



EAMC2018

XI Encontro Acadêmico de Modelagem Computacional

Laboratório Nacional de Computação Científica - LNCC/MCTIC

01 a 02 de Fevereiro de 2018.

Livro de resumos



Laboratório
Nacional de
Computação
Científica

XI Encontro Acadêmico de Modelagem Computacional

Laboratório Nacional de Computação Científica - LNCC/MCTIC

01 a 02 de Fevereiro de 2018. Petrópolis - RJ.

Comitê Organizador

Camila Martins Saporetti

Daniel Gaspar Gonçalves de Souza

Diego Tavares Volpatto

Lucas dos Anjos

Marianna Novaes Martins

Natanael Júnior Soares Bento

Natasha do Nascimento Paiva

Weslley da Silva Pereira

Comitê Científico

Daniel Gaspar Gonçalves de Souza - UCP

Gastão Florêncio Miranda Junior - UFS

Iury Higor Aguiar da Igreja - UFJF

Leandro Tavares da Silva - CEFET/RJ

Marcos Garcia Todorov - LNCC

Sicilia Ferreira Ponce Pasini Judice - FAETERJ

Thiago de Oliveira Quinelato - Unicamp

Thiago José Machado - UFPB

Trabalhos completos

	Page
<i>Modelagem do espalhamento sonoro por fontes equivalentes: a potência das fontes como critério da qualidade da solução</i> Yves J.R. Gounot	3
<i>Reposicionamento para um robô de sondagem utilizando redes convolucionais</i> Daniel Luna, Jonathan Araujo Queiroz, André Cavalcante, Allan Kardec Barros	11
<i>Estudo de aproximações numéricas na resolução de problemas de Riemann</i> Ingrid B. Carneiro, Sandra M.C. Malta, Márcio R. Borges	20
<i>Scientific computing in statistical mechanics: time decay of orientation order in 2D hard disk system</i> A.G. França, R.S. Grisotto, B.M. Rocha, F.A.A.M.N. Soares, H.A. Fernandes, P.F. Gomes	27
<i>On enhanced reduced models for advection-diffusion problems</i> L.A. Mansilla Alvarez, P.J. Blanco, R.A. Feijóo	34

XI ENCONTRO ACADÊMICO
MODELAGEM COMPUTACIONAL

Modelagem do espalhamento sonoro por fontes equivalentes: a potência das fontes como critério da qualidade da solução

Yves J.R. Gounot

¹Departamento de Matemática Aplicada (IME / UFF)
Caixa Postal 24.020-140 – Niteroi, RJ – Brasil

ygounot@im.uff.br

Abstract. *The present paper investigates the emitted power by the sources obtained with two variants of the Equivalent Source Method applied to an acoustic scattering problem. The two variants, which differ by the minimization technique used, are the least square method (LSM) and the full field equations (FFE). It is found that with LSM the total power produced by the source set is significantly lower than that of its individual sources, which is due to some sources absorbing what others emit, what does not occur with FFE. These features are used to explain why, although much more instable than FFE, LSM has a better ability to reproduce the boundary condition. Also, a genetic algorithm which searches simultaneously the source best position and amplitude - is used for better assessing the correlation between the acoustic power emitted by the source set and the solution accuracy these sources are able to produce.*

Resumo. *No presente trabalho investigam-se as características da potência acústica emitida por fontes obtidas com duas variantes do Método das Fontes Equivalentes (ESM) aplicado ao problema de espalhamento sonoro. Essas variantes, que diferem pela técnica de minimização usada, são o método dos mínimos quadrados (LSM) e o método ‘full field equations’ (FFE). Os resultados mostram que, com o LSM, a potência emitida pelo conjunto de fontes é significativamente menor do que as potências das fontes individuais, umas absorvendo o que outras emitem, o que não ocorre com o FFE. Essas características são usadas para explicar porque, apesar de uma instabilidade maior, o LSM pode fornecer soluções mais precisas do que o FFE. Também, um algoritmo genético - que busca simultaneamente os melhores posicionamentos e amplitudes para as fontes - é usado para melhor avaliar a correlação entre a potência do conjunto de fontes e a qualidade da solução que estas fornecem.*

1 Introdução

A modelagem de campos sonoros, que sejam radiados por uma estrutura vibrante ou resultantes do espalhamento de uma onda que incide num corpo, é classicamente realizada via o método dos elementos de contorno. Desenvolvido mais recentemente, o método das fontes equivalentes (*ESM*), mais simples, e logo, com um custo computacional menor, representa uma alternativa interessante. A ideia do *ESM* reside na substituição do corpo

por fontes pontuais - monopolos ou multipolos - colocadas ‘dentro’ do corpo e cujas amplitudes são determinadas de modo que o campo de velocidades gerado no ‘contorno’ reproduza da melhor forma possível o campo correspondente à situação real. Porém, o método apresenta duas desvantagens: não há sempre convergência da solução com o aumento do número de monopolos e a precisão da solução depende fortemente do posicionamento e do número de fontes, dois fatores intrinsecamente ligados (ver Gounot et al. (2005)). Ademais, não existem regras gerais para escolher um conjunto de fontes adequado e diversas abordagens já foram propostas para o posicionamento das fontes (ver, por exemplo, Kropp e Svensson (1995) ou ainda Song e Koopmann (1991)). Gounot e Musafir (2009) investigaram outra abordagem, que visa a resolver o problema de maneira global. Usaram um código computacional que combina o *ESM* com algoritmos genéticos para buscar simultaneamente as amplitudes e posições das fontes responsáveis pelo menor erro no contorno possível. No presente trabalho, soluções do problema de espalhamento tridimensional obtidas com duas variantes do *ESM* - uma usando o *método dos mínimos quadrados (LSM)* e outra as ‘*Full Field Equations*’ (*FFE*) desenvolvidas por Ochmann (1999) - são comparadas. Resultados sobre a potência das fontes são usados para explicar porque, apesar de uma instabilidade numérica significativamente maior, o *LSM* pode, em certos casos, fornecer soluções mais precisas do que o *FFE*.

2 Formulação Teórica

2.1 O Problema de Espalhamento

Quando a superfície de um corpo imerso em um fluido vibra com uma frequência ω , um campo de pressão p é radiado (Fig. 1a). No domínio da frequência, p é a solução do problema de valor de contorno de Neumann, isto é, p satisfaz à equação de Helmholtz (1)

$$\{\nabla^2 + k^2\}p(\mathbf{x}, \omega) = 0 \quad (1)$$

$$\frac{\partial p(\mathbf{x}, \omega)}{\partial n} = -ik\rho_0 c_0 \bar{u}_n(\mathbf{x}, \omega) \quad (2)$$

onde $k = \omega/c_0$ e c_0 é a velocidade de propagação do som, e também à condição de contorno (2), onde ρ_0 é a densidade média do meio, \bar{u}_n é a velocidade normal prescrita em S , \mathbf{x} e \mathbf{x}_S denotando pontos no domínio de propagação e no contorno, respectivamente. O problema de espalhamento gerado pela incidência de uma onda num corpo pode ser considerado como um problema de radiação (ver Fig. 1b), no qual a pressão espalhada p_{sc} é solução do problema de valor de contorno de Neumann (Eq. 1-2) e a velocidade normal prescrita é, para um corpo rígido, dada pelo oposto do componente normal em S da velocidade que seria gerada pela onda incidente na ausência do corpo.

2.2 O Método das Fontes Equivalentes

O método das fontes equivalentes substitui o corpo real - radiador/espalhador - por um conjunto de M fontes virtuais posicionadas em pontos \mathbf{y} na região Ω_I que devem, idealmente, produzir no contorno S a mesma velocidade normal do que no problema original.

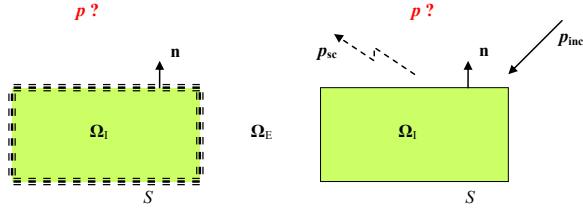


Figura 1: Representação dos problemas de radiação (a) e de espalhamento (b).

Os campos de pressão e de velocidade gerados por essas fontes são dados respectivamente pelas equações (3) e (4), onde A_m denota as amplitudes das fontes (a serem determinadas) e $g(\mathbf{x}, \mathbf{y})$ é uma função que descreve sua radiação.

$$p_{sc}(\mathbf{x}) = \sum_{m=1}^M A_m g(\mathbf{x}, \mathbf{y}_m) \quad (3)$$

$$v_n^{sc}(\mathbf{x}) = \frac{-1}{i\omega\rho_0} \sum_{m=1}^M A_m \frac{\partial g(\mathbf{x}, \mathbf{y}_m)}{\partial n} \quad (4)$$

As funções de radiação mais comumente usadas são a função de Green de campo livre para fontes monopolares, ou, para fontes multipolares, expansões em funções de ondas esféricas. Ao usar um conjunto de fontes com um número finito de monopolos, a condição de contorno não é exatamente satisfeita e um erro de velocidade local ε_v é gerado. Para uma configuração de fontes dada, as amplitudes ‘ótimas’ A_m são obtidas minimizando o erro global de velocidade no contorno. As duas variantes do *ESM* comparadas neste trabalho diferem pela técnica de minimização usada: o *método dos mínimos quadrados (LSM)* - o mais frequentemente usado, e um chamado ‘*full field equations*’ (*FFE*) desenvolvido por Ochmann (1999). Enquanto o *LSM* consiste em minimizar a soma do quadrado dos erros de velocidade no contorno, no *FFE*, ε_v é minimizado usando o método dos resíduos ponderados, no qual as funções de ponderação são o conjugado complexo das funções de ondas esféricas. Esta última variante apresenta a vantagem de gerar matrizes com diagonal dominante, responsável por uma estabilidade numérica muito maior que no caso da variante *LSM*. Ao minimizar o erro global de velocidade, obtém-se ‘a solução’, i.e., as M amplitudes A_m das fontes responsáveis pela melhor reconstituição da condição de contorno.

3 Simulações Numéricas e Resultados

3.1 O caso em estudo

Os resultados apresentados nesse trabalho são relativos ao problema de espalhamento acústico no qual uma onda plana incide normalmente a um corpo rígido com a forma de um cubo de lado L igual ao comprimento de onda da onda incidente (i.e., $kL = 2\pi$). Os monopolos são distribuídos uniformemente em suportes simples de implementação (um linear paralelo ao vetor de onda \mathbf{k} e um circular) situados no plano médio do cubo. O tamanho dos suportes é determinado multiplicando às dimensões do cubo por um fator de redução a estritamente inferior a 1. Quanto menor a , menor o suporte das fontes, e

consequentemente, para um mesmo M , quanto mais denso e concentrado no centro do cubo será o conjunto. A fim de poder avaliar a precisão das soluções obtidas, é usado o erro de contorno normalizado,

$$e_{BC} = \int_S |\varepsilon_v|^2 dS / \int_S |\bar{u}_n|^2 dS. \quad (5)$$

3.2 A Potência Acústica

A potência W_S gerada dentro de uma superfície fechada pode ser computada a partir da intensidade acústica em S como expresso pela equação (6), na qual $*$ denota o conjugado complexo, p e v_n representam a pressão e a componente normal da velocidade total.

$$W_S = \frac{1}{2} \Re \left\{ \oint_S p v_n^* dS \right\} \quad (6)$$

$$W_m = \frac{1}{2\rho_0} \Re \{ p(\mathbf{y}_m) q_m^* \} \quad (7)$$

Para um corpo espalhador rígido, a condição de contorno é de velocidade normal nula e, portanto, $W_S = 0$. Entretanto, ao usar um método numérico, o erro de velocidade ε_v gerado em S é responsável pela emissão (ou absorção) de uma potência sonora líquida W_{ESM} , obtida substituindo v_n por ε_v na equação (6). W_{ESM} pode ser diretamente computada a partir das amplitudes A_m , dado que a potência emitida por uma fonte pontual monopolar de ‘strength’ q_m posicionada em \mathbf{y}_m é expressa na equação (7) onde $p(\mathbf{y}_m)$ é a pressão acústica total em \mathbf{y}_m e $q_m = (i\omega)^{-1} A_m$. Esse procedimento para obtenção da potência acústica é exato e será usado para a computação de W_{ESM} . Todas as potências são aqui normalizadas pela potência da onda incidente.

3.3 Resultados e discussão

A figura 2 mostra, para os suportes de fontes circular e linear, a influência do número de monopolos (M variando de 2 a 50) no erro de contorno e_{BC} , na potência gerada pelo conjunto de fontes W_{ESM} e na amplitude média das fontes (normalizada), computados com as duas variantes do método das fontes equivalentes, o *LSM* e o *FFE*. Em primeiro lugar, podemos constatar a estabilidade e robustez do *FFE*, o e_{BC} decrescendo (mesmo que lentamente) com o aumento do número de monopolos. Ademais, as soluções obtidas são geradas por conjuntos de fontes apresentando potência e amplitude média quase constantes. Quanto à influência do tamanho do suporte, as melhores soluções são obtidas para $a = 0.5$, i.e. quando as fontes não são perto demais do contorno nem concentradas demais no centro. Quanto ao *LSM*, os resultados mostram uma forte instabilidade quando M ultrapassa um certo valor (que depende sensivelmente da posição das fontes), valor além do qual as amplitudes e potências podem atingir valores extremamente altas, o que origina uma degradação da qualidade das soluções. Apesar disto, o resultado talvez mais relevante é que, abaixo deste valor de M , as soluções *LSM* são significativamente melhores do que as *FFE*. Ademais, nesta faixa de M inicial, os resultados mostram que quanto menor o suporte, quanto melhores as soluções. É precisamente por poder ocasional soluções mais precisas com um número de fontes baixo que os conjuntos gerados pelo *LSM* merecem mais ampla investigação. Se considerar apenas as soluções *LSM* boas (as

com $e_{BC} < 0.7$, limite materializado pela linha verde na Fig. 2), aparece que os conjuntos que as geram sempre apresentam uma potência total emitida (ou absorvida) inferior a 25% da potência acústica recebida pelo corpo. Outra característica relevante é que isto ocorre apesar das fontes terem uma amplitude média muito elevada, da ordem de até 10^3 vezes maior do que a amplitude das fontes FFE. A diferença de comportamento observado com o suporte circular e o linear (a degradação ocorrendo com valores de M menores com o suporte linear) pode ser explicada pelo fato de que, para um M dado, o conjunto de fontes linear é muito mais denso e concentrado que o circular.

