

# Análise do Emprego de Ferramentas de Computação Natural no Problema de Separação de Misturas com Não-Linearidade Posterior

Filipe O. Pereira<sup>1</sup>, Everton Z. Nadalin<sup>1</sup>, Ricardo Suyama<sup>2</sup>, Romis Attux<sup>1</sup>

**Resumo**—Neste trabalho, investiga-se a aplicação de duas técnicas de otimização bio-inspiradas – um algoritmo imunológico e um método de enxame de partículas – ao problema de separação de misturas com não-linearidade posterior (PNL, de *post-nonlinear mixtures*). As técnicas são analisadas em dois cenários distintos e os resultados obtidos atestam a sua robustez e a sua eficiência no contexto do problema em questão. Consideramos que o estudo apresentado também se reveste de interesse como uma avaliação do potencial de ferramentas bio-inspiradas em domínios de processamento marcados pela existência de não-linearidades e pelo uso de critérios baseados na teoria de informação.

**Palavras-chave**—Separação Cega de Fontes, Misturas com Não-Linearidade Posterior, Computação Bio-Inspirada.

**Abstract**— In this work, we investigate the application of two bio-inspired optimization techniques – an artificial immune system and a particle swarm method – to the problem of source separation under post-nonlinear (PNL) mixture models. The techniques are analyzed in two distinct scenarios and the obtained results attest their robustness and their efficiency in the context of the focused problem. We consider that the presented study is also relevant as an assessment of the potential of bio-inspired search tools in signal processing domains characterized by the existence of nonlinearities and the use of information-theoretic criteria.

**Index Terms**— Blind Source Separation, Post-Nonlinear Mixtures, Bio-Inspired Computing.

## I. INTRODUÇÃO

O problema de Separação Cega de Fontes (Blind Source Separation – BSS) se associa à idéia de recuperar um conjunto de sinais - denominados fontes - a partir de observações que são misturas desconhecidas deles. Até o fim da última década, a maioria das técnicas propostas havia sido desenvolvida para resolver o problema de BSS no contexto de misturas lineares e instantâneas. Porém, em alguns cenários práticos de interesse [1], como uma consequência, por exemplo, de certas

características de sensores e amplificadores, é de grande importância lidar com modelos de mistura não-lineares. Tendo isso em vista, tornou-se como objeto de estudo o problema de BSS não-linear, e, desse esforço, emergiu o modelo de mistura com não-linearidade posterior (*post-nonlinear* – PNL) [2], o qual é particularmente atraente por permitir, no caso em que as fontes são estatisticamente independentes, a sua recuperação por meio de metodologias de Análise por Componentes Independentes (*Independent Component Analysis* – ICA).

Uma forma bastante direta de aplicar a noção de ICA a problemas de BSS é lançar mão de um critério baseado na informação mútua entre as estimativas das fontes. Em [3], uma estratégia desse tipo foi usada, em contraste com a abordagem original de Taleb e Jutten [2], tendo por base uma metodologia de otimização baseada num algoritmo bio-inspirado – um sistema imunológico artificial denominado *opt-aiNet* [4]. Em relação a metodologias mais clássicas, a estratégia de computação natural apresentava vantagens interessantes como 1) o emprego exclusivo de informação de ordem zero sobre a função custo estabelecida – evitando a necessidade de calcular derivadas da função, que pode não se prestar facilmente a tais cálculos, uma vez que podem estar envolvidas estimativas genéricas de densidades de probabilidade – e 2) apesar de não garantirem convergência global, estes métodos possuem uma significativa capacidade de exploração do espaço de busca, tornando menor a probabilidade de convergência para mínimos locais ruins em relação às técnicas baseadas em gradiente.

Tendo em mente as vantagens mencionadas anteriormente, propomos, neste trabalho, a realização de uma análise de duas ferramentas de otimização baseadas na computação natural – um algoritmo imunológico e um método de enxame de partículas – ao problema de separação de misturas PNL. Visto que os resultados apresentados em [3] foram promissores, utilizaremos estas técnicas que tendem a levar a uma menor complexidade computacional, com o intuito de obter resultados semelhantes como em [3] e assim buscaremos analisar o desempenho desses métodos à luz de métricas como taxa de convergência global em múltiplas realizações e, ademais, teceremos elementos de comparação com a metodologia adotada em [3]. Com isso, pretendemos formar um quadro mais claro da aplicabilidade de métodos bio-inspirados a tarefas de fronteira em processamento adaptativo.

