to assess it, Figure 3 shows the vapour field average (where blue regions correspond to pure liquid and red region correspond to pure vapour) at seven cross sections of the orifice and its standard deviation for the simulation  $P_{inj} = 160 \text{ MPa} - P_{back} = 5 \text{ MPa}$ , averaging a period equal to 15 times the time needed by a fluid particle to cover the whole orifice length.

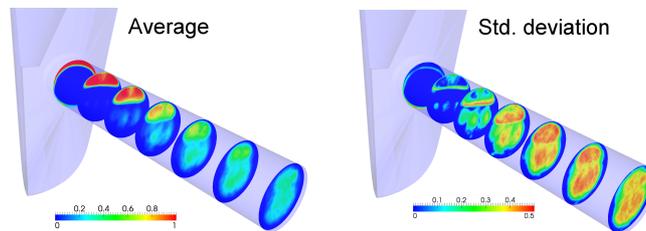


Figure 3: Vapour phase average and standard deviation.

As expected, the average field shows as the vapour bubbles are distributed along the upper part of the orifice, being the origin of the vapour phase the inlet corner region of the channel. The vapour bubbles flow along the upper surface of the orifice wall reaching the exit although it is possible to see vapour phase in the center and lower part of the orifice near the outlet section. Furthermore, the standard deviation values indicate that cavitation is always present near the orifice inlet, whereas cavitation appearance as the bubbles reach the exit varies time step to time step, showing that cavitation is a transient phenomenon.

The influence of the turbulence on the cavitation appearance can be seen in Figure 4, where vapour phase has been depicted for several cross sections of the orifice. Vapour phase is clearly divided in two cores, which can be also observed representing the velocity vectors in one of the cross sections chosen of the orifice. The direction of the vectors shows a pair of twin vortices, one of them in the left part, and another one in the right part of the section.

This behavior, which has been observed by other authors [5], justifies the separation of the vapour phase in two regions.

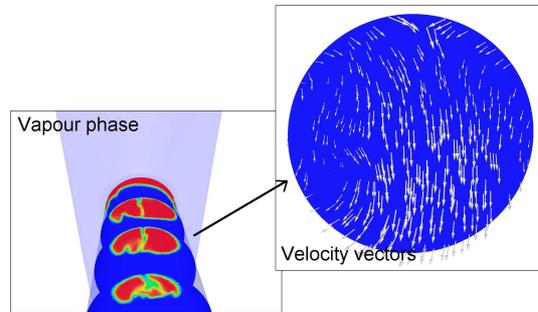


Figure 4: Vapour phase and velocity vectors for the simulation  $P_{inj} = 160$  MPa –  $P_{back} = 5$  MPa.

Up to now it has been proved that the turbulence developed inside a diesel injector nozzle has a strong influence on the vapour phase appearance. Nevertheless, there is a certain interaction or interdependence between both phenomena, since cavitation also enhances turbulence [6]. Indeed, if the vapour phase within the nozzle is compared with the vorticity (see Figure 5) it is possible to state that the highest values of turbulence is due to the presence of vapour bubbles, reaching the maximum values in the interphase liquid-vapour region.

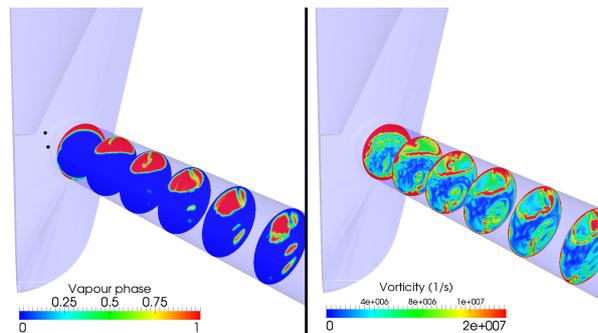


Figure 5: Vapour phase distribution and vorticity ( $P_{inj} = 160$  MPa –  $P_{back} = 5$  MPa).

## 5 Conclusions

- A code to model cavitation phenomenon by Large Eddy Simulation has been validated in a diesel injector nozzle at real operating conditions, showing an acceptable level of confidence.
- Cavitation appearance is strongly influenced by the turbulence and vortices development, justifying the separation of the vapour phase in two cores in the zone near the orifice inlet.
- Cavitation enhances turbulence, since the highest values of vorticity in the nozzle are found at the interphase liquid-vapour.

## Acknowledgments

This work was partly sponsored by “*Vicerrectorado de Investigación, Desarrollo e Innovación*” of the “*Universitat Politècnica de Valencia*” in the frame of the project “*Estudio numérico de la cavitación en toberas de inyección Diesel mediante Grid Computing (Cavigrid)*”, reference N° 2597 and by “*Ministerio de Ciencia e Innovación*” in the frame of the project “*Estudio teórico-experimental sobre la influencia del tipo de combustible en los procesos de atomización y evaporación del chorro Diesel (PROFUEL)*”, reference TRA2011–26293. This support is gratefully acknowledged by the authors. The authors would also like to thank the computer resources, technical expertise and assistance provided by the Universidad de Valencia in the use of the supercomputer “Tirant”.

## Nomenclature

$D$ : diameter

$P_{\text{back}}$ : discharge back pressure

$P_{\text{inj}}$ : injection pressure

$\Delta P$ : pressure drop,  $\Delta P = P_{\text{inj}} - P_{\text{back}}$

## References

- [1] Payri R., García J. M., Salvador F. J. and Gimeno J., Using spray momentum flux measurements to understand the influence of diesel nozzle geometry on spray characteristics *Fuel*, 84:551–561, 2005.
- [2] Piomelly U., Large-eddy simulation: achievements and challenges *Progress in Aerospace Sciences*, 35(4):335–362, 1999.
- [3] Payri R., Tormos B., Gimeno J. and Bracho G., The potential of Large Eddy Simulation (LES) code for the modeling of flow in diesel injectors *Mathematical and Computer Modeling*, 52:1151–1160, 2011.
- [4] Salvador F. J., Hoyas S., Novella R. and Martínez-López J., Numerical simulation and extended validation of two-phase compressible flow in Diesel injector nozzle *Proceedings of the Institution of Mechanical Engineers, Part D, Journal of Automobile Engineering*, 225:545–563, 2011.
- [5] Masuda R., Fuyuto T. and Nagaoka M., Validation of a diesel fuel spray and mixture formation from nozzle internal flow calculation *SAE paper 2005-01-2098*.
- [6] Desantes J. M., Payri R., Salvador F. J. and De la Morena J., Influence of cavitation phenomenon on primary break-up and spray behavior at stationary conditions *Fuel*, 89:3033–3041, 2010.

# An Approach to the Application of Shift-and-add Algorithms on Engineering and Industrial Processes

J.L. Sanchez-Romero<sup>†</sup>\*, A. Jimeno-Morenilla<sup>†</sup>, R. Molina-Carmona<sup>‡</sup>,  
and J. Perez-Martinez<sup>†</sup>

(†) Department of Computer Technology and Data Processing,  
University of Alicante, 03690 Alicante, Spain.

(‡) Department of Computing Sciences and Artificial Intelligence,  
University of Alicante, 03690 Alicante, Spain.