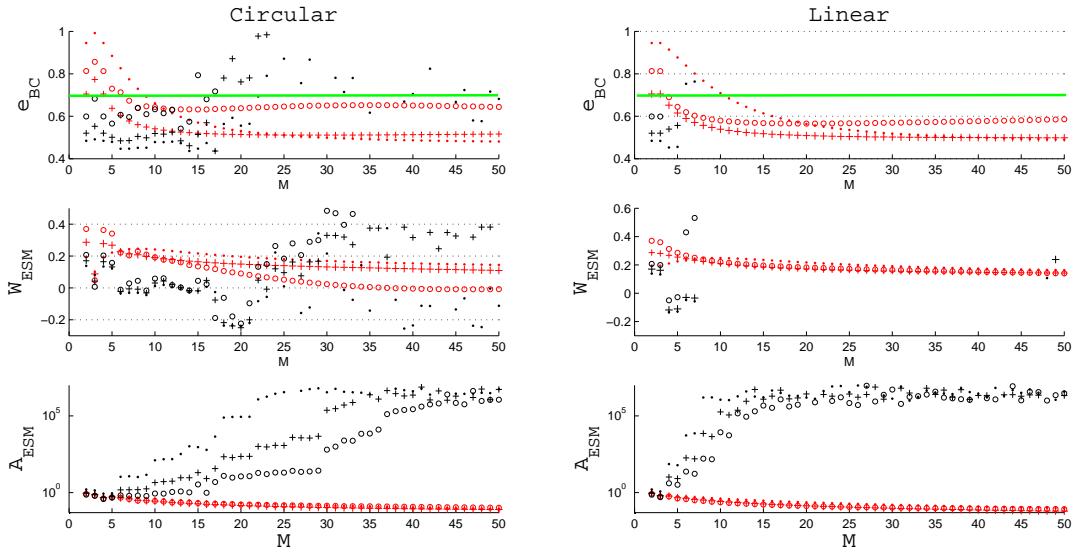


Figura 2: e_{BC} , W_{ESM} e A_{ESM} , obtidos com LSM (preto) e com FFE (vermelho) em função do número de monopolos para a configuração circular (1^a coluna) e linear (2^a coluna), para 3 tamanhos de suportes: $a = 0,25$ (.), $0,5$ (+) e $0,75$ (o)

A figura 3 ilustra, no caso da configuração circular com 16 monopolos, as características (amplitude, fase e potência das fontes individuais) representativas das soluções que o FFE e o LSM fornecem. Os resultados mostram que as fontes LSM apresentam uma heterogeneidade em amplitude e em fase muito maior do que as fontes FFE . Também, elas se arrumam aproximadamente em uma série de dipolos locais (2 monopolos justapostos sendo basicamente em oposição de fase) enquanto, no caso do FFE , o comportamento ‘dipolar’ só se manifesta globalmente, o conjunto se arrumando em 2 pacotes de fontes em oposição de fase. Esta característica se reflete ao nível das potências individuais, mostrando, no caso das fontes LSM , uma importante interação energética, algumas fontes absorvendo o que outras emitem. Consequentemente, com o LSM , a potência total (última coluna das figuras W_m) é sempre menor do que a média das amplitudes individuais. Quanto ao FFE , a interação entre fontes é muito menor, e a potência total é basicamente a soma das potências individuais. A maior interação que ocorre com as fontes LSM é, em princípio, uma característica desejável para melhor reconstituir a condição de contorno com baixos níveis de potência global, o que contribui para explicar as melhores soluções obtidas com LSM .

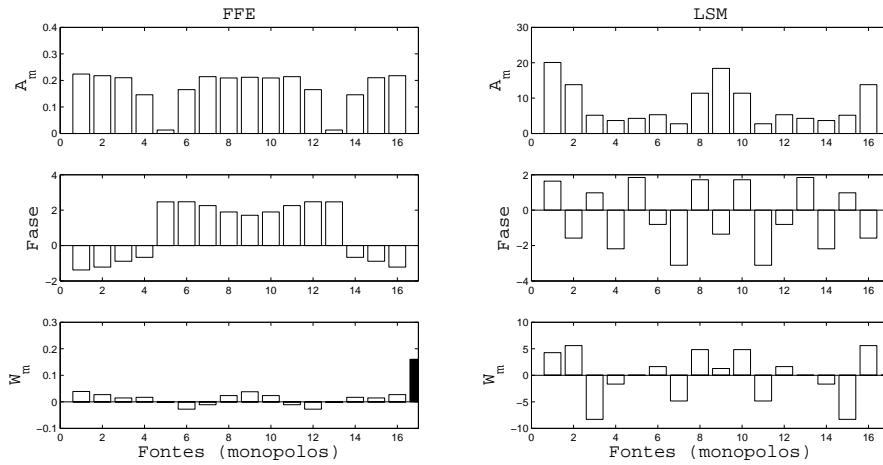


Figura 3: Amplitudes, fases e potências das fontes obtidas com *FFE* (1^a col.) e *LSM* (2^a col.) para a configuração circular com 16 monopolos.

3.4 Abordagem pelos Algoritmos Genéticos

Um código, *ESGA*, que combina o *ESM* com algoritmos genéticos (*AG*) foi desenvolvido a fim de, ao minimizar o erro de contorno e_{BC} , determinar simultaneamente as amplitudes complexas e as posições ótimas das fontes responsáveis pela melhor solução que um conjunto de M monopolos pode fornecer. A vantagem dos algoritmos genéticos é que eles limitam os riscos de cair num mínimo local (para mais detalhes, ver Goldberg (1989)). A Figura 4 mostra 4 evoluções típicas (para conjuntos de 2, 3, 4 e 5 monopolos) da função de custo minimizada (e_{BC}) ao longo de 500 gerações/iterações.

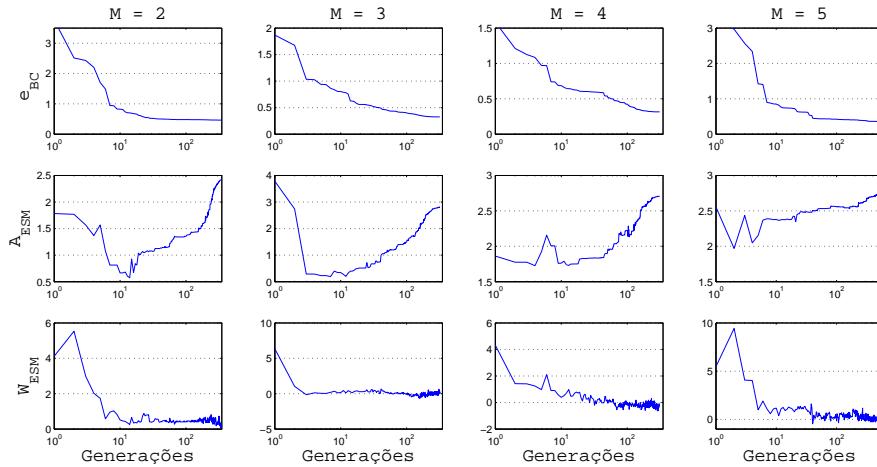


Figura 4: Erro de contorno minimizado pelo *ESGA* (1^a linha) e amplitude média (2^a linha) e potência (3^a linha) das fontes soluções, para conjuntos de 2, 3, 4 e 5 monopolos.

A cada geração, a melhor solução foi armazenada para computar A_{ESM} , a amplitude média, e W_{ESM} , a potência desses conjuntos. Vale insistir sobre o fato de que, nesse processo, o algoritmo minimiza apenas o e_{BC} , e não A_{ESM} e W_{ESM} . Todos os ensaios

feitos apresentam o mesmo aspecto geral: na fase inicial da evolução, A_{ESM} e W_{ESM} decrescem fortemente junto com o e_{BC} . Numa segunda fase, na qual a minimização do e_{BC} se torna mais lenta, observe-se que a potência se mantém quase constante, enquanto ocorre uma inversão na evolução da amplitude média das fontes, que começa a crescer até um valor que praticamente não depende do número de monopolos do conjunto. Os valores para as quais tendem as potências são compatíveis com os encontrados (ver Fig.2), da ordem de 30%. Isto mostra que um conjunto de fontes apropriado deve gerar uma potência baixa, porém obtida com fontes de amplitude relativamente altas, ou seja, que devem ocasionar interação energética importante entre si, um resultado conforme ao mostrado na Figura 3 ao nível das fontes individuais.

4 Conclusão

Os resultados mostram que, apesar de uma forte instabilidade, o *LSM* pode levar, quando o número de fontes é relativamente baixo, a soluções mais precisas do que o *FFE*. Esta vantagem é devida ao fato de que as fontes *LSM* apresentam uma forte interação energética entre si, o que não ocorre com o *FFE*. Quantitativamente, foi observado que, para todas as soluções de boa precisão (independente do número de monopolos, da forma e do tamanho do suporte), a potência do conjunto de fontes não pode ultrapassar 25% da potência recebida pelo corpo espalhador. Além disso, os resultados obtidos com o *ESGA* mostram que esta potência relativamente baixa deve ser gerada por fontes de amplitude altas. Esses dois critérios, que garantem uma boa qualidade da solução *LSM*, deverão ser aproveitados num trabalho futuro, para - ao integrá-los como penalidades na função de aptidão a ser minimizada - acelerar a convergência do algoritmo genético *ESGA*.

Referências

- D.E. Goldberg. Genetic algorithms in search, optimization, and machine learning. In: Addison-Wesley (editor), *Genetic Algorithms in search, optimization, and machine learning*. Massachusetts Ltd., 1989.
- Y.J.R. Gounot e R.E. Musafir. Genetic algorithms: a global search tool to find optimal equivalent source sets. In: *Journal of Sound and Vibration*, páginas 282–298. 2009.
- Y.J.R. Gounot, R.E. Musafir, e J.G. Slama. A comparative study of two variants of the equivalent sources method in scattering problems. In: *Acustica United with Acta Acustica*, páginas 860–872. 2005.
- W. Kropp e P.U. Svensson. Application of the time domain formulation of the method of equivalent sources to radiation and scattering problems. In: *Acustica*, páginas 528–543. 1995.
- M. Ochmann. The full-field equations for acoustic radiation and scattering. In: *Journal of the Acoustical Society of America*, páginas 2574–2584. 1999.
- L. Song e G.H. Koopmann. Numerical errors associated with the method of superposition for computing acoustic fields. In: *Journal of the Acoustical Society of America*, páginas 2625–2633. 1991.

XI ENCONTRO ACADÊMICO
MODELAGEM COMPUTACIONAL

Reposicionamento para um Robô de Sondagem utilizando Redes Convolucionais

Daniel Luna¹, Jonathan Araujo Queiroz¹, André Cavalcante¹, Allan Kardec Barros¹

¹Laboratório de Processamento da Informação Biológica

Universidade Federal do Maranhão (UFMA)

– Av. dos Portugueses, 1966, Bacanga– CEP 65080-805 – São Luís, MA – Brasil

dantmec@gmail.com, queirozjth@gmail.com, abcborges@gmail.com, akduailibe@gmail.com

Abstract. *Convolutional Neural Networks (CNN) has been successfully used for positioning operations on standalone platforms, for environments whose scenario complexity and image pre-processing capabilities are decisive factors for the success of the classification (repositioning attitudes). The objective of the present study is to develop an autonomous approximation system with the base classification of images by a CNN. The results show the superior CNN (accuracy 82%) to a method that uses Decision Threshold and Markers (accuracy 51.8%), developed an initial test of the approach system.*

Resumo. *As Redes Neurais Convolucionais (do inglês Convolutional Neural Networks-CNN) tem sido utilizadas com sucesso para operações de posicionamento em plataformas autônomas, para ambientes cuja complexidade do cenário e os recursos de pré-processamento da imagem são fatores decisivos para o sucesso da classificação (atitudes de reposicionamento). O objetivo do presente estudo é desenvolver um sistema de aproximação autônomo com base na classificação de imagens por uma CNN. Os resultados obtidos mostram a CNN superior (Acurácia 82%) a um método que utiliza Limiar de Decisão e Marcadores (Acurácia 51,8%), desenvolvido para teste inicial do sistema de aproximação.*

Palavras-chave: *Reposicionamento, Redes Convolucionais, Marcadores, Automatização, Imagem.*

1 Introdução

É crescente o uso de plataformas robóticas que se readaptem a ambientes e aplicações, utilizando visão computacional. A readaptação dos algoritmos embarcados nestas plataformas se faz necessária devido à variação nas condições de luminosidade, formas e cores dos objetos de interesse e mudanças nas características do cenários que as cercam.

Umas das formas de contornar esses problemas é através da extração de características de uma imagem, como forma de reduzir a quantidade de informação a ser processada. Métodos como *contorno, cor e limiar* tem sido utilizados por Viren Pereira e Sequeira (2014), *Wavelets* por Kamarul Hawari Ghazali e Hussain (2007), PCA (*Dependent*

Component Analysis) por Xiao (2010) ou ICA (*Independent Component Analysis*) por Wenzhi Liao (2008), para redução de dimensionalidade ou extração de características. O posterior uso de algoritmos de aprendizado de máquina (ANN- *Artificial Neural Networks* ou sistemas *Neuro-Fuzzy*), tornam os algoritmos readaptáveis à diferentes aplicações. Algumas dessas técnicas possuem variações à translação de objetos na imagem, ruídos ou variações de luminosidade (Viren Pereira e Sequeira, 2014).

No entanto, como mostrado por Goodfellow et al. (2016), Alessandro Giusti (2014) e Xie (2016) as *CNN* (*Convolutional Neural Networks*) ou Redes Neurais Convolucionais (classe específica de Redes Neurais) possuem a capacidade de extração de características de cor e forma, invariantes à translação, ruídos ou variações de luminosidade, problema comum ao processamento de imagem mostrado por Gonzalez (2007).

Neste contexto, é utilizada uma *CNN* e um método baseado em limiar de decisão para a classificação do grau de inclinação do sistema de sondagem descrito na Fig. 1.