Este trabalho está estruturado da seguinte forma: na seção II, são expostos os fundamentos do problema de Separação Cega de Fontes com misturas PNL; na Seção III, discute-se a

<sup>1</sup> Departamento de Eng. de Computação e Automação Industrial, <sup>2</sup> Departamento de Microonda e Óptica, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, São Paulo, Brasil. E-mails: {filipe.nadalin,attux}@dca.fee.unicamp.br; {rsuyama}@dmo.fee.unicamp.br; Os autores agradecem à CAPES e à FAPESP o apoio financeiro recebido.

aplicação de uma metodologia baseada em informação mútua e estatísticas de ordem, o que leva à formulação do paradigma de separação que serve de base a nosso esforço. Na seção IV, discutem-se os algoritmos de otimização escolhidos, enquanto a seção V é devotada à apresentação e análise de resultados e a seção VI às conclusões.

## II. SEPARAÇÃO DE MISTURAS COM NÃO-LINEARIDADE POSTERIOR ATRAVÉS DE ICA

Na formulação clássica de um problema de separação de fontes via análise de componentes independentes (ICA), procura-se estimar, de modo não-supervisionado, um conjunto de sinais de interesse, supostos mutuamente independentes e não-gaussianos, a partir de versões misturadas dos mesmos. Seja  $s(t) = [s_1(t), \dots, s_N(t)]^T$  o vetor de sinais das fontes e  $x(t) = [x_1(t), \dots, x_N(t)]^T$  o vetor de misturas (consideramos um número igual de misturas e fontes). Caso se lide com misturas instantâneas e lineares – o caso mais usual na literatura [1] –, matematicamente, as misturas são combinações lineares das fontes e podem ser representadas na forma matricial:

$$\mathbf{x}(t) = \mathbf{A}s(t) \quad (1)$$

onde  $\mathbf{A}$  denota a matriz de mistura. Neste caso, uma possibilidade natural é realizar a separação multiplicando o vetor de mistura por uma matriz de separação  $\mathbf{W}$ :

$$\mathbf{y}(t) = \mathbf{W}\mathbf{x}(t) \quad (2)$$

A aplicação de ICA ao problema de separação se dá a partir da idéia de escolher  $\mathbf{W}$  de modo que os elementos de  $\mathbf{y}(t)$  sejam estatisticamente independentes [1]. Nos casos em que a inversão do modelo de mistura é possível, isso leva à recuperação das fontes a menos de ambigüidades de permutação e fator de escala [5].

Uma extensão para o caso não-linear que permite a manutenção desse arcabouço teórico é baseada na introdução de não-linearidades inversíveis posteriores à mistura linear [2]. Esse modelo, mostrado na Fig. 1, recebe o nome de modelo de mistura com não-linearidade posterior (PNL). Matematicamente, o processo de mistura é dado por:

$$\mathbf{x}(t) = \mathbf{f}(\mathbf{A}s(t)) \quad (3)$$

onde  $\mathbf{f}(\cdot) = [f_1(\cdot), \dots, f_N(\cdot)]^T$  são as não-linearidades inversíveis e sem memória associadas às diversas saídas do processo de mistura linear. A matriz  $\mathbf{A}$  também deve ser inversível para que a separação seja viável.

Um candidato natural a sistema separador nesse caso é:

$$\mathbf{y}(t) = \mathbf{W}\mathbf{g}(\mathbf{x}(t)) \quad (4)$$

Onde  $\mathbf{g}(\cdot) = [g_1(\cdot), \dots, g_N(\cdot)]^T$  são funções não-lineares que devem ser corretamente ajustadas para “anular o efeito” de  $\mathbf{f}(\cdot)$ , ou seja, fazer com que a cascata de funções  $g_i(\cdot)$  e  $f_i(\cdot)$ , para  $i = 1, \dots, N$ , sempre seja uma função linear.