November 30, 2011

## 1 Introduction

CORDIC (COordinate Rotation DIgital Computer) is an iterative algorithm to approximate mathematical functions [1]. It was originally developed for working with radix-2 operands. The method was extended to support hyperbolic and linear coordinates [2].

In recent years, a renewed interest in decimal computing has arisen, since it is essential for many applications and scopes [3]. In [4] and [5] the use of CORDIC for BCD operands was proposed. This method cannot compare to binary CORDIC with regard to latency, since the binary method requires a smaller number of iterations so as to obtain the same precision. Therefore, the advantages of using the algorithm with decimal operands would be reduced to omit conversion between decimal and binary representation and, consequently, to avoid loss of precision.

---

\*e-mail:sanchez@dtic.ua.es

In [6], we proposed a new CORDIC which works with decimal operands. Besides avoiding the tasks of conversion from decimal to radix-2 representation and vice versa, which implies neither conversion error nor time consumption implicit in these stages, the specific features of this recent decimal CORDIC allow a significant reduction in the number of required iterations. From this point on, the new decimal CORDIC will be referred to as ND-CORDIC, whereas binary CORDIC and the older decimal CORDIC [4] will be referred to as B-CORDIC and D-CORDIC, respectively.

In this paper, we will analyze the requirements of some engineering and industrial problems in terms of type of operands and functions to approximate. We will propose the application of different CORDIC versions to these problems. We will make a comparison between the different methods applied in terms of the obtained precision and the number of iterations required.

## **2 Application of CORDIC on Industrial Processes.**

### **2.1 Specific characteristics of data in industrial and engineering environments**

A trending issue in CAD/CAM is the utilization of decimal operands, since in the earlier stages of the product design engineers work with decimal magnitudes, which are propagated throughout the whole manufacturing process. Therefore, decimal formats should be considered so as to represent data and operate with them.

### **2.2 Computation of specific operations in industrial processes**

One of the most important tasks in the manufacturing process is machining. It can be described as the definition of an object to be manufactured by means of predefined tools. The more usual operation in machining is the removal of material by using a cutting tool. In the machining process, the main task is the tool path computation: obtaining a trajectory of tool centres that defines the required object to be machined with a given precision. Therefore, for every object point to be machined, the position of the tool centre must

be determined. This task involves a high computational cost.

An example of a hard-to-compute distance is the one required when using a toroidal tool for machining [7]. A toroidal cutting tool is characterized by a major radius  $R$ , a minor radius  $r$ , and the coordinates of the torus center  $(T_x, T_y)$ , as shown in Fig. 1. The distance from a point  $(p_x, p_y, p_z)$  to a toroidal tool centre is given by the following expression:

$$d(p_x, p_y, p_z) = (T_y - p_y) - \sqrt{(R + \sqrt{r^2 - (p_x - T_x)^2})^2 - p_z^2} \quad (1)$$

In equation 1, two square roots appear:

$$sqr1 = \sqrt{r^2 - (p_x - T_x)^2} \quad (2)$$

$$sqr2 = \sqrt{(R + sqr1)^2 - p_z^2} \quad (3)$$

The argument of each one of these square roots is the difference of two squares. The computation of such operations has a high computational cost. However, it can be efficiently performed by the CORDIC algorithm. In fact, CORDIC working in vectoring mode with hyperbolic coordinates directly provides as a result the square root of the difference of two squares [8]. Therefore, an algorithm including two instances of CORDIC can be developed so as to calculate the distance shown in Eq. (1):

```
Algorithm distance (px, py, pz, Tx, Ty, R, r: real): real;
Variables sqrt_1, sqrt_2, y1, x2, d: real;
y1 = px - Tx;
sqrt_1 = CORDIC_Hyp_Vec(x = r, y = y1);
x2 = R + sqrt_1;
sqrt_2 = CORDIC_Hyp_Vec(x = x2, y = pz);
d = Ty - py - sqrt_2;
return d;
End Algorithm
```

For B-CORDIC, some of the iterations must be performed twice so as to guarantee convergence. For D-CORDIC, repeating some iterations is also required so as to achieve convergence. In this case, the repetition are much more numerous than those for B-CORDIC. The experiments shown in [5] used 10 iterations for the initial stage, that is, for  $j = 1$ , and 9 iterations for the following ones. In case of ND-CORDIC, it is not necessary to repeat any

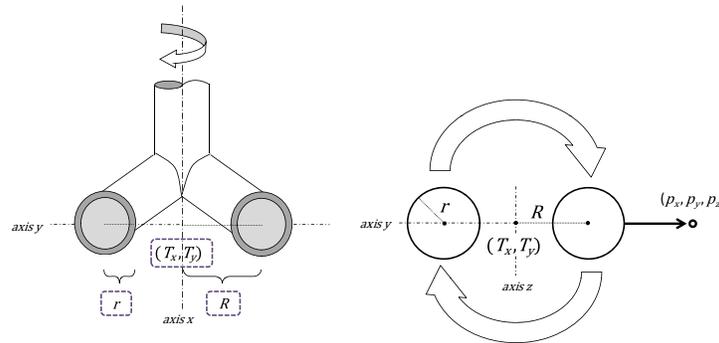


Figure 1: Structure of a toroidal tool and its main parameters.

stage, due to the fact that the elementary angle chosen in each stage of the vectoring mode is not fixed, but it is specifically selected according to the value of  $x_j$  and  $y_j$ .

Different tests were carried out so as to make a complete comparison with regard to latency and precision between B-CORDIC, D-CORDIC, and ND-CORDIC. The architecture for ND-CORDIC, as proposed in [6], was implemented on the hardware description language VHDL using the Xilinx ISE Design Suite 10.1 tool. The Virtex4 XC4VLX60 FPGA device was randomly chosen for simulation and synthesis. The architectures for D-CORDIC proposed in [4] and [5] and for B-CORDIC were also implemented. In all cases, a complete stage of the method was implemented.

Original data were selected in interval  $[0,1)$  and represented in BCD with 6 fractional digits. A homogeneous length of 28 bits was used for every number format, so 24 bits were used for the fractional digits of BCD and BCD XS3 numbers. Conversion from BCD to BCD XS3 and vice versa can be easily performed by means of a few logic gates. The use of BCD XS3 is proposed since addition, subtraction, and other operations such as detection of zero and nine's complement are simpler than for BCD [4]. The initial conversion from BCD to BCD XS3 and the final conversion the other way were also included for D-CORDIC and ND-CORDIC. For B-CORDIC, an initial conversion from BCD to binary and a final conversion the other way were also implemented. The results obtained with every algorithm were compared with those obtained by means of the direct application of Eq. (1), calculated in Matlab©. The mean relative error according to the execution time is depicted in Fig. 2. For ND-CORDIC, experiments were performed

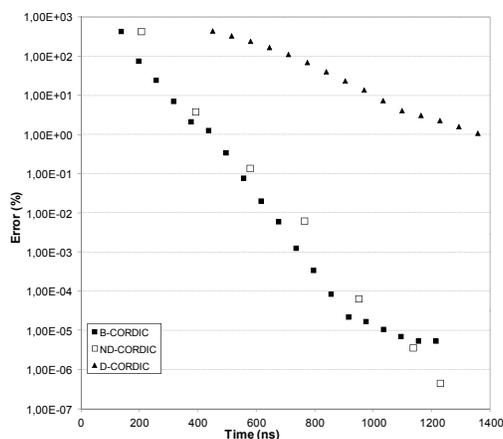


Figure 2: Mean relative error according to delay for each CORDIC architecture when calculating the toroidal tool distance; logarithmic scale.

with only six iterations of the first instance of CORDIC (calculation of  $\sqrt{rt_1}$ ). This optimization is due to the fact that there is not a significant loss of precision and, moreover, the overall latency is reduced.