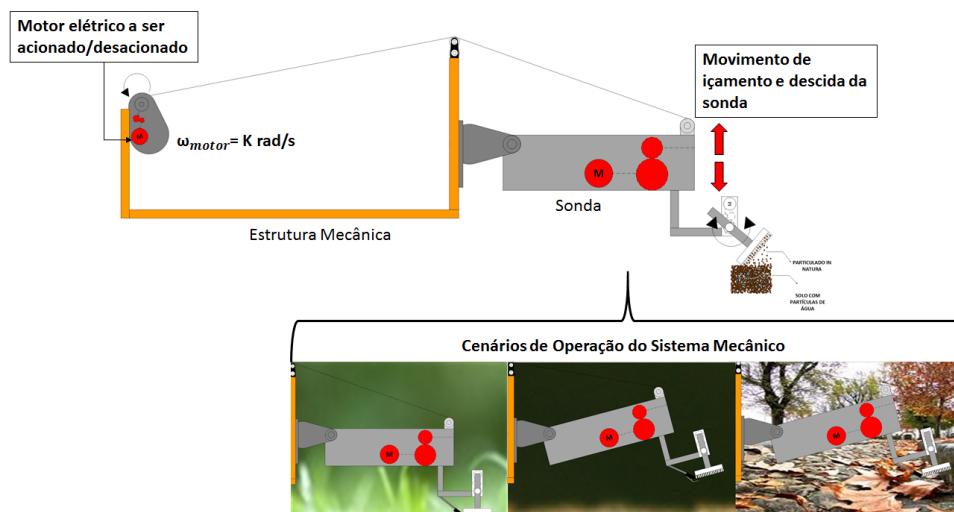


Figura 1: Diagrama do sistema de aproximação e cenários. Sistema mecânico caracterizado pela plataforma (*Estrutura Mecânica*), sistema a ser reposicionado (*Sonda*) e seus possíveis ambientes de operação (Cenários de Operação).

Os resultados das classificações das imagens adquiridas pelo sistema são apresentados de acordo com as métricas de acurácia, precisão e sensitividade.

1.1 Motivação

O trabalho proposto foi motivado devido ao desenvolvimento de uma plataforma robótica, cujo sistema de sondagem deveria realizar operações de posicionamento de forma autônoma.

O processamento de imagem foi escolhido devido ao número de informações contidas na imagem adquirida (sistema de sondagem, ambiente de operação e ponto desejado para o reposicionamento no cenário). A utilização do aprendizado de máquina foi motivado devido a variação das estatísticas das imagens adquiridas (objeto artificial em meio

a cenários naturais com variação de cor, forma e estatísticas dos pontos para os quais a sonda deveria ser reposicionada)(Torralba e Oliva, 2003).

2 Materiais e Métodos

Para realização dos experimentos na plataforma robótica, foi implementado um sistema de aquisição de imagens (Fig. 2), juntamente com a instalação de alguns marcadores de cor específica (Fig. 5). Os testes dos métodos que contemplam tanto as redes Convolucionais, como o método baseado no limiar de decisão são feitos com base nos sistemas de aquisição implementados (Figuras 2 e 5).

2.1 Materiais

Inicialmente, a aquisição da imagem é realizada por uma câmera com sensor de imagem do tipo *CMOS* (*Complementary Metal Oxide System*), com resolução aproximada de 2MP (dois Mega Pixels), sistema de cor RGB (Matrizes *Red*, *Green* *Blue*) com profundidade pixel de 8 bits (valores de cada pixel variando entre 0 e 255). O diagrama do processo e do hardware de classificação podem ser observados na Figura 2.

O drive de acionamento é composto por uma placa de desenvolvimento *Arduino®* modelo *UNO*, dispositivo de acionamento do motor elétrico (tensão de operação de 5 Volts) e motor elétrico *DC* (*Direct Current*) com tensão de operação nominal em 5 Volts e corrente de operação em torno de 800 miliampéres.

2.2 Método

Os sistemas de aquisição da imagem (*CÂMERA*), processamento da imagem (*SERVIDOR*) e acionamento do motor elétrico (*DRIVE DE ACIONAMENTO*), utilizados no processo de aproximação estão descritos na Fig. 2:

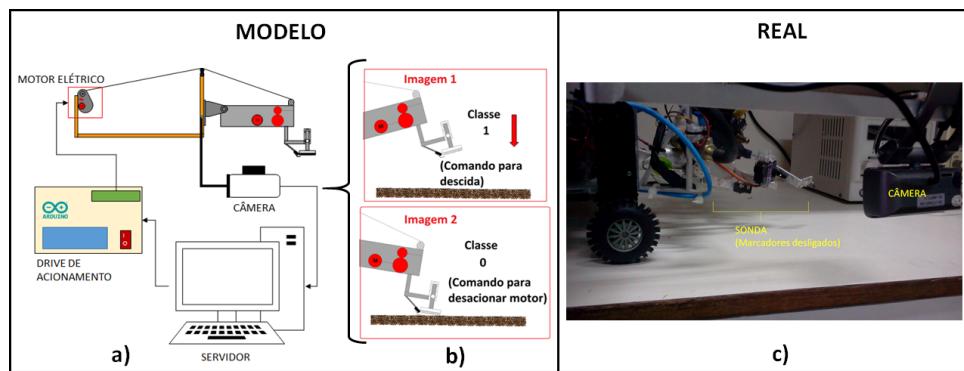


Figura 2: Diagrama de funcionamento de sistema de aproximação. a) Método de aquisição e classificação para reposicionamento do sistema de sondagem ; b) Imagens das classes de inclinação e a descrição dos comandos de reposicionamento; c) Sistema Real de Aproximação.

Conforme a figura 2 o processo de aproximação se divide em 3 etapas sendo: aquisição da imagem por uma câmera posicionada na parte lateral da estrutura mecânica (primeira etapa- Fig. 2a) e classificação da imagem por um algoritmo implementado no servidor (segunda etapa). Com base no valor numérico de classificação é gerado um sinal para o acionamento ou parada do motor elétrico (terceira etapa- Fig. 2b), ocorrendo o reposicionamento da sonda (Fig. 2a).

O modelo de classificação proposto pelos autores contempla dois níveis de afastamento da sonda (Fig. 2b), correspondendo a classe 1 (sonda distante do ponto desejado) e classe 0 (sonda no ponto desejado). A classe 1 gera, na saída do dispositivo de acionamento, um sinal de corrente elétrica para o acionamento do motor elétrico e a classe 0 suspende o fornecimento desse sinal.

O braço mecânico (caracterizado como *Sonda* na Fig. 1) é um mecanismo de 1 Grau de Liberdade (G.D.L), reposicionado angularmente através de um sistema de içamento por cabo (movido pelo motor elétrico), gerando o movimento de subida e descida, movendo a sonda ao ponto desejado.

2.3 Modelo de classificação utilizando a CNN

Os processos de extração de características e posterior classificação pela *CNN* estão descritos na Fig. 3:

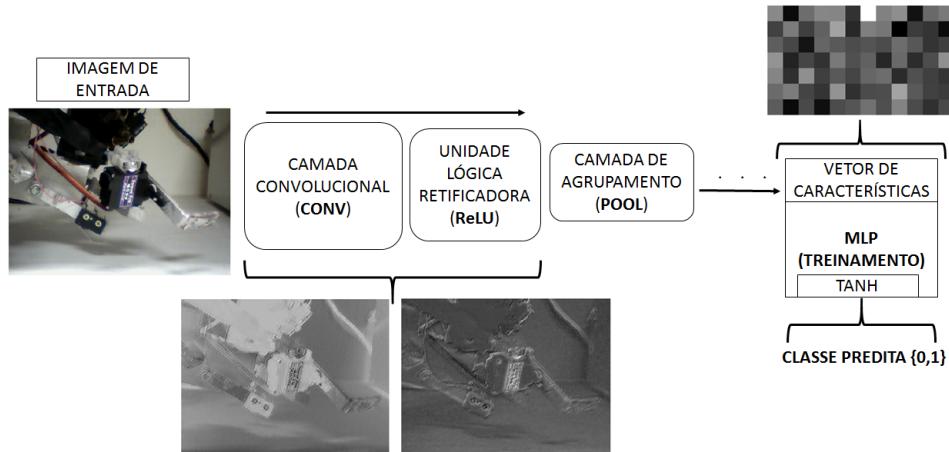


Figura 3: Diagrama de blocos da *CNN*. Imagem do sistema mecânico real sendo convoluída com o filtros na camada convolucional *CONV*, retificada pela camada *ReLU* e agrupada ou sub-amostrada na camada *POOL*, sendo gerado então um conjunto de valores caracterizado como *VETOR DE CARACTERÍSTICAS* inserido na rede *Multi-Layer-Perceptron* (MLP) tipo *Fully-Conected*.

A teoria das Redes Convolucionais propostas no trabalhos de Goodfellow et al. (2016), Xie (2016) e Wu (2017) mostram as principais estruturas das camadas convolucionais (*Conv*), de Retificação Lógica (*ReLU*) e Agrupamento (*Pooling*). O processo de convolução da imagem com o núcleo do filtro na camada convolucional (Fig. 3) é dado por:

$$ma[i][j] = \frac{\sum_m \sum_n k_{i,j} i m_{ch(i+m),(j+n)}}{\sum_n k_{(ch)(n)}} + rcc_{ch(i,j)}, \quad (1)$$

sendo $ma_{[i],[j]}$ os elementos que formam o *Mapa de Ativação MA* a matriz de saída da camada convolucional, m e n os índices dos elementos $k_{(m,n)}$ do núcleo do filtro e $i m_{ch(i+m),(j+n)}$ os elementos do respectivo canal (ch) da imagem de entrada (Topo da Fig. 3). Em seguida, sofrem uma combinação linear e são adicionadas ponderações $rcc_{ch(i,j)}$ dos outros planos de cor à MA.

O modelo de geração do mapa de ativação está descrito como:

$$\mathbf{MA} = Red(\mathbf{CMC}) + Green(\mathbf{CMC}) + Blue(\mathbf{CMC}), \quad (2)$$

sendo a **CMC** a matriz formada pelos elementos de contribuição rcc_{ch} do respectivo plano de cor ou canal ch .

Após a geração dos mapas de ativação (Eq. 2), estes sofrerão uma retificação lógica (camada *ReLU-Rectified Logic Unit*), cuja função aplicada ao elemento é a função *sigmoide-logística* ou função de retificação, descrita para cada elemento do mapa de ativação MA, ou seja:

$$ma_{[i][j](ReLU)} = \ln(1 + e^{ma_{[i][j]}}). \quad (3)$$

Por fim, a matriz ou mapa de ativação retificado \mathbf{MA}_{ReLU} (Eq. 3), passa por um processo de agrupamento ou (*max-pooling*), dado por:

$$ma_{[i][j](Pool)} = \max(ma_{[i][j](ReLU)} : ma_{[i+m][j+n](ReLU)}), \quad (4)$$

sendo $ma_{[i][j](Pool)}$ o respectivo elemento do mapa de agrupamento \mathbf{MA}_{pool} , gerado pelo valor máximo encontrado em cada sub-matriz formada pelos elementos $ma_{[i][j](ReLU)} : ma_{[i+m][j+n](ReLU)}$. As sub-matrizes selecionadas para extração dos seus valores máximos não são sobrepostas. O elemento $ma_{[i][j](ReLU)}$ é o primeiro da sub-matriz (contida no mapa de ativação \mathbf{MA}_{ReLU}) e $ma_{[i+m][j+n](ReLU)}$, o último desta sub-matriz.

O método de aprendizado da *CNN* é baseado no (*Steep Gradient Descent-SGD*) ou *método da descida mais íngreme*.

2.3.1 Banco de Dados (*CNN*)

A organização das imagens para alimentação da rede foi realizada com aquisição de 10.000 imagens ou (*frames*) em 5 diferentes cenários (2.000 imagens por cenário), estruturados de acordo com o diagrama da Fig. 4.

O banco de imagens foi dividido e rotulado em imagens *classe 0* e imagens *classe 1*. A classe 0 corresponde ao sistema mecânico próximo ao solo (ponto desejado), em diferentes cenários. A classe 1 correspondem a toda estrutura mecânica distante do ponto desejado (Fig. 4), fixa em um cenário.

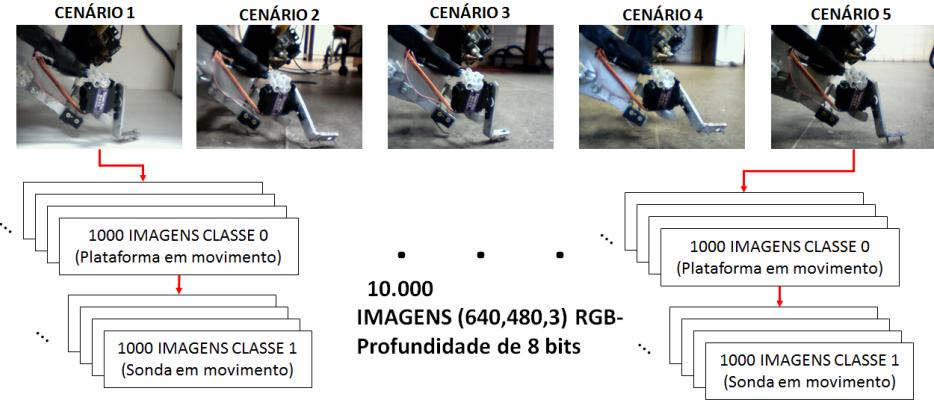


Figura 4: Estrutura do Banco de Dados usado para o treinamento da rede Convolucional.

As imagens do banco de dados possuem dimensão original de 640X480X3 pixels (RGB), que foram reduzidas pelas camadas convolucionais à vetores de características com 256 elementos para entrada na rede tipo *Fully Connected*.

A estrutura final da rede implementada é composta por 5 camadas convolucionais (*Conv*) e 5 camadas de agrupamento (*Pool*). A configuração da MLP se resume à 2 camadas escondidas, somadas à uma camada de entrada e uma camada de saída.

2.4 Modelo de Classificação utilizando Marcadores

O método de extração de características, baseado em limiar de decisão, utiliza a área formada entre os marcadores de frequência fixa no espectro visível. Esses são implementados na forma de Diodos Emissores de Luz (*LEDs*), descritos na Fig. 5a.

O diagrama do sistema mecânico com os marcadores (enumerados como *LED1*, *LED2* e *LED3*-Ponto desejado para o reposicionamento), assim como o método de classificação estão descritos na Fig. 5.

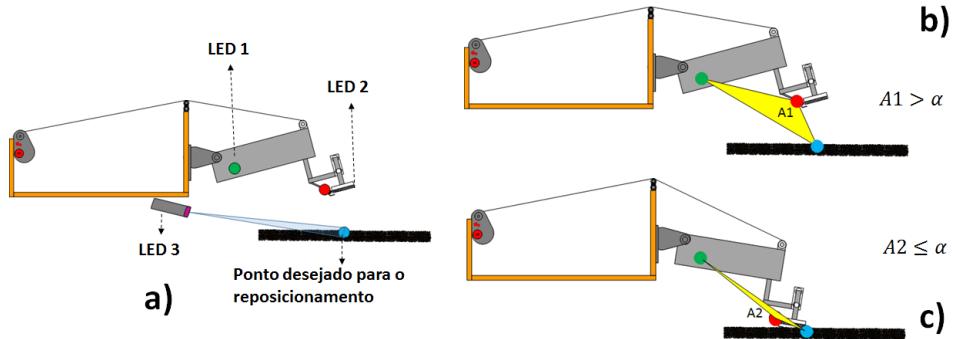


Figura 5: Diagrama dos Marcadores. a) Localização dos Marcadores; b) Área A_1 entre os marcadores que ultrapassa o limiar α (sonda distante do ponto desejado); c) Área A_2 entre os marcadores que não ultrapassa o limiar α (sonda estática no ponto desejado).