As observed, the error for ND-CORDIC is lower than that for D-CORDIC at any time interval. On the other hand, accuracy achieved by ND-CORDIC is similar to the one obtained with B-CORDIC and, indeed, ND-CORDIC provides a precision of  $10^{-6}$  with a latency similar to the one of B-CORDIC. Besides, when a high number of iterations is performed, ND-CORDIC achieves an error lower than this limit, while B-CORDIC is not able to.

### 3 Conclusion

In this paper, a proposal for the application of CORDIC on engineering and industrial processes has been shown. A manufacturing task with a high computational cost has been analyzed and, as a result, the application of CORDIC has been proposed for improving performance. Moreover, specific features of operands involved in these tasks suggest the utilization of CORDIC versions specially developed for working with decimal operands. Experimentation has been performed for comparing the original, binary CORDIC (B-CORDIC) with two versions of CORDIC for decimal operands (D-CORDIC and ND-CORDIC). The results confirm that ND-

CORDIC requires fewer iterations than both B-CORDIC and D-CORDIC so as to obtain a required precision. In addition, ND-CORDIC offers similar global performance than B-CORDIC and, in some cases, accuracy is even better than the one obtained with B-CORDIC.

## Acknowledgments

This research was supported by the Conselleria de Educacion of the Valencia Region Government under grant number GV/2011/043.

## References

- [1] Volder J. The CORDIC Trigonometric Computing Technique, *IRE Trans. Electron. Comput.*, **EC-8**(3):330–334, 1959.
- [2] Walther J.S. A unified algorithm for elementary functions, *Proc. AFIPS Spring Joint Computer Conf.*, 379–385, 1971.
- [3] Cowlishaw M.F. Decimal Floating Point: Algorithm for Computers, in: *Proc. 16th IEEE Symp. Computer Arithmetic*, 104–111, 2003.
- [4] Schmid H. *Decimal Computation* (John Wiley & Sons, New York, 1974).
- [5] Schmid H., and Bogacki A. Use decimal CORDIC for generation of many transcendental functions, *EDN*, **Feb.**, 64–73, 1973.
- [6] Sanchez J.L., Mora H., Mora J., and Jimeno A. Function approximation on decimal operands, *Digital Signal Processing*, **11**(2):354–366, 2011.
- [7] Jimeno A. *Modelado topologico del proceso de fabricacion (Topological modelling of the manufacturing process)*, Ph.D. Thesis, University of Alicante, Alicante, Spain, 2003.
- [8] Muller J.M., *Elementary Functions. Algorithms and Implementation* (Birkhäuser, 2006).

# Lattice Boltzmann method in bioreactor design and simulation

V. Štumbauer\* ; K. Petera†, and D. Štys\*

(\*) Institute of Physical Biology,

University of South Bohemia,

373 33 Nové Hradý, Czech Republic

(†) Czech Technical University in Prague,

Technická 4, 166 07 Prague 6, Czech Republic.

November 30, 2011

## 1 Introduction

Microalgae has gained a lot of attention due to its wide area of applicability. It has been shown that it has a high potential as a source of renewable energy – i.e. biofuels and also in other environment related applications like CO<sub>2</sub> sequestration, bioremediation, waste water treatment and others.

Mainly due to the complex multi-scale processes governing the microalgae growth in multi-phase, multi-component flow, the bioreactor design still is rather an empirical process. In our previous work [3, 4] we have introduced an alternative approach to bioreactor simulation based on random walk with spatially dependent dispersion coefficient devised from a CFD simulation and with the work presented we would like to contribute to more deterministic bioreactor design methods with a method based on parallel LBM fluid flow simulation and parallel stochastic implementation of PSF. Within the scope of our work we focus on the so-called Couette-Taylor photobioreactor(CTBR) [5].

---

\*e-mail: stumbav@gmail.com

## 2 Materials & Methods

The Lattice Boltzmann method has gained a lot of attention in diverse areas because of its ability to simulate complex flows, straight-forward implementation and parallelism potential. LBM is still lacking a lot in boundary treatment and there exist various approaches to the treatment of the curved velocity boundary required to simulate fluid flow in CTBR. In this work we have employed and validated the interpolation scheme of Mei [2]. The so-called Photosynthetic Factory model (PSF) [1] has been employed for microalgae growth simulation in the 3D flow field of homogeneously irradiated CTBR. Irradiance is the sole input parameter of the model and the model exhibits the important property of 'integrating' fast irradiance changes. Thus in dense well-mixed cultures all microalgae cells may be exposed to irradiance intermittently and grow as if exposed to an averaged irradiance value. Both LBM and PSF solver have been implemented on CPU as well as the parallel architecture of CUDA in order to compare the performance advantage of the method.

## 3 Results

### 3.1 LBM solver

As a means of validation for the LBM solver in case of CTBR a cylindrical Couette flow was simulated at various  $\frac{r_1}{r_2-r_1}$  and compared with the analytical velocity profile – see figure 1. Computations have been done on a grid of  $64 \times 64$  nodes (cross section). Speedup of LBM solver between CPU implementation and parallel GPU implementation has been measured at various grid sizes. Smaller grid sizes ( $16 \times 16 \times 16$ ) show much lower speedup than larger ones – speedup for larger grids was approximately constant ( $\doteq 20$ ).

### 3.2 PSF solver

To validate the PSF solver a comparison of simulated results with the analytical solution for  $x_a$  of the PSF model at constant irradiance has been performed. Figure 2 depicts comparison of averaged simulated results for 8192 microalgae cells and analytical solution at various levels of constant illumination normalized by the optimal illumination  $I_{opt}$ . Comparison is based on results from [1] ( $\alpha = 5\beta$ ,  $\gamma = 5\delta$ ).

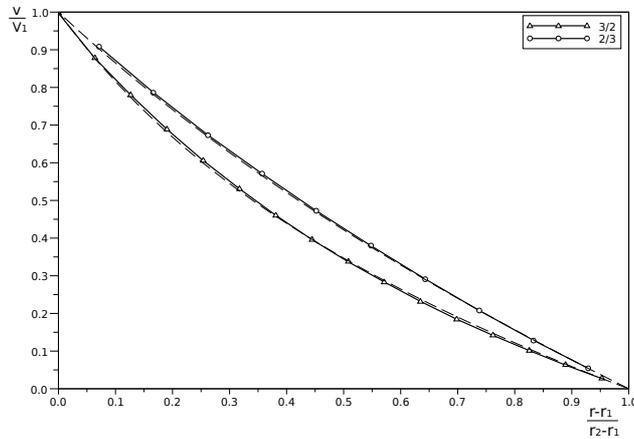


Figure 1: Comparison of simulated cylindrical Couette flow velocity profile with analytical solution at different  $\frac{r_1}{r_2-r_1}$ .