Pode ser observada as diferentes áreas entre os marcadores (enumerados como

LED1, *LED2* e *LED3*- Fig. 5a) com a sonda próxima e distante do ponto desejado para o reposicionamento (Fig. 5b e c).

O cálculo da área é dado por:

$$A = \frac{1}{2} \|\det(\mathbf{M})\|, \quad (5)$$

sendo A a área formada entre os marcadores e a matriz \mathbf{M} formada pela média das coordenadas (baricentros) dos pixels de máxima intensidade. A matriz \mathbf{M} é dada por:

$$\mathbf{M} = \begin{pmatrix} \bar{x}_{red} & \bar{y}_{red} & 1 \\ \bar{x}_{green} & \bar{y}_{green} & 1 \\ \bar{x}_{blue} & \bar{y}_{blue} & 1 \end{pmatrix}, \quad (6)$$

sendo \bar{x} , \bar{y} as coordenadas (médias) dos pixels máximo valor, nos respectivos planos *red*, *green*, *blue*.

Baseado na Eq. (5), o critério de classificação nas proximidades do ponto desejado é dado por:

$$\mathbf{C} = \begin{cases} \text{Classe 1,} & A > \alpha \\ \text{Classe 0,} & A \leq \alpha \end{cases} \quad (7)$$

sendo α o limiar de decisão. Quando a área entre os marcadores ultrapassar o limiar α , a sonda deve ser reposicionada para o ponto desejado. Caso contrário, será gerado um comando para finalizar o processo de reposicionamento.

3 Resultados

Foram utilizadas 10.000 imagens sendo 9.000 de treinamento e 1.000 para teste (CNN), comparado à 1.000 imagens de teste para o método dos marcadores (performance descrita na Tabela 1). Os resultados apresentados na Tabela 1, avaliam os classificadores em relação as métricas de acurácia (*performance* geral do classificador), precisão (habilidade do classificador em rotular corretamente as amostras) e sensitividade (habilidade do classificador para encontrar amostras positivas).

Métricas	Marcadores	CNN
Acurácia ($\frac{tp+tn}{tp+tn+fp+fn}$)	51,80%	82%
Precisão (classificação) ($\frac{tp}{tp+fp}$)	55,475%	91%
Sensitividade ($\frac{tp}{tp+fn}$)	52%	64%

Tabela 1: Comparação entre os dois classificadores pelas métricas de Acurácia, Precisão e Sensitividade. Os verdadeiros-positivos correspondem à (tp), os verdadeiros-negativos à (tn), os falso-positivos à (fp) e falso-negativos à (fn).

Os valores de área nas diferentes situações para o método dos marcadores estão descritos na Figura 6:

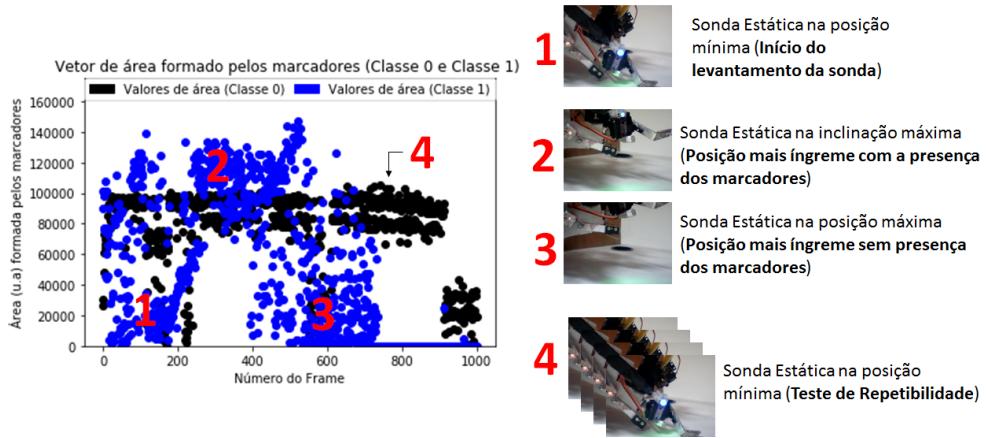


Figura 6: Gráfico dos valores de área agrupados em dois vetores (pontos azuis e pontos pretos), cujos números (em vermelho) indicam: (1) Sonda no início da operação de içamento; (2) Posição cujos marcadores encontram-se nas extremidades da imagem; (3) Sonda fora da área contemplada pela imagem (situação de emergência); (4) Sonda estática na inclinação mínima (verificação da repetibilidade-pontos pretos).

4 Discussões

A CNN mostrou-se superior em todas as métricas utilizadas (Tabela 1), não necessitando do uso de marcadores, sendo testada em um cenário para o qual essa não foi previamente treinada o que justifica seu uso. O desenvolvimento deste sistema de aproximação (Figs. 2 e 3), permite uma serie de aplicações em sistemas mecânicos, diferindo apenas as características a serem extraídas pela CNN (sofrendo novo treinamento e/ou pequenas reconfigurações, caso necessário).

Como desvantagens a CNN necessita de um vasto banco de dados, sendo necessário, para uso em plataformas agrícolas ou ambientes urbanos, ferramentas que automatizem a aquisição e rotulação das imagens (esforço este já realizado por alguns dos autores) (Dan, 2017).

Os valores de inclinação mínima ($\theta_{inclinação} \approx -15^\circ$) e inclinação máxima ($\theta_{inclinação} \approx 30^\circ$) correspondem à inclinação da sonda em relação à plataforma (Fig. 6). Na segunda operação a sonda permanece estática na inclinação mínima (correspondente aos pontos pretos no gráfico).

Pode ser observada instabilidade na extração de características de baixo nível (área entre os marcadores) quando a sonda encontra-se estática na posição mínima (posição de número (4), Fig. 6). Os erros de classificação estão em valores próximos ou maiores ao limiar α fixado em 100000 u.a (pontos pretos, Fig. 6), devendo assumir valores menores à 40000 u.a. (erros atribuídos à vibrações e variações de luminosidade durante a aquisição). Refinos na Eq.(6) e Eq.(7) assim como adição de filtros durante o processamento da imagem (anteriores ao cálculo da área) podem aumentar a acurácia e precisão do classificador.

A posição de número (3) na figura 6 (caracterizada como situação de emergência), obteve valores de área abaixo do limiar α , gerando o comando de parada da sonda (co-

mando necessário neste tipo de situação).

Referências

- Desenvolvimento de uma ferramenta automática para geração de um banco de dados com base em sistema mecânico simulado para treinamento de uma deep neural network-dnn. In: LNCC (editor), *Anais do X Encontro Acadêmico de Modelagem Computacional*, páginas 29–39. 2017.
- Dan C. Ciresan Fang-Lin He Juan P. Rodríguez Flavio Fontana Matthias Faessler Christian Forster Jürgen Schmidhuber Gianni Di Caro Davide Scaramuzza Luca M. Gambardella Alessandro Giusti, Jérôme Guzzi. A machine learning approach to visual perception of forest trails for mobile robots. *IEEE Robotics and Automation Letters*, 1: 661–667, 2014.
- Rafael Woods E. Richard. Gonzalez, C. *Digital Image Processing. Third Edition*. Person, 2007.
- Ian Goodfellow, Yoshua Bengio, e Aaron Courville. *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.
- Mohd. Marzuki Mustafa Kamarul Hawari Ghazali, Mohd Fais Mansor e Aini Hussain. Feature extraction technique using discrete wavelet transform for image classification. *IEEE Research and Development, 2007. SCOReD 2007. 5th Student Conference on*, 2007.
- Antonio Torralba e Aude Oliva. Statistics of natural image categories. *Network: Computation in Neural Systems*, 14(3):391–412, 2003. ISSN 0954-898X.
- Vandyk Amsdem Fernandes Viren Pereira e Junieta Sequeira. Low cost object sorting robotic arm using raspberry pi. *IEEE Global Humanitarian Technology Conference - South Asia Satellite (GHTC-SAS)*, páginas 1–6, 2014.
- Jinshan Jiang Wenzhi Liao. Image feature extraction based on kernel ica. *IEEE Image and Signal Processing, 2008. CISP '08. Congress on*, 2008.
- Jianxin Wu. *Introduction to Convolutional Neural Networks*. National Key Lab for Novel Software Technology, Nanjing University, China, 2017. <https://cs.nju.edu.cn/wujx/paper/CNN.pdf>.
- Binjie Xiao. Principal component analysis for feature extraction of image sequence. *Computer and Communication Technologies in Agriculture Engineering (CCTAE), 2010 International Conference On*, 2010.
- L. Yang Z.C.Song. Wying Hu. Xie, J. A theory of generative convnet. *arXiv:1602.03264v3*, páginas 1–10, 2016.

XI ENCONTRO ACADÊMICO
MODELAGEM COMPUTACIONAL

Estudo de aproximações numéricas na resolução de problemas de Riemann.

Ingrid B. Carneiro¹, Sandra M. C. Malta¹, Márcio R. Borges¹

¹Laboratório Nacional de Computação Científica (LNCC)
Caixa Postal 25.651-075 – Petrópolis, RJ – Brasil

ingridbc@lncc.br, smcm@lncc.br, mrborges@lncc.br

Resumo. A utilização de um esquema de resolução de primeira ou de segunda ordem para a solução de problemas de Riemann pode exibir soluções que apresentam suavização excessiva da frente de onda ou introduzir oscilações espúrias na descontinuidade. Neste trabalho comparamos a eficácia de métodos clássicos de primeira ordem e de segunda ordem, como os métodos Upwind e Lax-Wendroff, com esquemas de alta ordem que possuem propriedade TVD (Total Variation Diminishing), na resolução dos problemas lineares e não-lineares, predominantemente advectivos. As aproximações numéricas são comparadas em termos da norma- L^2 do erro e da precisão em resolver as regiões de salto (descontinuidades).

1 Introdução

Em problemas de transporte predominantemente advectivos, modelos matemáticos são constituídos de uma ou de um conjunto de equações diferenciais hiperbólicas, cuja solução é obtida através da aplicação de um método numérico, tal como o Método dos Volumes Finitos. Entretanto, para problemas de Riemann, esquemas tradicionais de primeira ordem apresentam excessiva difusão numérica, enquanto métodos de mais alta ordem apresentam oscilações espúrias. Desta forma, neste trabalho, objetivamos utilizar e analisar esquemas de alta ordem que possuam propriedade TVD. Para tal, são aplicados limitadores de fluxo, buscando combinar métodos de primeira ordem com aqueles de ordem superior.

2 Aproximações Numéricas

Seja o problema hiperbólico:

$$\begin{cases} s_t + f(s)_x = 0, & (x, t) \in \mathbb{R} \times (0, \infty), \\ s(x, 0) = s^0(x), \\ s(0, t) = s^c, \end{cases} \quad (1)$$

onde f é a função de fluxo convexa, s é a quantidade conservada e $s^c \in \mathbb{R}$ é uma constante. Consideramos s e $f(s)$ funções suficientemente regulares para o estudo desenvolvido

neste trabalho. Sendo $(x, t) \in [0, L] \times [0, T]$ com $L, T \in \mathbb{R}$ positivos, subdividimos L em M volumes finitos, denominados células, com espaçamento constante Δx , isto é,

$$x_{i-1/2} = (i-1)\Delta x \quad , \quad x_{i+1/2} = i\Delta x \quad \text{e} \quad \Delta x = x_{i+1/2} - x_{i-1/2} = \frac{L}{M}, \quad (2)$$

com $i = 1, \dots, M$. Desta forma, o valor médio de $s(x, t)$ na célula i em um tempo fixo $t = t^n = n\Delta t$, $n \in \mathbb{N}$, é definido por:

$$s(x_i, t^n) = s_i^n \equiv \frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} s(x, t^n) dx. \quad (3)$$

Na aproximação da solução do problema (1) usamos métodos numéricos da forma:

$$s_i^{n+1} = s_i^n + \lambda[F_{i-1/2}^n - F_{i+1/2}^n], \quad (4)$$

sendo $\lambda \equiv \frac{\Delta t}{\Delta x}$ e $F_{i+1/2}^n$ o fluxo numérico, ou seja, uma aproximação para o fluxo $f(s)$, ao longo de $x = x_{i+1/2}$. A seguir, apresentamos alguns métodos numéricos usualmente empregados na aproximação de $F_{i+1/2}^n = F(x_{i+1/2}, t^n)$ em (4).

2.1 Métodos de primeira e segunda ordem

Método Upwind (LeVeque, 1992): Erro de truncamento: $\mathcal{O}(\Delta t) + \mathcal{O}(\Delta x)$.

$$F_{i-1/2}^n = f(s_{i-1}^n), \quad \text{se } f' > 0. \quad (5)$$

Método Lax-Wendroff (Cunha, 2003): Erro de truncamento: $\mathcal{O}(\Delta t^2) + \mathcal{O}(\Delta x^2)$.

$$F_{i-1/2}^n = \frac{1}{2}[f(s_{i-1}^n) + f(s_i^n)] - \frac{1}{2}\frac{\Delta t}{\Delta x}f'(s_i^n)[f(s_i^n) - f(s_{i-1}^n)]. \quad (6)$$

Como será exposto na próxima seção, o método Upwind (primeira ordem) produz soluções excessivamente suaves não conseguindo capturar o choque presente na solução analítica. Por outro lado, o método Lax-Wendroff (segunda ordem) apresenta oscilações próximo às descontinuidades. Portanto, utilizam-se métodos de mais alta ordem, combinados com limitadores de fluxo, de modo que estes tenham a propriedade de serem TVD (definido a seguir neste texto) e construídos com base no algoritmo REA (reconstruir, evoluir e projetar), descrito abaixo (Le Veque, 2002; Toro, 2013):

- **Reconstruct (reconstrução):** A solução aproximada \bar{s}_i^n (constante por partes) é reconstruída por uma função polinomial por partes $\tilde{s}^n(x)$, em cada célula $I_i := (x_{i-1/2}, x_{i+1/2})$ no tempo t^n , ou seja:

$$\tilde{s}^n(x) = \bar{s}_i^n + \phi_i^n(x - x_i) \quad \text{para } x_{i-1/2} \leq x \leq x_{i+1/2}, \quad (7)$$

onde

$$x_i = \frac{1}{2}(x_{i-1/2} + x_{i+1/2}) = x_{i-1/2} + \frac{1}{2}\Delta x, \quad (8)$$

e $\phi_i^n(\cdot)$ é um *limitador de fluxo*, que será definido na sequência desse texto.