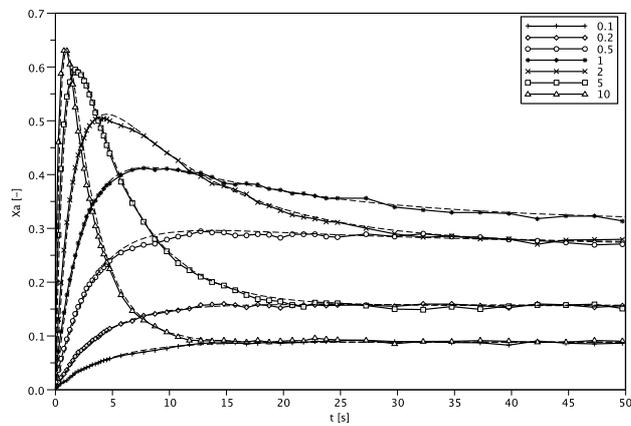


Figure 2: Comparison of simulated PSF results and analytical solution.

## 4 Conclusions

Within the presented work an alternative approach to bioreactor design and simulation based on parallel implementation of Lattice Boltzmann method for fluid flow and parallel stochastic implementation of Photosynthetic Factory for microalgae growth has been presented. Simulation approach has been validated on the so-called Couette-Taylor photobioreactor. Fluid flow simulation results have been matched against analytical velocity profiles – the cylindrical Couette flow. Parallel stochastic approach to the model of Photosynthetic Factory has been successfully validated by comparison of simulated results and analytical results at constant irradiance.

## Acknowledgement

This work was supported by the project “Jihočeské výzkumné centrum akvakultury a biodiverzity hydrocenóz”(CENAKVA), OP VaVpI.

## References

- [1] Eilers, P.H.C., Peeters, J.C.H. Dynamic behaviour of a model for photosynthesis and photoinhibition. *Ecological Modelling*, 69:113–133, 1993.
- [2] Mei R., Shyy W., Luo L.S. Lattice Boltzmann method for 3D flows with curved boundary. *Journal of Computational Physics*, 161:680–699, 2000.
- [3] Papáček Š., Matonoha C., Štumbauer V., Štys D. Modelling and simulation of photosynthetic microorganism growth: random walk vs. finite difference method. *Mathematics and Computers in Simulation*, in press
- [4] Papáček Š., Štumbauer V., Štys D., Petera K., Matonoha C. Growth impact of hydrodynamic dispersion in a Couette-Taylor bioreactor. *Mathematical and Computer Modelling*, 7-8 (54):1791–1795.
- [5] Papáček Š., Štys D., Dolínek P., Petera K. Multicompartment/CFD modelling of transport and reaction processes in Couette-Taylor photobioreactor. *Applied and Computational Mechanics*, 1:577–586, 2007.

# Improvements of true anomaly iteration method for preliminary orbit determination\*

Francisco Chicharro, Alicia Cordero and Juan R. Torregrosa<sup>†</sup>

Instituto de Matemática Multidisciplinar, Universitat Politècnica de València,  
Camino de Vera, s/n, 46022, Valencia, Spain

November 30, 2011

## 1 Introduction

From position and velocity coordinates for several given instants, it is possible to determine orbital elements for the preliminary orbit of an artificial satellite, taking only into account mutual gravitational attraction forces between the Earth and the satellite. Different methods developed for this purpose need to find a solution of a nonlinear function. In classical methods it is usual to employ fixed point or secant method. We focus our attention in the method of iteration of the true anomaly, in which the secant classical method is replaced by more efficient methods.

### The two-body problem

Let  $M$  and  $m$  be the masses of the primary (the Earth) and secondary (the satellite) bodies, respectively, and  $G$  the universal gravitational constant. Moreover, let  $\mu$  be the gravitational parameter, defined by  $\mu = G(M + m) \approx GM$ . By applying second Newton's law for each mass, and adding both

---

\*This research was supported by Ministerio de Ciencia y Tecnología MTM2010-18539 and by Vicerrectorado de Investigación, Universitat Politècnica de València PAID-06-2010-2285.

<sup>†</sup>e-mail: frachilo@teleco.upv.es (F. Chicharro), acordero@mat.upv.es (A. Cordero), jr-torre@mat.upv.es (J. R. Torregrosa).

expressions, we obtain  $\ddot{\vec{r}} + \frac{\mu}{r^3}\vec{r} = 0$ , where  $|\vec{r}| = r = \frac{p}{1+e\cos\nu}$ . This expression can be interpreted as a conic, where  $p$  is the semiparameter of the conic,  $e$  is its eccentricity and the polar angle  $\nu$  is called true anomaly. The scalar velocity of the satellite, given by Vis-Viva Equation (see [1]), can be obtained by  $v = \sqrt{\mu\left(\frac{2}{r} - \frac{1}{a}\right)}$ , where  $a$  is the semimajor axis of the conic. Many observations of the satellites motion determines that the orbit is bounded. For elliptic motion,  $0 < e < 1$ , and  $0 < a < \infty$ . The orbital elements set  $(a, e, i, \omega, \Omega, \tau)$  determinate the position and the velocity of a satellite in space, or vice versa. Some of them are described in Figure 1.

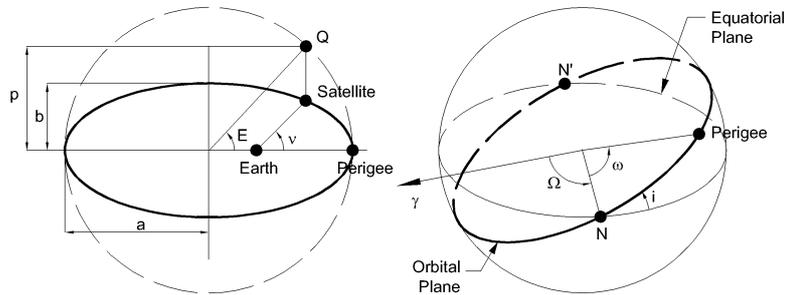


Figure 1: The Orbital Plane and Space

## 2 Determination of a preliminary orbit from two positions and time

From two position vectors of the satellite, and the time interval between the measurements,  $\vec{r}_1(t_1)$  and  $\vec{r}_2(t_2)$ , we can find the velocity of the satellite by  $\vec{v}_1 = \frac{\vec{r}_2 - f\vec{r}_1}{g}$ . Once  $\vec{r}_1$  and  $\vec{v}_1$  are known, orbital elements are easily obtained.

### True Anomaly Iteration

Let us set an initial estimation of  $\nu_1$  and check if the eccentricity and the semimajor axis satisfy the conditions of elliptic motion; if not, a new initial estimation of  $\nu_1$  is obtained. The function  $F$  is defined by the difference between the estimated and the real time in the first position of the satellite. While  $F$  is not near zero (i.e. estimated and real times are not similar), we apply the secant method, incrementing the estimation of  $\nu_1$  a small amount. Finally, the velocity of the satellite is obtained by the series  $f$  and  $g$ .

**Derivative-free high-order methods**

In order to improve the efficiency of true anomaly iteration method, we are going to replace the secant method (SEC) by derivative-free high-order methods. These methods are based on Newton’s method, where the derivative is replaced by some approximations. The first one is Steffensen’s method (SFT) (see [2]), with the iterative expression  $x_{k+1} = x_k - \frac{[f(x_k)]^2}{f(x_k+f(x_k))-f(x_k)}$  and quadratic convergence.

The method of Liu, Zheng and Zhao (LZZ) (see [3]) is a fourth-order scheme obtained by the composition of STF and Newton’s method. Its iterative expression is  $x_{k+1} = y_k - \frac{f[x_k,y_k]-f[y_k,z_k]+f[x_k,z_k]}{(f[x_k,y_k])^2} f(y_k)$ , where  $y_k$  is the STF iteration, and  $z_k = x_k + f(x_k)$ . The method of Cordero and Torregrosa (CT) (see [4]) is a fourth-order method designed as the composition of STF and Newton’s method. Its iterative expression is  $x_{k+1} = y_k - \frac{f(y_k)}{\frac{f(y_k)-f(z_k)}{y_k-z_k} + \frac{f(y_k)}{y_k-x_k}}$ . M8 method (see [5]) is an eighth-order scheme obtained by the double composition of STF and Newton’s method, with Padé approximants of degrees one and two, respectively. The iterative expression is  $x_{k+1} = u_k - \frac{f(u_k)}{b_2-b_1b_4}$ , where  $u_k$  is the previous fourth-order step,  $b_1 = f(u_k)$ ,  $b_2 = f[y_k, u_k] - b_3(y_k - u_k) + f(y_k)b_4$ ,  $b_3 = f[y_k, u_k, z_k] + b_4f[y_k, z_k]$ , and  $b_4 = \frac{f[y_k,u_k,x_k]-f[y_k,u_k,z_k]}{f[y_k,z_k]-f[y_k,x_k]}$ .

**3 Numerical Results**

Numerical computations have been carried out using variable precision arithmetic, with 500 significant digits, in MATLAB R2009B. The stopping criterion used is  $|x_{k+1} - x_k| < 10^{-498}$ . Each method is applied to several reference orbits, that can be found in [1]. The study is based on the initial estimation of  $\nu_0$  that causes the maximum number of iterations with the secant method. In Table 1 the difference between the two last iterations  $|x_{k+1} - x_k|$ , the number of iterations *iter*, the approximated order of convergence (*ACOC*) defined in [6] and the *e - time* calculated after 100 performances of the program are.

The composed methods LZZ, CT and M8 reach higher computational order of convergence than SEC and STF. The number of iterations decreases as the order of convergence increases. Moreover, CT has a better behavior than LZZ in terms of the mean elapsed time and M8 needs the minimum number of iterations to get the result and it is the fastest one. In general, modified

Ref	Parameter	SEC	STF	LZZ	CT	M8
I	$ x_{k+1} - x_k $	3.3885e-319	1.0233e-186	2.8184e-305	9.3325e-216	2.8949e-136
	iter	56	12	7	6	5
	ACOC	1.00	2.00	4.00	4.00	8.24
	e-time	18.2864	2.7509	2.5310	2.2068	1.4625
II	$ x_{k+1} - x_k $	3.9525e-323	1.7378e-258	1.3490e-173	1.7378e-133	1.6239e-74
	iter	63	15	7	6	5
	ACOC	1.00	2.00	4.00	4.00	7.75
	e-time	15.7550	3.0288	2.1762	1.8823	1.4507
III	$ x_{k+1} - x_k $	3.6308e-318	1.2023e-233	1.2589e-160	1.2882e-262	1.3317e-260
	iter	105	28	7	6	5
	ACOC	1.00	2.00	4.00	4.00	8.00
	e-time	22.5885	7.0931	2.5951	2.2656	1.6436

Table 1: Numerical Results for Reference Orbit I, II and III

methods are more stable and efficient than the classical one.

## References

- [1] P. R. Escobal, Methods of Orbit Determination. Florida, Robert E. Krieger Publishing Company, 1965.
- [2] J.M. Ortega, W.G. Rheinboldt, Iterative solutions of nonlinear equations in several variables, Academic Press, New York, 1970.
- [3] Z. Liu, Q. Zheng, P. Zhao. A variant of Steffensen’s method of fourth-order convergence and its applications. *Applied Mathematics and Computation*, 216: 1978-1983 (2010).
- [4] A. Cordero, J. R. Torregrosa. A class of Steffensen type methods with optimal order of convergence. *Applied Mathematics and Computation*, 217: 7653–7659 (2011).
- [5] A. Cordero, J. L. Hueso, E. Martínez, J. R. Torregrosa. Fourth and eighth-order optimal derivative-free methods for solving nonlinear equations. Proceedings of the 11th International Conference on Computational and Mathematical Methods in Science and Engineering, CMMSE 2011, Vol I 365-373. ISBN 978-84-614-6167-7.
- [6] A. Cordero, J. R. Torregrosa. Variants of Newton’s method using fifth-order quadrature formulas. *Applied Mathematics and Computation*, 190: 686–698 (2007)

# Modelling the attitude dynamics of the Basque population towards ETA before the passing of the *Law of Political Parties* (LPP)

Miguel Peco\*, Francisco-J. Santonja<sup>†</sup>, Ana-C. Tarazona<sup>‡</sup>,  
Rafael-J. Villanueva<sup>‡\*</sup>, Javier Villanueva-Oller<sup>⊗</sup>

(\*) Instituto Universitario General Gutiérrez Mellado (IUGM) de la UNED, Madrid (Spain)

(†) Departamento de Estadística e Investigación Operativa, Universitat de València, Valencia (Spain)

(‡) Instituto de Matemática Multidisciplinar, Universidad Politécnica de València, Valencia (Spain)

(⊗) Centro de Estudios Superiores Felipe II, Aranjuez, Madrid (Spain)

November 30, 2011

## 1 Introduction

ETA defines itself as "(...) a Basque socialist revolutionary organization for national liberation" [1]. One of its core demands is the creation of a Basque State, which would encompass the current three Basque Provinces and Navarra, in Spain, and three more French provinces. According to the Spanish Ministry of Internal Affairs, ETA has committed 829 murderers since 1968 [2].

In June 2002, the "Law of Political Parties" (LPP) was passed and its goals were "(...) to guarantee the democratic system and citizen's essential freedoms, by preventing some political parties from threatening democracy, justify racism and xenophobia or give political support to terrorist organizations" [3].

The LPP meant a substantial change in the anti-terrorist policy in Spain. In practical terms, once Batasuna was outlawed, this party could not present candidates to elections anymore. Considering Batasuna as the ETA's political wing, as the Supreme Court had stated, which meant that ETA was not going to be supported anymore from political institutions nor funded by public budgets.

The period of time considered for this study is from June 1995 to the passing of the LPP (June 2002). We consider this period because, in our opinion, the anti-terrorist policies were reasonably homogeneous.

---

\*Corresponding author, e-mails: mpeco.iugm@gmail.com, francisco.santonja@uv.es, actarazona@asic.upv.es, rjvillan@imm.upv.es, jvillanueva@pdi.ucm.es

This paper is organized as follows. In Section 2, data from Euskobarometro about the attitude of the Basque Country population towards ETA are retrieved and processed [4, Table 20]. Section 3 is devoted to build a model describing the attitude dynamics towards ETA in the Basque Country. Model parameters are estimated in Section 4 by fitting the model with the Euskobarometro data, predictions for the subsequent years are given and some conclusions are drawn.