- **Evolve (evolução):** Evolui a equação de maneira exata ou aproximada com os dados iniciais do passo de Reconstrução para obter $\tilde{s}(x, t^{n+1})$.
- **Average (cálculo da média):** Calcula-se a média desta função sobre cada célula para obter as novas células médias no tempo t^{n+1} .

$$\bar{s}_j^{n+1} = \frac{1}{\Delta x} \int_{I_j} \tilde{s}(x, t^{n+1}). \quad (9)$$

2.2 Métodos com limitadores de fluxo

Os esquemas com limitadores de fluxo são métodos numéricos baseados na reconstrução linear por partes, ou seja, possuem a acurácia dos métodos de segunda ordem para soluções suaves e satisfazem a propriedade TVD (LeVeque, 2002), isto é:

$$TV(s^{n+1}) \leq TV(s^n), \quad \text{onde} \quad TV(s) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_{-\infty}^{\infty} |s(x) - s(x - \varepsilon)| dx. \quad (10)$$

Assim, a expressão para $F_{i+1/2}^n$ em (4) torna-se:

$$F_{i+1/2}^n = F_L^n(s; i + 1/2) + \phi(s; i + 1/2)[F_H^n(s; i + 1/2) - F_L^n(s; i + 1/2)], \quad (11)$$

onde $F_L^n(s; i + 1/2)$ é um fluxo de primeira ordem, $F_H^n(s; i + 1/2)$ fluxo de segunda ordem e $\phi(s; i + 1/2)$ é o limitador de fluxo, que será definido a seguir.

Uma das formas de verificar se $s(x, t^n)$ é suave consiste em utilizar o quociente entre os valores da derivada em pontos consecutivos acrescido de uma tolerância ε no denominador, evitando-se assim divisão por zero (Cunha, 2003). Logo, define-se

$$\theta_{i-1/2}^n = \frac{\Delta s_{i-3/2}^n}{\Delta s_{i-1/2}^n + \varepsilon} \quad \text{e} \quad \Delta s_{i-1/2}^n = s_i^n - s_{i-1}^n, \quad (12)$$

de onde observamos que os valores de s_i^n são suaves próximos de $x_{i-1/2}$ quando $\theta_{i-1/2} \approx 1$. Podemos assim reescrever $\phi(s; j)$ em (11) como uma função de θ . Logo, se $\phi(\theta) = 1$ tem-se um método de segunda ordem e se $\phi(\theta) = 0$ um método de primeira ordem. As funções limitadoras de fluxo frequentemente encontradas na literatura são:

- **Minmod (MM):** Dada a partir de dois argumentos,

$$MM(a, b) = \begin{cases} a, & \text{se } |a| < |b| \text{ e } ab > 0, \\ b, & \text{se } |a| > |b| \text{ e } ab > 0, \\ 0, & \text{se } ab \leq 0. \end{cases} \quad \phi(\theta) = MM(1, \theta), \quad (13)$$

- **Superbee (Sb)**

$$\phi(\theta) = \max(0, \min(1, 2\theta), \min(2, \theta)), \quad (14)$$

- **Monotonized Central (MC)**

$$\phi(\theta) = \max(0, \min((1 + \theta)/2, 2, 2\theta)), \quad (15)$$

- **van Leer (vL)**

$$\phi(\theta) = \frac{\theta + |\theta|}{1 + |\theta|}. \quad (16)$$

3 Resultados Numéricos

Nesta seção, apresentamos simulações numéricas objetivando analisar a precisão dos esquemas numéricos, introduzidos anteriormente, para a solução de problemas de Riemann, com condição inicial da forma $s(x, 0) = s_l$, se $x \leq 0$, e $= s_r$, se $x > 0$, onde s_l e s_r são estados constantes à esquerda e à direita respectivamente. Foram realizados dois experimentos (Exemplos 1 e 2) com os seguinte dados: malha de $M = 500$ células; domínios com comprimentos $L_1 = 1,5\text{ m}$ e $L_2 = 1,0\text{ m}$; tempo total de simulação $T = 0,2\text{ s}$; condições de contorno $s(0, t) = s^c = 0,0$ e $NC_1^{CFL} = 0,1$, $NC_2^{CFL} = 0,8$. Aqui NC_1^{CFL} e NC_2^{CFL} são os números de Courant definidos em função da condição de CFL (LeVeque, 2002), como:

$$NC^{CFL} = \max_{s \in [s^{\min}, s^{\max}], x \in [0, L]} \left| \frac{df}{ds} \right| \frac{\Delta t}{\Delta x} \leq 1,$$

onde s^{\min} e s^{\max} são as quantidades conservadas mínima e máxima para um dado t^n .

Exemplo 1: O problema de advecção linear (Toro, 2013) é definido como sendo:

$$\begin{cases} s_t + s_x = 0, & t > 0 \\ s(x, 0) = s_0(x), \end{cases} \quad (17)$$

com condição inicial,

$$s(x, 0) = \begin{cases} 0, & x \leq 0,3 \\ 1, & 0,3 \leq x \leq 0,7 \\ 0, & x \geq 0,7 \end{cases} \quad (18)$$

Este problema possui saltos em $x = 0,3$ e em $x = 0,7$; portanto a solução analítica se dá através do deslocamento da solução inicial. Como pela condição de Rankine-Hugoniot (Trangenstein, 2009), tem-se $s_{choque} = \frac{f(s_l) - f(s_r)}{s_l - s_r} = 1$, então as descontinuidades estão em $x = 0,3 + 0,2 \times 1 = 0,5$ e $x = 0,7 + 0,2 \times 1 = 0,9$.

Exemplo 2: O problema de Burgers é dado por:

$$\begin{cases} s_t + \left(\frac{s^2}{2} \right)_x = 0, & t > 0 \\ s(x, 0) = s_0(x), \end{cases} \quad (19)$$

com condição inicial,

$$s(x, 0) = \begin{cases} 1, & 0,1 < x < 0,3 \\ 0, & \text{caso contrário} \end{cases} \quad (20)$$

Este problema apresenta a propagação de uma onda de rarefação em $x = 0,1$ e de uma onda de choque em $x = 0,3$. Portanto, a velocidade de choque é dada por $s_{choque} = \frac{f(s_l) - f(s_r)}{s_l - s_r} = 1/2$ e a velocidade de rarefação $s_{r,f} = f'(s_r) = s = 1$.

As soluções exata e numérica dos exemplos são mostradas nas Figuras 1 a 5. Observando os resultados plotados nas Figuras 1 e 4, nota-se que as aproximações do método

Upwind apresentam forte suavização próximo às descontinuidades devido aos erros de aproximação, como mencionado em (Cunha, 2003; LeVeque, 2002). Já o método de Lax-Wendroff (LW), por ser de segunda ordem, exibe menor difusão numérica quando comparado ao método de primeira ordem Upwind; porém são observadas oscilações espúrias próximas às regiões de salto, como visto na Figura 2. Portanto, este método não será considerado nos estudos seguintes. Ao introduzirmos os limitadores de fluxo, o método de Lax-Wendroff apresenta resultados satisfatórios para ambos os exemplos, conforme as Figuras 3 e 5.

Na Tabela 1 comparamos a precisão dos métodos Upwind e Lax-Wendroff com limitador de fluxo MinMod (LW-MM), Superbee (LW-Sb), MC (LW-MC) e vanLeer (LW-vL) a fim de averiguar a eficiência da propriedade TVD via o erro na norma- L^2 para os Exemplos 1 e 2. Dos resultados concluímos que, em geral, o método LW-Sb apresenta as melhores aproximações, com NC_1^{CFL} e NC_2^{CFL} , para os dois exemplos. No entanto, nota-se que o método LW-MC é um pouco mais preciso para o Exemplo 1, com NC_2^{CFL} .

Tabela 1: Norma L^2 do erro para os diferentes métodos nos Exemplos 1 e 2.

Método	erro L^2 - ex. 1	erro L^2 - ex. 2	Método	erro L^2 - ex. 1	erro L^2 - ex. 2
Upwind	0,16665	0,05117	Upwind	0,12055	0,04777
LW-MM	0,10783	0,03713	LW-MM	0,10343	0,04635
LW-Sb	0,07948	0,03409	LW-Sb	0,10409	0,04623
LW-MC	0,09051	0,03529	LW-MC	0,10112	0,04639
LW-vL	0,09290	0,03581	LW-vL	0,10250	0,04646

(a) Norma L^2 de erros para NC_1^{CFL} . (b) Norma L^2 de erros para NC_2^{CFL} .

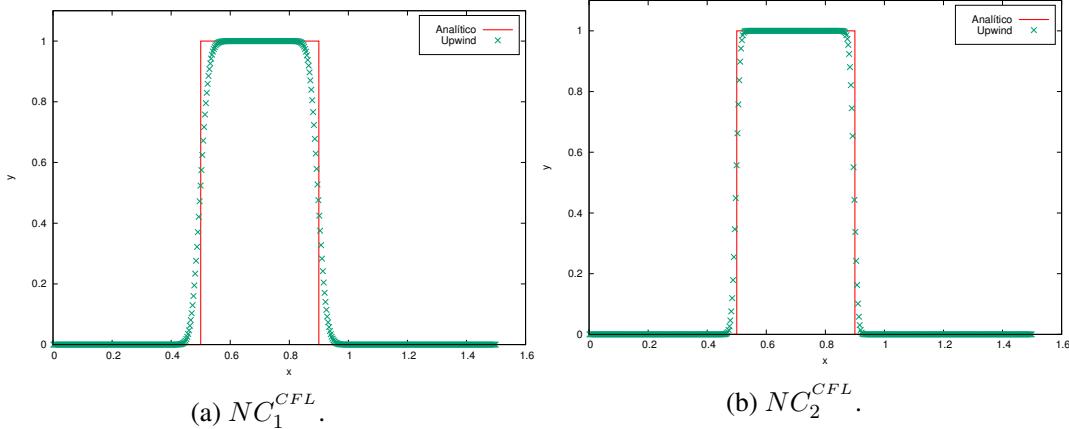


Figura 1: Método Upwind para o Exemplo 1.

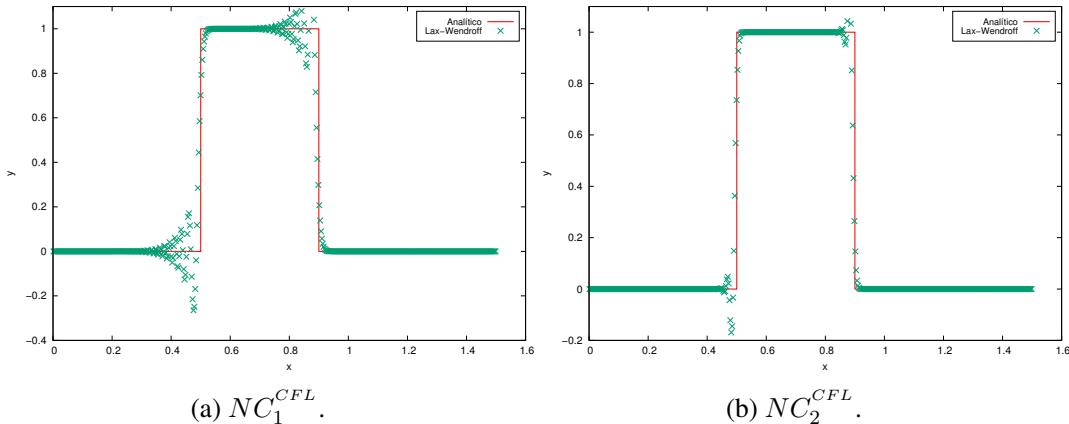


Figura 2: Método Lax - Wendroff (LW) para o Exemplo 1.

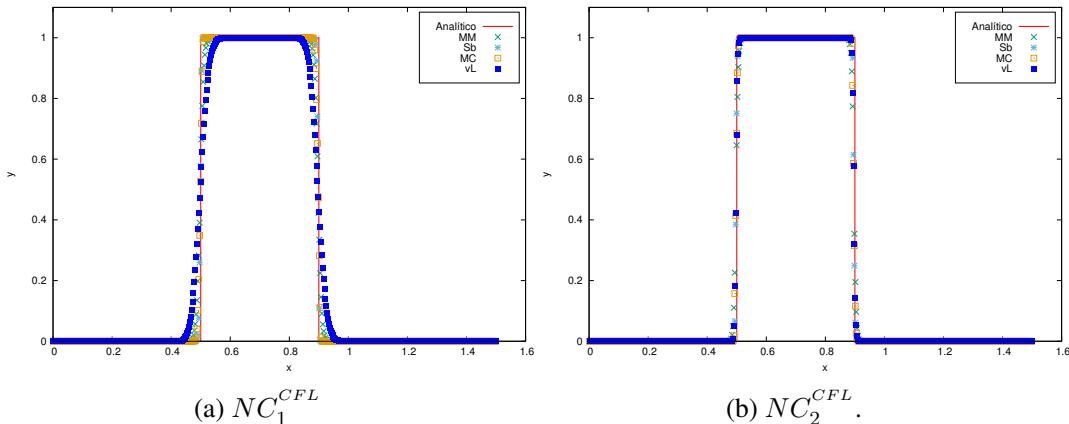


Figura 3: Método LW com limitador de fluxo para o Exemplo 1.

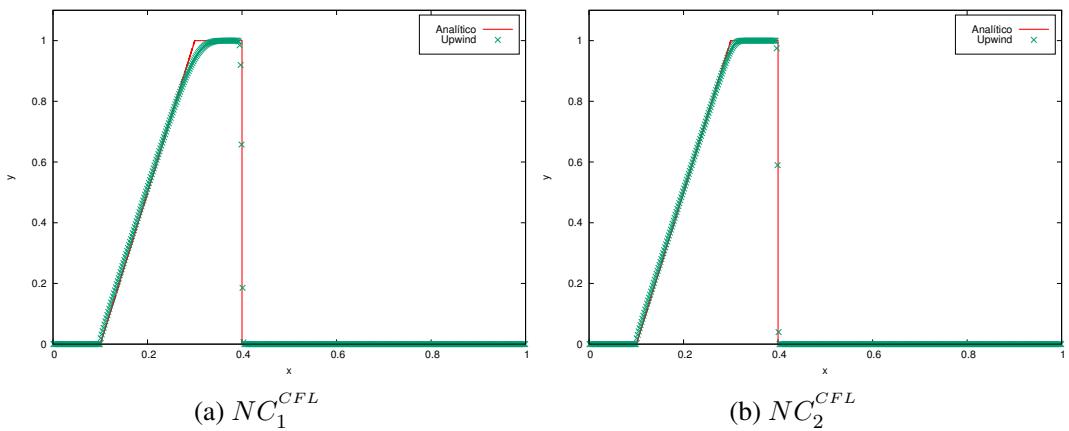


Figura 4: Método Upwind para o Exemplo 2.