## 2 Data

We have retrieved data from the Euskobarometro of November 2010 on the attitude of the Basque Country population towards ETA [4, Table 20]. The eight types of attitudes towards ETA that appear in the Euskobarometro are: Total support; Justification with criticism; Goals yes / Means no; Before yes / Not now; Indifferent; ETA scares; Total rejection; No answer. In order to simplify the model (the number of subpopulations) we group the eight attitudes in only three:

1. Support: people who have an attitude of support towards ETA. We consider the people with attitudes of "Total support" and "Justification with criticism" make up of this group.
2. Rejection: people who have an attitude of rejection against ETA. In this group we include the people with attitudes "Goals yes / Means no", "Before yes / Not now", "ETA scares" and "Total rejection". It could be dubious to include in this group the attitude "Goals yes / Means no", however, the fact is that there are parties and associations in the Basque Country with similar goals as ETA and they have a rejection attitude towards ETA because its violent means.
3. Abstention: people who have no opinion or have an indifferent attitude towards ETA, that is, the "Indifferent" and the "No answer" groups.

Data grouped in these three groups appear in Table 1 from June 1995 to June 2002, before the passing of the LPP.

## 3 Model building

Bearing in mind Table 1, we distinguish three main different attitudes towards ETA and divide the population of the Basque Country into the following three subpopulations (time  $t$  in years):

- $A_1(t)$ , the percentage of people in the Basque Country which have an attitude of support towards ETA at time instant  $t$ ,
- $A_2(t)$  is the percentage of people which have an attitude of rejection towards ETA at time  $t$ ,

Survey date	Support (%)	Rejection (%)	Abstention (%)
Jun 1995	7	85	8
Dec 1995	5	87	8
Dec 1996	6	87	7
Dec 1997	6	86	8
Dec 1998	5	85	10
Jun 1999	11	76	13
Jun 2000	8	87	5
Dec 2000	7	87	6
Jun 2001	3	90	7
Dec 2001	4	88	8
Jun 2002	2	96	2

Table 1: Percentage of people in the Basque Country with respect to their attitude towards ETA from Jun 1995 to Jun 2002, when the LPP was passed.

- $A_3(t)$  corresponds to the percentage of population in the Basque Country whose attitude towards ETA is not defined, abstain or simply they do not want to say their opinion, at time  $t$ .

$A_1(t)$ ,  $A_2(t)$  and  $A_3(t)$  are the variables of the mathematical model. The assumptions used to build the equations of the model are:

- A subpopulation  $A_i$ , whose people share a particular attitude towards a phenomenon, can influence the people's attitude of another subpopulation,  $A_j$ , towards the same phenomenon. This influence can be provoked either by direct contact, i.e., when people from  $A_i$  and  $A_j$  interact, or by indirect contact, i.e., through the interaction of a person in  $A_i$  with his environment.
- Regarding this latter way, in this context, it is assumed that the environment of a person in  $A_j$  is made up of the flows and channels of information able to reach his sensorial system. Note that reaching sensorial system does not imply necessarily reaching perception. Thus, alteration in that environment can provoke either changes in the attitude of that person in  $A_j$  or not. Environment alteration can be provoked, in its turn, by the behaviour of people from the other subpopulations among other factors, attitude being itself considered as a part of that behaviour.
- It is assumed that all people could access to all relevant information channels and flows, i.e., there is in principle a homogeneous environment affecting people of all the subpopulations. However, the interaction of a person with the environment varies on an individual basis, depending on both situational and non-situational factors. The individual initial attitude itself towards the subject of influence, for instance, is a non-situational factor which modulates environment influence, acting on that initial attitude either as an enabler or as a shield.

- It is not the goal of this work to clarify those factors of variation, but only to show the eventual changes in attitudes of the object populations and, if possible, to attribute those changes to the influence of other subpopulations, either directly or indirectly. However, a diffuse idea as to the involved processes, environment effectiveness differences etc., as a whole can be obtained from the model. The non-linear term  $\beta_{ij}A_iA_j$  being the one that models these influences, it is the parameter  $\beta_{ij}$  that, in some way, measures that environment effectiveness and includes the rest of the above mentioned factors.

Then, the system of differential equations that models the evolution of attitudes towards ETA in Basque Country over time is given by

$$A_1'(t) = (\beta_{21} - \beta_{12})A_2(t)A_1(t) + (\beta_{31} - \beta_{13})A_3(t)A_1(t), \tag{1}$$

$$A_2'(t) = (\beta_{12} - \beta_{21})A_2(t)A_1(t) + (\beta_{32} - \beta_{23})A_3(t)A_2(t), \tag{2}$$

$$A_3'(t) = (\beta_{13} - \beta_{31})A_3(t)A_1(t) + (\beta_{23} - \beta_{32})A_3(t)A_2(t). \tag{3}$$

The above system of differential equations can be represented by the diagram of Figure 1.

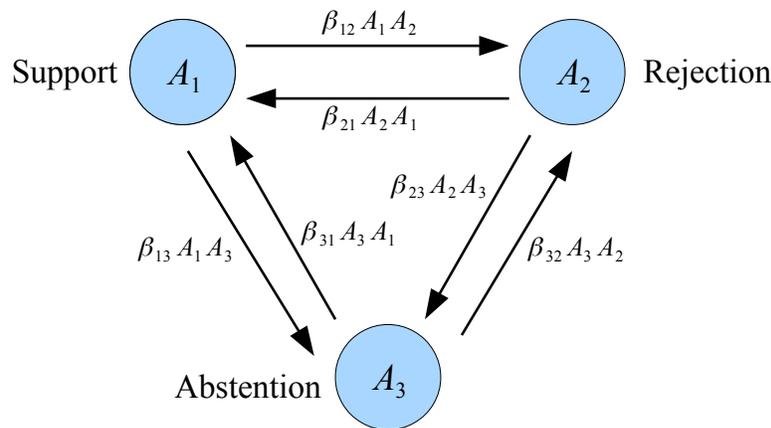


Figure 1: Graph depicting the model (1)-(3). Circles are the subpopulations and arrows represent the flow of people who change their attitude towards ETA over time.

## 4 Estimation of model parameters

The model has six unknown parameters  $\beta_{ij}, i, j = 1, 2, 3, i \neq j$  and we should estimate them taking into account that the model has to be as close as possible of data in Table 1.

Parameter	Value	Parameter	Value
$\beta_{12}$	0.0815425	$\beta_{21}$	0.0627668
$\beta_{13}$	0.000421055	$\beta_{31}$	0.182483
$\beta_{23}$	0.0317568	$\beta_{32}$	0.0216873

Table 2: Estimated model parameters.

To do that, we designed an algorithm in *Mathematica* [5] in order to compute the parameters which best fit the model with data of Table 1 in the least square sense. The values of these parameters appear in Table 2.

We can see the goodness of fitting graphically in the Figure 2.