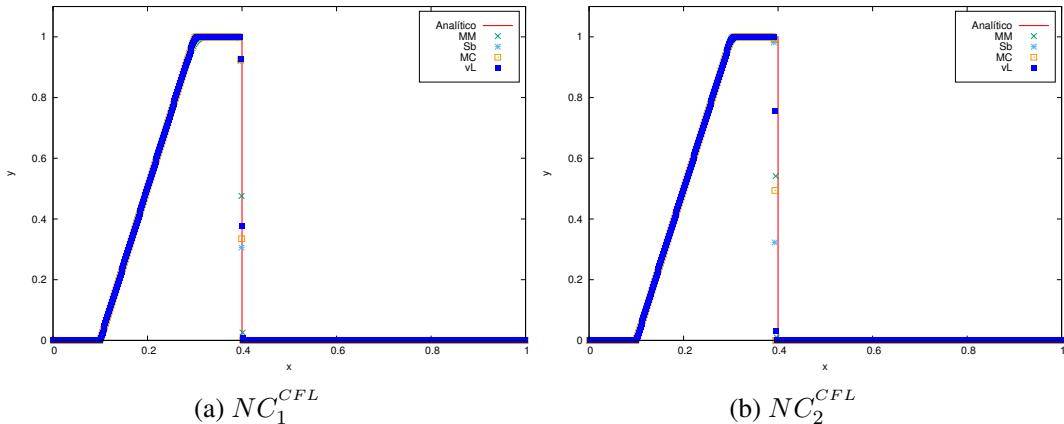


Figura 5: Método LW com limitador de fluxo para o Exemplo 2.

4 Conclusões

Neste trabalho, foram aplicados os métodos de primeira e de segunda ordem para a resolução dos problemas de Riemann, além dos métodos de alta ordem que possuem a propriedade TVD. Analisando os resultados numéricos, observamos que os métodos de segunda ordem produzem oscilações espúrias na presença de problemas com descontinuidades, enquanto que os métodos de primeira ordem resolvem a descontinuidade, mas apresentam suavizações devido ao erro de discretização (Cunha, 2003; LeVeque, 2002). Portanto, os métodos com limitadores de fluxo exibem melhores resultados para problemas com descontinuidades, mesmo para valores de NC^{CFL} baixos. Dentre os métodos de alta ordem, destacam-se os métodos LW-Sb e LW-MC, que descrevem melhor o comportamento das descontinuidades nos casos analisados neste estudo.

Referências

- M Cristina C Cunha. *Métodos numéricos*. Editora da UNICAMP, 2003.

Randall J LeVeque. *Finite volume methods for hyperbolic problems*, volume 31. Cambridge university press, 2002.

R.J. LeVeque. *Numerical Methods for Conservation Laws*. Birkhäuser Basel, 1992.

Eleuterio F Toro. *Riemann solvers and numerical methods for fluid dynamics: a practical introduction*. Springer Science & Business Media, 2013.

John A Trangenstein. *Numerical solution of hyperbolic partial differential equations*. Cambridge University Press, 2009.

XI ENCONTRO ACADÊMICO
MODELAGEM COMPUTACIONAL

Scientific computing in Statistical Mechanics: time decay of orientation order in 2D hard disk system

A.G. França¹, R.S. Grisotto² B.M. Rocha³, F.A.A.M.N. Soares³, H.A. Fernandes¹, P.F. Gomes¹

¹Instituto de Ciéncia Exatas, Universidade Federal de Goiás
Jataí, GO – Brasil

²Instituto de Computação – Universidade Estadual de Campinas
Campinas, SP – Brasil

³Instituto de Informática – Universidade Federal de Goiás
Goiânia, GO – Brasil

andreygfranca@gmail.com, paulofisicajatai@gmail.com

Abstract. *In this work we successfully applied diverse computational techniques to calculate important quantities in one problem in Statistical Mechanics: the hard disk system in two dimensions. We calculated the global and local orientation order and the time decay constant of the global orientation correlation function. We also computed the Voronoi construction to visualize the spatial distribution of the local orientation order.*

Resumo. *Neste trabalho apresentamos diversas técnicas computacionais no cálculo de grandezas de interesse em um sistema de Mecânica Estatística: discos rígidos em duas dimensões. Calculamos a ordem de orientação local e global e a constante de decaimento da função de correlação da orientação global. Fizemos também a diagramação de Voronoi para visualização da distribuição espacial da orientação local.*

1 Introduction

Solidification and vaporization of water are two common examples of phase transitions in physics. The temperature and pressure defines the phase of the water: solid, liquid or gas. Another parameter called order parameter is used to verify which phase the system is. In the case of the water it is the molecule organization: in the solid phase the molecules form an ordered hexagonal state, in the liquid state the molecules are close to each other but with no organization and in the gas state there is no correlation between the molecules. One interesting system to study phase transitions is the hard disks in two dimensions [Isobe and Krauth (2015)], composed of N disks of radius R confined in a 2D planar rectangular box of sides L_x and L_y (see Fig. 1(a)). Each disk k has its center at position (x_k, y_k) . The superposition between two disks is forbidden so the minimal distance between them is $2R$. One parameter is the local orientation of each disk k defined

as [Deutschlander et al. (2014); Engel et al. (2013)]:

$$\psi_k(x_k, y_k) = \frac{1}{6} \sum_{l=1}^6 e^{6i\theta_{kl}}, \quad (1)$$

where $i = \sqrt{-1}$ and the sum is taken over the six nearest-neighbors. The angle θ_{kl} is between a line passing through the center of the two discs and the x axis (see Fig. 1(b)). The ψ_k is defined so that $\psi_k = 0$ in the liquid phase and $\psi_k = 1$ in the solid one. The spatial distribution of the parameter through the volume is also important. The order parameter can have different ranges defined by its dependency on the distance r [Deutschlander et al. (2014)]: exponential decay $e^{-r/\xi}$ means short range, algebraic one $r^{-\xi}$ means quasi-long range and a constant variation means a long range. So the liquid state in the 2D hard disk system has short range and the solid state has long range for ψ_k . However, a third phase was theoretical and experimentally observed: the so called hexatic phase [Strandburg (1988)] (see Fig. 2(a)) and it has quasi-long range orientational order [Bernard and Krauth (2011)]. The hexatic phase is an exclusive feature of 2D systems [Isobe and Krauth (2015)].

1.1 Global orientation order

The global orientation is the absolute value of the average of the local one:

$$\Psi = \left| \frac{1}{N} \sum_{k=1}^N \psi_k \right|. \quad (2)$$

The phase of the system depends on the density $\eta = (N\pi R^2)/(L_x L_y)$. For $\eta < 0.700$ the system is liquid (L), for $0.700 < \eta < 0.716$ it is in the coexistence between the liquid and hexatic (L-H) phases (see Fig. 2(a)). For $0.716 < \eta \lesssim 0.720$ the phase is hexatic (H). For $\eta \gtrsim 0.72$ the phase is solid (S), which has an hexagonal symmetry (see Fig. 4(a)). The liquid phase has random positions for the disks (see Fig. 4(b)). We study the time behavior of the orientation order correlation function:

$$C(\delta) = \sum_{t=0}^Q \Phi(t)\Phi(t+\delta), \quad (3)$$

which should be normalized so $C(0) = 1$. The normalized orientation order is $\Phi(t) = \Psi(t) - \Psi_\infty$ where $\Psi_\infty = \lim_{t \rightarrow \infty} \Psi(t)$ ¹. The correlation function has an exponential decay [Krauth and Wilson (2009)] of the form $C(\delta) \approx e^{-t/\tau}$ where the decay time τ is a function of the number of disks, $\tau = \tau(N)$. This function depends of the phase:

$$\tau(N) = \begin{cases} N \log N, & \text{liquid phase} \\ cN^\alpha, & \text{hexatic phase} \\ N, & \text{solid phase} \end{cases} \quad (4)$$

¹In practical terms, $\Psi_\infty = \Psi(t = Q)$ where Q is the number of evaluated Monte Carlo iterations.

In this work we study the time decay τ of the global orientation in the liquid-hexatic transition as function of N in the 2D hard disk system. We use different computation technique to evaluate the time evolution and the orientation order. The goal is to verify the exponential decay of the orientation order correlation function and then obtain the behavior of the time constant as function of the number of disks N . To our knowledge, and complete study of the τ as function of N in the hexatic-liquid coexistence phase ($0.70 < \eta < 0.716$) has not been done.

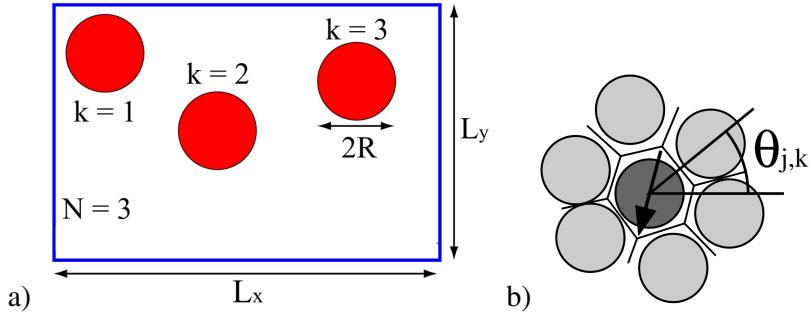


Figure 1: (a) Illustration of the 2D hard disk system. (b) Illustration of the angle θ_{jk} between two discs. Source: Ref. [Bernard (2011)].

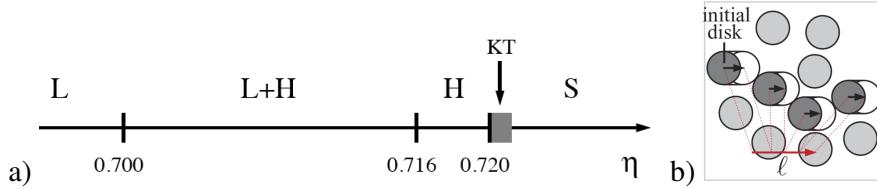


Figure 2: (a) Identification of the phase as function of the density η . (b) Illustration of the movement in the Event-chain algorithm. Collective move with the total displacement ℓ . Source for both figures: Ref. [Bernard (2011)].

1.2 Simulation Method

We used two methods for the time evolution of the system. The first one is the Event Chain Monte Carlo method [Michel et al. (2014)] where many disks move in sequence in each iteration. Each disk moves until it reaches another disk (when their distance = $2R$). Then the first disk stops and the second one starts to move until it reaches the third disk. The process repeats until the sum of the distances of all disks are equal to an input value ℓ [Krauth and Wilson (2009)]. The method is efficient because many disks move in one iteration. The other method is the parallel Markov Chain Monte Carlo one, calculated with the Hard Particle Monte Carlo module (HPMC) [Anderson et al. (2016)] of HOOMD-BLUE package² [Anderson et al. (2008); Glaser et al. (2015)]. This implementation has been used in many type of problems including disks [Anderson et al. (2017)].

²glotzerlab.engin.umich.edu/hoomd-blue.

To calculate ψ_k we need the distances between the disk k and all other disks. Using brute force one requires $N - 1$ calculations. Then, to calculate the average Ψ , another N calculations will be done with total of $(N - 1)N \simeq N^2$ calculations when N is large. The distances are also required in the time evolution of the system. We implemented this algorithm in C language. However, for large N we used the method List of Cells [Frenkel and Smit (2002)] which is faster. In this one a grid of N cells is created so there is one disk per cell. When we need to get the neighbor of a disk, we just look into the neighbor cells. The scaling of this method is N . We implemented this list of cells method in C++.

2 Results

The pseudo code of our C++ implementation is illustrated in the Algorithm 1. The distance calculations are performed in the functions `newL` and `abs_Psi`, where the `listcell` structure is used to identify the neighbor disk. The matrix `LxLy` keeps the positions $x(t), y(t)$ of all disks in the iteration t^3 . In the Fig. 3(a) there is a simple example of the list with 9 cells. To get the neighbor of a disk, just check the neighbor cells (each cell has pointer pointing to them and a linked list to keep the id of the disks). The Fig. 3(b) shows a comparison of execution time between brute force and list of cells methods. For small N , brute force is better because there is a cost to create the cell grid. However, for $N = 128^2$ the list of cells is already one order of magnitude better, and this improvement increases with N .

Algorithm 1 Time evolution with Event Chain Monte Carlo

```

1: procedure MCCE
2:   Input:  $N, \ell, \eta, S, Q, L_x$ 
3:   listcell  $\leftarrow$  CREATE_LIST( $N$ )
4:   for  $k \in S$  do
5:      $LxLy \leftarrow$  INITIAL( $N, \eta, L_x$ )
6:     ADD_DISKS( $N$ , listcell)
7:     seed  $\leftarrow k$ 
8:     for  $t \in Q$  do
9:        $LxLy \leftarrow$  NEWL( $LxLy, N, \eta, \ell, seed, listcell$ )
10:       $Psi(t) \leftarrow$  ABS_PSI( $\eta, LxLy, listcell$ )
11:      file  $\leftarrow$  time_Psi.csv
12:      write( $t, Psi$ , file)

```

In total, the code has six input parameters: (1) number of disks N , (2) the density η , (3) the number of iterations Q , (4) the side L_x of the box, (5) the number S of samples in order to take the average (and decrease the error) and (6) the total distance ℓ . We used $S = 5$, $L_x = 100$ and $\ell = \sqrt{N}$ and $Q = 30\,000$ in all simulations. All the graphics presented here were done in Python (Matplotlib package). The way to achieve the liquid state configuration shown in Fig. 4(b) was to evaluate the system with $Q = 200$ iterations and $\eta = 0.5$ using our C++ implementation and an initial hexagonal configuration (Fig. 4(a)).

³Although the concepts of the method and the definition of Ψ are not too complicated, the code itself presents many numerical challenges [Bernard (2011)].

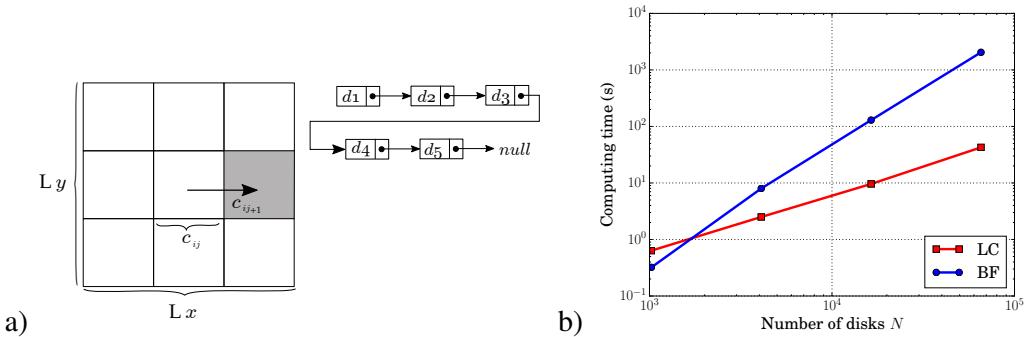


Figure 3: (a) System with 9 cells, each one has pointer pointing to the neighbors and a linked list. (b) Comparison between the two methods to calculate the distances between the disks: brute force in blue and list of cells in red.

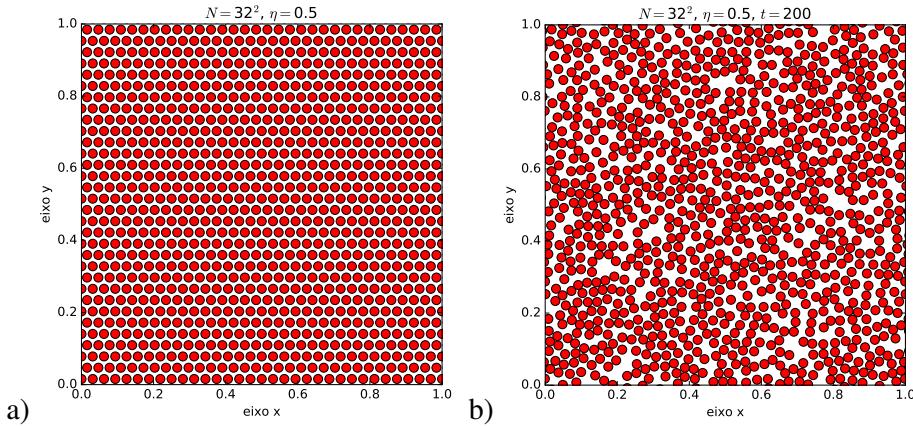


Figure 4: Simulation with $N = 32^2$ and $\eta = 0.5$ using our C++ implementation. (a) Initial configuration with triangular symmetry. (b) Configuration with $t = 200$.

The correlation functions $C(\delta)$ were calculated for $N = 64^2, 128^2, 256^2, 512^2$ and 1024^2 . The results shown in the Fig. 5(a) were noisy. To find the decay constant it is needed more samples. Another result is shown in the Fig. 5(b), which is a kind of graphic used to visually identify the phases [Bernard and Krauth (2011); Engel et al. (2013)]. First a Voronoi construction is created where each cell contains only one disk. Then, the local orientation order ψ_k is calculated for each disk and a color coding is created. The disk in each cell are removed and each cell are painted with the color of its disk orientation order [Bernard (2011)]. If ψ_k is aligned with Ψ we have $\cos \theta_k = 1$ and if they are perpendicular $\cos \theta_k = 0$, using the associated vectors with ψ_k and Ψ . Note that in Fig. 5(b) there are $N = 128^2$ disks. The number of sides of each cell is the number of its neighbors cells. As the solid phase of the hard disk system has hexagonal symmetry, the average number of neighbors cell is 6, which explains why most of the cells are hexagons. As the density is $\eta = 0.7$, the configuration of Fig. 5(b) is liquid.

The functions of HOOMD-BLUE package are organized in modules, which should be installed as regular Python packages (within Conda environment). The results

for $C(\delta)$ are shown in the Fig. 5(c) for $\eta = 0.705$ and $N = 64^2, 128^2, 256^2$ and 512^2 . To calculate the decay time τ we adjust an exponential function $f(\delta) = A_0 + A_1 e^{-\delta/\tau}$ to the data. The blue circles in Fig. 5(d) are the obtained values of τ . The second step is to determine α from Eq. 4. So we take the natural logarithm $\ln \tau = B + \alpha \ln N$ (with $B = \ln c$) and make a linear fit for $\ln \tau$ vs. $\ln N$. The slope will be the constant α . Indeed, they follow a linear behavior as the red solid line indicates in Fig. 5(d)⁴.

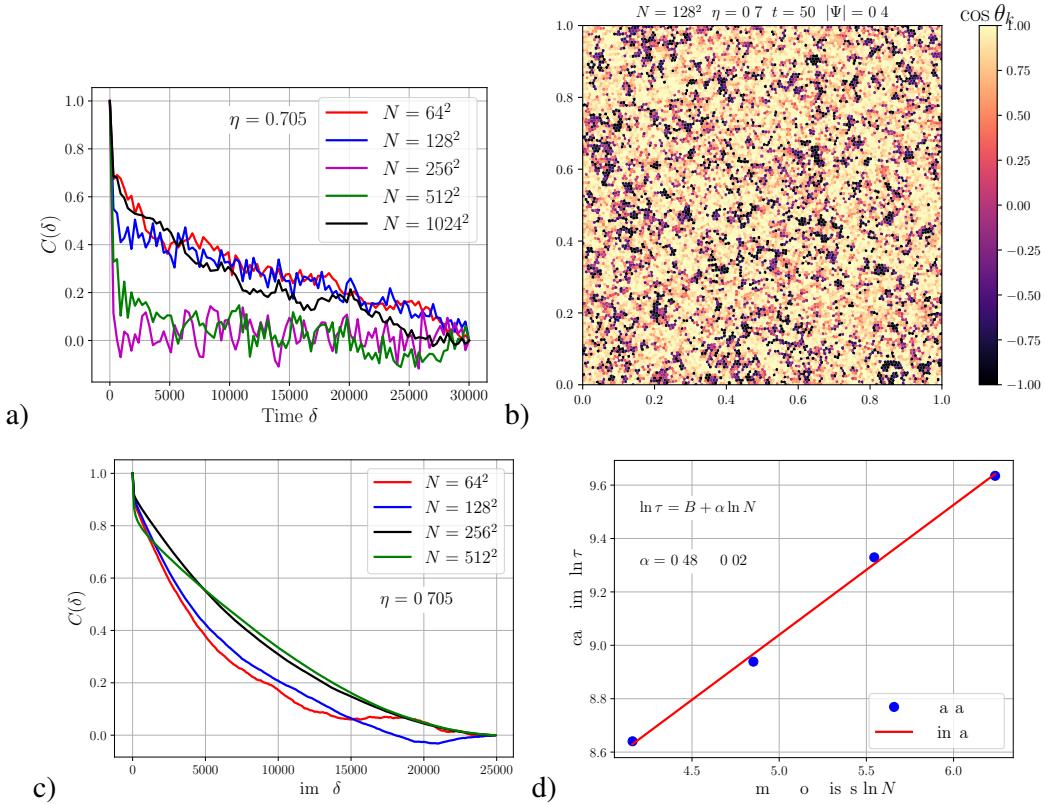


Figure 5: Time behavior of correlation function $C(\delta)$ (Eq. 3). (a) Using our C++ implementation. (b) Voronoi construction where each cell contains only one disk. (c) Using HOOMD. (d) Linear fit of the decay τ vs. the number of disks N using the results from HOOMD case.

3 Conclusions

We compute the correlation function of the orientation order in the 2D hard disk system and verified its exponential behavior. We then obtain the expected behavior of the decay constant τ as function of the number of disks N for $\eta = 0.705$. We successfully applied various computation techniques to this Statistical mechanical problem, showing the great benefit which one can have from the interplay between Statistical Mechanics and computational modeling.

⁴The code that generated the results presented here are available at Github: github.com/AndreyGFranca/mcec.

4 Acknowledgements

This work was funded in part by CNPq (A. G. Fran  a and R.S. Grisotto). We thank W. Krauth for the suggestions.

References

- J. A. Anderson, C. D. Lorenz, and A. Travesset. *Journal of Computational Physics*, 227: 5342, 2008.
- J. A. Anderson, M. E. Irrgang, and S. C. Glotzer. *Computer Physics Communications*, 204:21–30, 2016. doi: <http://dx.doi.org/10.1016/j.cpc.2016.02.024>.
- J. A. Anderson, M. E. Irrgang, and S. C. Glotzer. *Physical Review X*, 7:021001, 2017. doi: <https://doi.org/10.1103/PhysRevX.7.021001>.
- E. Bernard. PhD thesis, Universit   Pierre et Marie Curie - Paris VI, The address of the publisher, 2011. <https://tel.archives-ouvertes.fr/tel-00637330/en>.
- E. P. Bernard and W. Krauth. *Physical Review Letters*, 107:155704, 2011. doi: <http://dx.doi.org/10.1103/PhysRevLett.107.155704>.
- S. Deutschlander, A. M. Puertas, G. Maret, and P. Keim. *Physical Review Letters*, 113: 127801, 2014. doi: <http://dx.doi.org/10.1103/PhysRevLett.113.127801>.
- M. Engel, J. A. Anderson, S. C. Glotzer, M. Isobe, E. P. Bernard, and W. Krauth. *Physical Review E*, 87:042134, 2013. doi: <http://dx.doi.org/10.1103/PhysRevE.87.042134>.
- D. Frenkel and B. Smit. *Understanding Molecular Simulation*. Academic Press, 2002.
- J. Glaser, T. Dac Nguyen, J. A. Anderson, P. Lui, F. Spiga, J. A. Millan, D. C. Morse, , and S. C. Glotzer. *Computer Physics Communications*, 192:97, 2015. doi: <http://dx.doi.org/10.1016/j.cpc.2016.02.024>.
- M. Isobe and W. Krauth. *The Journal of Chemical Physics*, 143:084509, 2015. doi: <http://dx.doi.org/10.1063/1.4929529>.
- E. P. Bernard W. Krauth and D. B. Wilson. *Physical Review Letters*, 80:056704, 2009. doi: <http://dx.doi.org/10.1103/PhysRevE.80.056704>.
- M. Michel, S. C. Kapfer, and W. Krauth. *The Journal of Chemical Physics*, 140:054116, 2014. doi: <http://dx.doi.org/10.1063/1.4863991>.
- K. J. Strandburg. *Review of Modern Physics*, 60:161, 1988. doi: <http://dx.doi.org/10.1103/RevModPhys.60.161>.

On enhanced reduced models for advection-diffusion problems

L.A. Mansilla Alvarez^{1,2}, P.J. Blanco^{1,2}, R.A. Feijóo^{1,2}

¹National Laboratory for Scientific Computing, LNCC/MCTIC
Av. Getúlio Vargas 333, 26561-075, Petrópolis, Brazil.

²National Institute of Science and Technology in Medicine Assisted
by Scientific Computing, INCT-MACC, Petrópolis, Brazil.

lalvarez@lncc.br, pjblanco@lncc.br, feij@lncc.br

Abstract. Surrogate models based on the definition of one main direction, in which the dynamics is more valuable, appears in many engineering areas due their capacity to provide accurate solutions while reducing the computational burden in large-scale simulations. In this work we explore a family of reduced-order models which are defined through the combination of classical finite element approximants to approach the dynamics in the dominant (main) direction and the tackle of transversal dynamics via high-order polynomials. This approach proves to be capable to accurately predict the global dynamics while reducing drastically the problem size and, therefore, the computational time.

1 Introduction

The trade-off between computational cost and the accurate numerical approximation (through, for example, by element methods) is well known and constitutes an intensive field of research in several areas of engineering. Problems that exhibit a spatial dimension predominant over the others, like problems defined on pipe-like domains, have been focus on many efforts in the search for efficient numerical solutions (Blanco et al., 2009; Hughes and Lubliner, 1973; Osiadacz, 1987). An interesting, but limited, approach is based on reducing the problem to a one dimensional setting along the mainstream after dropping the transversal dynamics. Examples of this are the popular Euler equations in fluid dynamics that reduce drastically the problem size but focusing only on the mainstream dynamics.

In a different direction, strategies in where the full kinematics is considered but where the approximate solution lives in a finite-dimensional space specially designed for each problem, appealing to *a priori* knowledge or previous solutions, appears as an alternative way to reduce the computational burden. Some strategies of this type are the well known Proper Orthogonal Decomposition (Berkooz et al., 1993; Kunisch and Volkwein, 2001), Reduced Basis (Drohmann et al., 2011; Peterson, 1989) and the recently proposed Hierarchical Modeling (Aletti et al., 2015; Perotto, 2014) which share, in some way, a reduction in the computational cost but in detriment of a high computational effort in a *offline* step, where is defined the finite-dimensional space in where the continuous problem is approximate.

In this work, is studied a numerical methodology capable to be straightforwardly extended to 3D problems, that reduces the computational burden without significantly loss of transversal information and general enough to be independent of any *offline* step. Moreover, the strategy features a computational implementation based in the use of rather standard concepts and numerical tools. This approach, based in (Blanco et al., 2015; Mansilla Alvarez et al., 2017) and coined as Transversally Enriched Pipe Element Method (TEPEM), is based on the split of transversal and dominant dependence of the physical field of interest. In this regard, low order interpolants are employed for the dominant (axial) direction, while high-order interpolants are considered for the transversal direction. The descriptive capabilities of high-order interpolants to accurately approximate the transversal dynamics in this kind of problems allows us to consider the TEPEM as an efficient methodology placed in between a general purpose finite element method and specially oriented reduced-order approaches.

2 Enriched reduced-model scheme

In this section will be introduced the basic ingredients of the TEPEM. The structure of this approach is based in the discretization of the computational domain in pipe regions, aligned with the dominant direction of the dynamics, to exploit this characteristic by differentiating the way in which physical fields are approximated according to the spatial direction.

2.1 Geometric discretization

The main geometrical assumption at the TEPEM scope is the capability to perform a pipe-type partition of the computational domain. Such partition is composed by pipe-elements (parallelograms in the 2D case) which split the axial and transversal directions of the dynamics. More precisely, let us consider a domain $\Omega \subset \mathbb{R}^2$ as show in Figure 1. A key aspect of the TEPEM is the special discretization based on pipe-type elements, in which each element has its axial direction given by the axial direction of Ω (dominant direction). We divide the domain Ω into a partition $\mathcal{T}_h(\Omega)$ (parametrized by the axial mesh parameter h associated to the axial length of the elements).

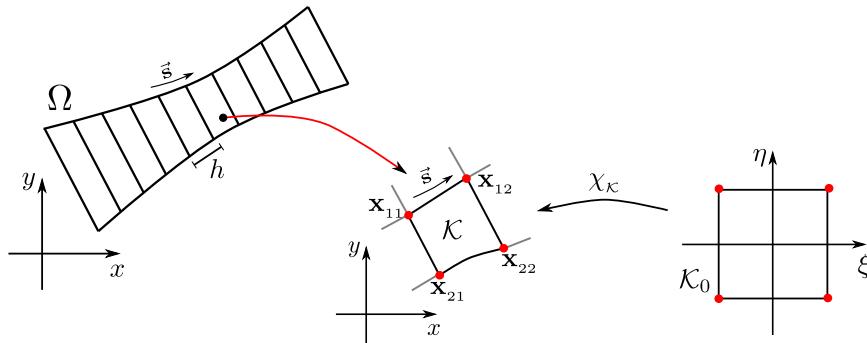


Figure 1: Schematic view of pipe partition in a two-dimensional geometrical domain. Each pipe element, denoted by \mathcal{K} , is related with the reference element in the $\xi\eta$ -plane.