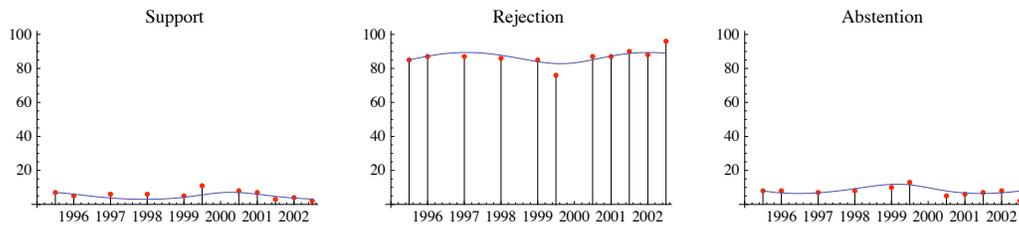


Figure 2: Graph representing the fitting. The lines are the corresponding model functions and the points are data from Table 1.

Solving the model with the parameters of Table 2, it can be predicted what would have happened if the LPP was not passed by the Spanish Parliament. In Figure 3, we can see the model predictions (line) for every subpopulation from June 2002 until June 2005 and data from Euskobarometro (points) in the same time interval. It can be observed some differences between the model prediction and the data. It may lead to say that LPP had some effect on the attitude of the Basque Country population towards ETA.

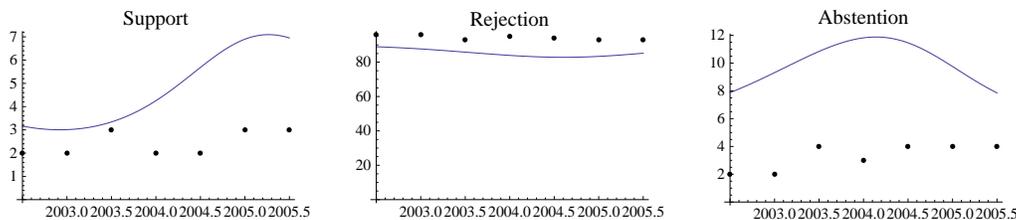


Figure 3: Graph of model prediction from June 2002 until June 2005 (line) with data from Euskobarometro. We can observe the differences between the model prediction and the data indicating that LPP had some effect on the attitude of the Basque Country population towards ETA.

## References

- [1] Communiqué of ETA January 8th 2011. Video published in GARA. Available in [http://www.gara.net/bideoak/110108\\_video/](http://www.gara.net/bideoak/110108_video/). A transcription of this Communiqué of ETA (in Spanish) is available in the web site [http://www.elpais.com/elpaismedia/ultimahora/media/201101/10/espana/20110110elpepunac\\_1\\_Pes\\_PDF.pdf](http://www.elpais.com/elpaismedia/ultimahora/media/201101/10/espana/20110110elpepunac_1_Pes_PDF.pdf).
- [2] Ministerio del Interior. (Spanish Ministry of Internal Affairs). Available in [http://www.mir.es/DGRIS/Terrorismo\\_de\\_ETA/ultimas\\_victimas/p12b-esp.htm](http://www.mir.es/DGRIS/Terrorismo_de_ETA/ultimas_victimas/p12b-esp.htm).
- [3] Ley Orgánica 6/2002 del 27 de junio. (Organic Law 6/2002, of Juny 27th). Available in <http://www.boe.es/boe/dias/2002/06/28/pdfs/A23600-23607.pdf>.
- [4] Euskobarómetro Survey November 2011. Available in the web site [http://alweb.ehu.es/Euskobarometro/index.php?option=com\\_docman&task=cat\\_view&gid=15&Itemid=97](http://alweb.ehu.es/Euskobarometro/index.php?option=com_docman&task=cat_view&gid=15&Itemid=97)
- [5] <http://www.wolfram.com/products/mathematica>.

# Analysis and assessment of models of pharmaceutical expenditure prediction.

Ruth Usó-Talamantes\*, Maria Caballer-Tarazona\*, Isabel Barrachina-Martínez\*\*, José Luis Trillo-Mata\* and David Vivas-Consuelo\*\*

\*Conselleria de Sanitat. Direcció general de Farmacia y Productos Farmacéuticos. Oficina de Farmacoeconomía.

\*\*Centro de Economía y Gestión de la Salud. Universitat Politècnica de València

## 1. Introduction

The Spanish National Health System is financed by taxes and decentralized into 17 regions. The services provision integrates hospital and primary health care in health districts. One of these regions is the Valencian Community. Here there are 24 health districts connected by a single integrated EHR. The patients are identified by a Health Card which is the nexus where the health information system connects.

Each health district receives an annual per capita quantity to attend to its assigned population, The flow of incoming patients from other health districts and outgoing to others generates an inter-centre turnover to establish the corresponding economic compensation [1]. In this context, systems capable of explaining utilization, costs and achievements as a function of morbidity have been developing.

In 2010, 110 million prescriptions were invoiced by the Valencian public health system. Regarding the pharmaceutical information, 97% of population is included in the EHR and 100% of prescriptions are electronic.

In addition, prescription drugs have become a major and growing component of health care - specifically pharmaceutical expenditure represents around 27% of public health spending in the Valencian region. However, studies aimed at understanding population patterns of drug consumption, cost and morbidity are uncommon.

Due to the fact that patient complexity is a major determinant of expenditure, methodologies aimed at allocating an average pharmaceutical expenditure to each health status are very useful in order to control and rationalise pharmaceutical expenditures, and the need to explain variations in the use and expenditures in medicines is increased.

Information systems must be capable of integrating data about services provided and given information about population served, but above all these systems must be capable of identifying health problems and those persons who need special attention because of their health status.

In recent years several models have been developed aimed at adjusting health risk and patient management. First we found the model based on chronic conditions classifications from the ATC-code of medicines [2-5]. Second, the Rx models also based on prescriptions [6]. Then we have the case mix models based on diagnoses Clinical Risk Groups (CRGs) [7], Adjusted Clinical Groups [8] and Diagnostic Cost

Groups, DCG. Finally the mixed models combining both, Diagnostic Cost Groups, DCG/Rx [9] and Adjusted Clinical Groups ACG/Rx [8, 10], have proven their efficacy.

In general, systems which measured the different cases according to the diagnostic, try to cluster the cases depending on their homogeneous level of resource consumption and at the same time have clinical significance.

Clinical Risk Groups (CRGs) is a relatively new system that classifies individuals into mutually exclusive categories and, using enrolment data, claims or other encounter level data and assigns each person to a severity level if s/he has a chronic health condition.

This system is used in order to classify patients in mutually exclusive categories based on the information of all contacts realized in any service of health care. All groups can be classified into 9 statuses from healthy to catastrophic condition. Inside these main groups various severity levels can be found. There are four grouping possibilities from 1076 groups to 38. CRG system can be used to understand use and consumption patterns of health services and for developing applications for risk and price adjusting [11].

In this paper we design a multivariate model in order to explain pharmaceutical expenditure in primary health care, comparing one system based on CRG with another based on chronic condition classification.

## **2. Data and Methods**

### *2.1. Data*

We used information from a population of 265,740. The period of analysis was 1 year. The variables considered were: annual pharmaceutical expenditure by patient, age, pharmaceutical co-payment status (0 for without co-payment and 1 for a co-payment of 40%), 18 groups of ATC-code as proxy of 18 chronic conditions; 9 Status of ACRG\_3 Groups. In Spain there is only co-payment for medicines. There are two categories of co-payment - pharmacy gratuity for pensioners and population without resources, and co-payment of 40% of the price for the remaining

The information was obtained from Health Electronic Record (EHR) [12] recently established in the Valencian Community.

### *2.4. Mathematical models*

Two models were developed to explain pharmaceutical expenditure in primary health care. The dependent variable was always the Neperian log of annual pharmaceutical expenditure by patient. The logarithm of the pharmaceutical expenditure is estimated as a better approach to its normal distribution of the variable, due the asymmetric probabilistic function of this variable.

As independent variables for the first model we take the following dummy variables:  $X_1$  sex (0 man, 1 woman);  $X_2$  pharmaceutical co-payment status (1 free; 0 with co-payment);  $X_3$  to  $X_{11}$  variables for groups from 1 to 8 or more chronic conditions:

$$Y = \alpha + X_1 \cdot \beta_1 + X_2 \cdot \beta_2 + X_3 \cdot \beta_3 + \dots + X_{11} \cdot \beta_{11} \quad (1)$$

In the second model we also include sex ( $X_1$ ) and pharmaceutical co-payment status ( $X_2$ ) and from  $X_3$  to  $X_{10}$  8 dummy variables for ACRG\_ groups:

$$Y = \alpha + X_1 \cdot \beta_1 + X_2 \cdot \beta_2 + X_3 \cdot \beta_3 + \dots + X_{10} \cdot \beta_{10} \quad (2)$$

The models were estimated by means of ordinary least squares (OLS). The goodness of fit is determined through the corrected  $R^2$  and the *Snedecor F*, and for each of coefficients *t Student* values are obtained to determine the level of significance.

The linear regression model allows us to explain the neperian log of pharmaceutical expenditure (dependent variable Y) in function of each considered independent variable ( $X_i$ ). The mathematical expression is:

$$\text{Pharmaceutical expenditure} = e^{(\alpha + \beta_1 \cdot x_1 + \dots + \beta_n \cdot x_n)} \quad (3)$$

### 3. Results

Two logarithm regression models were obtained (Table 1). Statistical significance checked by an F-test and T Student was over 99% for all of them. The first model achieved the highest level of explanation, 55%. The second model only reaches 52%.

Small but significant differences by gender are appreciated. Men received less prescriptions than women. This appeared as negative coefficient sign. The co-payment status does not appear in the model because it is highly correlated with age.

This second model gives a lower coefficient than we would expect for groups 8-9, given that this CRG should have the highest coefficient.

We interpret this as a liability of the model in explaining the pharmaceutical expenditure in primary care in all cases. These groups include malignancies and catastrophic diseases and these patients are treated in hospitals.

**Table 1. Regression models**

Variables	MODEL 1	Average cost		MODEL 2	Average cost	
		Model 1.a (€)	Model 1.b (€)		Model 2.a (€)	Model 2.b (€)
Constant	1.009			1.363		
Sex	0.234			0.172		
Co-payment 0%	1.37			1.651		
Without chronic condition		13.64	3.47			
1 chronic condition	1.304	50.25	12.77			
2 chronic conditions	2.24	128.12	32.56			
3 chronic conditions	2.969	265.60	67.49			
4 chronic conditions	3.55	474.85	120.66			
6 chronic conditions	3.939	700.64	178.04			
7 chronic conditions	4.22	927.97	235.80			
8 chronic conditions	4.432	1147.11	291.49			
More than 8 chronic conditions	4.692	1487.72	378.04			
ACRG 1					24.19	4.64
ACRG 2				1.091	72.02	13.82
ACRG 3				1.496	107.99	20.72
ACRG 4				2.325	247.40	47.47
ACRG 5				2.645	340.70	65.37
ACRG 6				3.572	860.92	165.17
ACRG 7				4.189	1595.59	306.13
ACRG 8				3.089	531.13	101.90
ACRG 9				3.108	541.31	103.86
R <sup>2</sup>	0.559			0.527		

a Women with free pharmacy; b Women with pharmacy co-payment

#### 4. Discussion

Both models used differed in the clinical meaning and grouping methodology.

The first model is obtained grouping the number of chronic conditions from 1 to 8 or more from data provided by the electronic prescription system. The main advantage is the reliability of data but the clinical significance is poor.

The CRG model is more useful from the clinical point of view, but obtaining the information for grouping is complicated. This system is designed for adjusting the total health spending. Thus, it is appropriate to include the hospital pharmaceutical cost in the model.

The two models achieved similar and valid levels of explanation in the context of other systems.

However CRG has a high clinical utility, thus making further research to determine the predictive power of this model at local level seem logical and necessary.

## **5. References**

- [1] Caballer-Tarazona M, Moya-Clemente I, Vivas-Consuelo D, Barrachina-Martínez I. A model to measure the efficiency of hospital performance. *Mathematical and Computer Modelling* 2010; 52:1095-102.
- [2] Clark DO, Von Korff M, Saunders K, Baluch WM, Simon GE. A chronic disease score with empirically derived weights. *Med Care* 1995; 33:783-95.
- [3] Johnson RE, Hornbrook MC, Nichols GA. Replicating the chronic disease score (CDS) from automated pharmacy data. *J Clin Epidemiol* 1994; 47:1191-9.
- [4] Von Korff M, Wagner EH, Saunders K. A chronic disease score from automated pharmacy data. *J Clin Epidemiol* 1992; 45:197-203.
- [5] Lamers LM, van Vliet RC. The Pharmacy-based Cost Group model: validating and adjusting the classification of medications for chronic conditions to the Dutch situation. *Health Policy* 2004; 68:113-21.
- [6] Fishman PA, Goodman MJ, Hornbrook MC, Meenan RT, Bachman DJ, O'Keeffe Rosetti MC. Risk adjustment using automated ambulatory pharmacy data: the RxRisk model. *Med Care*, 2003:84-99.
- [7] Hughes JS, Averill RF, Eisenhandler J, Goldfield NI, Muldoon J, Neff JM, Gay JC. Clinical Risk Groups (CRGs): a classification system for risk-adjusted capitation-based payment and health care management. *Med Care* 2004; 42:81-90.
- [8] Hanley GE, Morgan S, Reid RJ. Explaining prescription drug use and expenditures using the adjusted clinical groups case-mix system in the population of British Columbia, Canada. *Med Care* 2010; 48:402-8.
- [9] Stam PJ, van Vliet RC, van de Ven WP. Diagnostic, pharmacy-based, and self-reported health measures in risk equalization models. *Med Care* 2010; 48:448-57.
- [10] Kuo RN, Lai MS. Comparison of Rx-defined morbidity groups and diagnosis-based risk adjusters for predicting healthcare costs in Taiwan. *BMC Health Serv Res* 2010; 10:126.
- [11] Inoriza JM, Coderch J, Carreras M, Vall-Llosera L, Garcia-Goni M, Lisbona JM, Ibern P. [Measurement of morbidity attended in an integrated health care organization]. *Gac Sanit* 2009; 23:29-37.
- [12] De la Poza-Plaza EB-M, I; Trillo-Mata, JL et al. . Sistema de Prescripción y dispensación electrónica en la Agencia Valenciana de Salud. . *El Profesional de la Información* 2011; ; 20:9.