The nature of these pipe elements allows us to define a sub-parametric mapping between an arbitrary element \mathcal{K} and the reference one $\mathcal{K}_0 = [-1, 1]^2$ (in the space $\xi\eta$) though the mapping

$$\chi_{\mathcal{K}}(\xi, \eta) = \sum_{i,j=1}^2 \mathbf{x}_{ij} L_i(\xi) L_j(\eta) \quad (1)$$

where $\{\mathbf{x}_{ij}\}$ ($i, j = 1, 2$) are the coordinates of the points that define the boundaries of each element (as can be seen in Figure 1) and $\{L_1, L_2\}$ is the classic basis for the linear polynomial space $\mathbb{P}_1([-1, 1])$. The η -axis, on the reference element, is aligned with dominant direction.

2.2 Field approximation

It is on the reference element \mathcal{K}_0 that we exploit the presence of a dominant direction. Each physical field w is approximated through the function w^h defined as element of the finite-dimensional space

$$\mathbb{T}_h^p = \left\{ w^h \in L^2(\Omega) : w^h \circ \chi_{\mathcal{K}}(\xi, \eta) = \sum_{j=1}^2 \sum_{i=1}^{p+1} w_{ij}^h \phi_i(\xi) L_j(\eta); \quad \mathcal{K} \in \mathcal{T}_h(\Omega) \right\} \quad (2)$$

where the integer parameter p is called of transversal enrichment, $\{L_1, L_2\}$ is the \mathbb{P}_1 classical basis and the functions $\{\phi_i : i = 1, \dots, p + 1\}$ are the Lagrange polynomial functions for the polynomial space \mathbb{P}_p defined over the set of Chebishev-Gauss-Lobatto (CGL) points. This is, each ϕ_i has the form

$$\phi_i(t) = \prod_{\substack{j=1 \\ j \neq i}}^{p+1} \frac{t - x_j}{x_j - x_i} \quad i = 1, \dots, r + 1, \quad (3)$$

where each node of the CGL set $\{x_i, i = 1, \dots, p + 1\}$ is defined as

$$x_i = -\cos\left(\frac{(i-1)\pi}{r}\right). \quad (4)$$

This low-high order interpolation strategy represents a natural and simple scheme of transversal enrichment, defined by the parameter p , without the introduction of auxiliar problems that must be solved in a *offline* step, like occurs in hierarchical modeling or reduced-basis approaches. The rationale of the proposed TEPEM approach is to combine the solution as the result of the combination of low order general purpose functions with enriched transversal functions to capture the major transversal phenomena. This drastically reduces the size of the discrete problem (and therefore the computational time), yet preserving accuracy.

3 Reduced setting for advection-diffusion equation

Let the generic advection-diffusion problem be defined as: Find $u \in \mathcal{U}$ such that

$$a(u, v) = F(v), \quad \forall v \in \mathcal{V} \quad (5)$$

where

$$a(u, v) = \int_{\Omega} \mu \nabla u \cdot \nabla v + (\mathbf{b} \cdot \nabla u + \sigma u)v \, d\Omega \quad (6)$$

$$F(v) = \int_{\Omega} fv \, d\Omega \quad (7)$$

with $\mu \in L^{\infty}(\Omega)$, $\mathbf{b} \in [W^{1,\infty}(\Omega)]^2$, $f \in L^2(\Omega)$. As classical in by-element methods, addressing this problem via the TEPEM implies the discrete problem

$$\mathbf{A}\mathbf{u} = \mathbf{f}, \quad \mathbf{A} \in \mathbb{R}^{N \times N}, \mathbf{u}, \mathbf{f} \in \mathbb{R}^N \quad (8)$$

where $N = (n_{1D}+1)(p+1)$, \mathbf{A} (equivalently \mathbf{f}) is associated to the discrete form of $a(\cdot, \cdot)$ ($F(\cdot)$) and the vector \mathbf{u} is composed by the nodal values of the solution field. Here n_{1D} stand by the number of pipe-elements in the partition of Ω and p by the transversal enrichment. The complete discrete problem is equivalent to a block tridiagonal linear system with dense sub-blocks when the dominant discretization is considered at the outermost level of approximation. The global matrix structure can be seen in Figure 2.

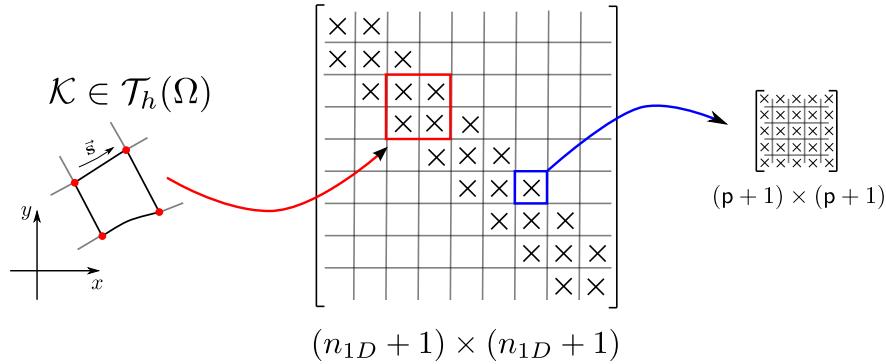


Figure 2: Algebraic structure of the linear system in the 2D scalar problem. Here n_{1D} stands for the number of pipe-type elements and p for the transversal interpolation order.

4 Numerical assessment

To show the effectiveness of the proposed approach, let us consider $\Omega \in \mathbb{R}^2$ the domain limited by the curves:

$$\begin{aligned} \Gamma_i &= \{(x, y) \in \mathbb{R}^2 : x = 0, 0 \leq y \leq 1\} \\ \Gamma_o &= \{(x, y) \in \mathbb{R}^2 : x = 2, 0 \leq y \leq 1\} \\ \Gamma_b &= \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 2, y = 1 - 0.25 \sin(2\pi x)\} \\ \Gamma_t &= \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 2, y = 2 + 0.25 \sin(2\pi x)\} \end{aligned}$$

and the following 2D advection-diffusion problem: Find $u \in \mathcal{U}$

$$\int_{\Omega} \mu \nabla u \cdot \nabla v + (\mathbf{b} \cdot \nabla u)v \, d\Omega = \int_{\Omega} fv \, d\Omega \quad \forall v \in \mathcal{V} \quad (9)$$

where $\mu = 1$, $\mathbf{b} = (100, 30 \sin(2\pi x))^T$, $f = 10$ and the linear manifold of admissible solutions (\mathcal{U}) and the associated linear space of admissible variations (\mathcal{V}) are defined as

$$\begin{aligned}\mathcal{U} &= \{u \in H^1(\Omega) : u|_{\Gamma_i} = -\sin(\pi y), u|_{\Gamma_b} = u|_{\Gamma_t} = 0\} \\ \mathcal{V} &= \{u \in H^1(\Omega) : u|_{\Gamma_i} = u|_{\Gamma_b} = u|_{\Gamma_t} = 0\}\end{aligned}$$

As no analytical solution is available for this problem, we consider as reference solution an approximate solution obtained via classical FEM on a very fine triangular mesh, made of 168 838 triangles and 84 838 DoFs. This reference solution will be simply denoted u (the TEPEM approximation will be denoted by u^h). The solution exhibits a clear predominant direction aligned with x -axis and also present high gradients near the superior and inferior boundaries as can be seen in Figure 3. This last characteristic allows us to show the capacity that transversal enrichment has to naturally deal with gradients in the transversal direction. Regarding to the numerical TEPEM convergence, Figure 3 make evident high convergence rates for both the field and its gradient and also the effectiveness of TEPEM approach in reducing the numerical error when increasing the DoFs when compared with classical FEM. Furthermore, Tables 1-2 show the reduction in the number of degrees of freedom and the numerical error in the TEPEM approximation against the reference FEM solution, respectively.

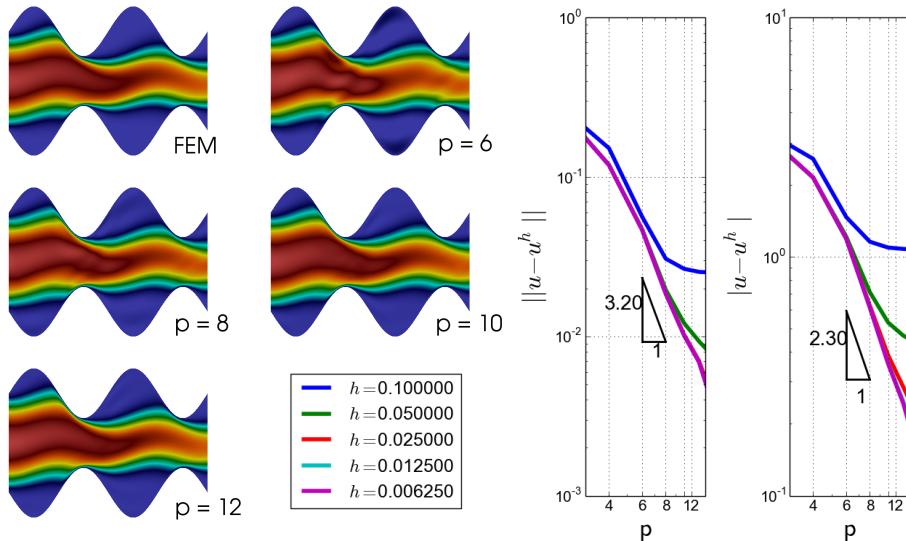


Figure 3: Comparison of the reference FEM solution and the TEPEM approximation for different degrees of transversal enrichment p . Also convergence rates for the field and its gradient is displayed. Usual notation was employed for the L^2 -norm ($\|\cdot\|$) and H^1 -seminorm ($|\cdot|$).

5 Final Remarks

The proposed approach, based on the enrichment of transversal polynomial interpolants (called TEPEM), can be seen as a simple yet promising methodology capable to deliver

Degrees of Freedom					
	$n_{1D} = 4$	$n_{1D} = 8$	$n_{1D} = 16$	$n_{1D} = 32$	$n_{1D} = 64$
p = 2	21	51	99	195	387
p = 4	35	85	165	325	645
p = 6	49	119	231	455	903
p = 8	63	153	297	585	1 161
p = 10	77	187	363	715	1 419

Table 1: Degrees of Freedom on TEPEM for different axial discretization and transversal enrichment.

		Number of slab-type elements				
		$n_{1D} = 4$	$n_{1D} = 8$	$n_{1D} = 16$	$n_{1D} = 32$	$n_{1D} = 64$
$\ u^h - u\ $	p = 2	0.321541	0.306105	0.305305	0.305265	0.305263
	p = 4	0.152767	0.121654	0.118924	0.118782	0.118774
	p = 6	0.127125	0.051941	0.046674	0.046641	0.046644
	p = 8	0.136135	0.032475	0.018757	0.018618	0.018628
	p = 10	0.133899	0.033999	0.010954	0.010204	0.010199
$ u^h - u $	p = 2	0.020043	0.019142	0.019021	0.019019	0.019018
	p = 4	0.013346	0.011148	0.011193	0.011167	0.011165
	p = 6	0.017969	0.006956	0.006204	0.006147	0.006142
	p = 8	0.023455	0.005951	0.003470	0.003282	0.003264
	p = 10	0.026526	0.006483	0.002351	0.002017	0.001979

Table 2: Numerical error in the field and its gradient between the TEPEM solution and the reference FEM solution obtained with an extremely fine triangular mesh. Usual notation was employed for the L^2 -norm ($\|\cdot\|$) and H^1 -seminorm ($|\cdot|$).

accurate results maintaining a reduced number of degrees of freedom. This, consequently, results in small algebraic problems, which leads to low computational cost. Moreover, the TEPEM can straightforwardly be extended to 3D problems. Finally, the results obtained in this work are extremely promising towards the application of the TEPEM strategy to more challenging and computationally demanding problems, are those encountered in computational hemodynamics.

6 Acknowledgement

This work was partially supported by the Brazilian agencies CNPq and FAPERJ. The support of these agencies is gratefully acknowledged.

References

- M. Aletti, S. Perotto, and A. Veneziani. Educated bases for the himod reduction of advection-diffusion-reaction problems with general boundary conditions. *MOX report*, 2015.
- G. Berkooz, P. Holmes, and J.L. Lumley. The proper orthogonal decomposition in the analysis of turbulent flows. *Annual review of fluid mechanics*, 25(1):539–575, 1993.
- P.J. Blanco, M.R. Pivello, S.A. Urquiza, and R.A. Feijoo. On the potentialities of 3d–

- 1d coupled models in hemodynamics simulations. *Journal of Biomechanics*, 42(7):919–930, 2009.
- P.J. Blanco, L.A. Mansilla Alvarez, and R.A. Feijóo. Hybrid element-based approximation for the Navier–Stokes equations in pipe-like domains. *Comp. Meth. Appl. Mech. Engrg.*, 283:971–993, 2015.
- M. Drohmann, B. Haasdonk, and M. Ohlberger. Adaptive reduced basis methods for nonlinear convection-diffusion equations. *Finite Volumes for Complex Applications VI-Problems & Perspectives*, 1:369–377, 2011.
- T.J.R. Hughes and J. Lubliner. On the one-dimensional theory of blood flow in the larger vessels. *Mathematical Biosciences*, 18(1-2):161–170, 1973.
- K. Kunisch and S. Volkwein. Galerkin proper orthogonal decomposition methods for parabolic problems. *Numerische mathematik*, 90(1):117–148, 2001.
- L.A. Mansilla Alvarez, P.J. Blanco, C.A. Bulant, E. Dari, A. Veneziani, and R.A. Feijóo. Transversally enriched pipe element method (tepem): An effective numerical approach for blood flow modeling. *Int. J. Num. Meth. Biomed. Engng.*, 33(4), 2017.
- A. Osiadacz. Simulation and analysis of gas networks. 1987.
- S. Perotto. A survey of hierarchical model (hi-mod) reduction methods for elliptic problems. In *Numerical simulations of coupled problems in engineering*, pages 217–241. Springer, 2014.
- J.S. Peterson. The reduced basis method for incompressible viscous flow calculations. *SIAM Journal on Scientific and Statistical Computing*, 10(4):777–786, 1989.