De posse desses modelos, o processo de separação passa a depender de dois aspectos fundamentais: a escolha de um

critério que permita quantificar o grau de independência entre as saídas do separador e de um método de parametrização das funções não-lineares do separador. Um ponto importante é que essa parametrização não deve permitir que funções não-

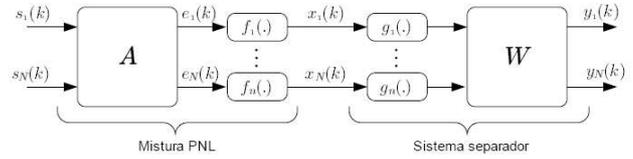


Fig.1. Sistema com Não-Linearidade Posterior

monotônicas sejam obtidas, pois isso pode distorcer o conteúdo de informação do qual inevitavelmente o critério escolhido dependerá. Neste trabalho, seguindo a metodologia proposta em [3], adotaremos a informação mútua como critério fundamental e uma parametrização polinomial com restrição de monotonicidade para as funções do sistema separador. Nessa abordagem, um ponto crucial se vincula à estrutura da função custo baseada na informação mútua: tratemos, pois, desse assunto em mais detalhe.

## III. A FUNÇÃO CUSTO BASEADA NA INFORMAÇÃO MÚTUA E O PROBLEMA DE OTIMIZAÇÃO DELA DECORRENTE

Para obter o sistema separador no âmbito de um problema de misturas PNL, podemos, como visto na seção anterior, recorrer à ICA. É necessário, portanto, uma função custo para quantificar o grau de dependência estatística entre os componentes do vetor de saída  $\mathbf{y}$  (omitiremos o índice temporal para simplificar a notação). Uma grandeza estatística interessante, nesse caso, é a informação mútua, definida como:

$$I(\mathbf{y}) = \sum_i H(y_i) - H(\mathbf{y}) \quad (5)$$

onde  $H(\mathbf{y})$  representa a entropia diferencial conjunta de  $\mathbf{y}$  e  $H(y_i)$  a entropia diferencial de cada um dos elementos de  $\mathbf{y}$ . Considerando a estrutura de separação mostrada na Fig. 1, pode-se expressar a informação mútua das saídas do separador, considerando que as funções  $g_i(\cdot)$  são inversíveis, da seguinte forma:

$$I(\mathbf{y}) = \sum_i H(y_i) - H(\mathbf{x}) - \log|\det \mathbf{W}| - E\left\{\log \prod_i |g'_i(x_i)|\right\} \quad (6)$$

onde  $g'_i(\cdot)$  denota a primeira derivada da  $i$ -ésima não-linearidade  $g_i(\cdot)$  do sistema separador. Analisando (6), percebe-se que, para minimizar a informação mútua, dever-se-ia estimar, a rigor, todas as entropias  $H(y_i)$  e  $H(\mathbf{x})$ . A entropia conjunta das misturas  $H(\mathbf{x})$  não depende, no entanto, dos parâmetros do sistema separador, o que torna desnecessário o seu cálculo do ponto de vista de otimização esboçado acima. Com isso, resta apenas resolver o problema fundamental de estimação da entropia de cada um dos elementos de  $\mathbf{y}$ , ou seja, das entropias marginais  $H(y_i)$ , para que se disponha dos alicerces de um arcabouço eficiente de separação de misturas PNL via ICA. Neste trabalho, a estimação das entropias marginais foi realizada por meio de uma metodologia baseada em estatísticas de ordem [6], que será descrita a seguir.

### A. Estimação de Entropia Utilizando Estatísticas de Ordem

A metodologia de estimação de entropia baseada em *estatísticas de ordem* [6] apresenta um grau significativo de robustez e uma complexidade computacional tipicamente menor que a associada a métodos baseados na estimação de densidades de probabilidade. Tal metodologia tem sido usada com sucesso tanto para separação de misturas lineares quanto de não-lineares.

Considere um conjunto de variáveis aleatórias  $Y_1, Y_2, \dots, Y_T$ . A cada realização  $y_1, y_2, \dots, y_T$  deste conjunto, o valor  $m = \max y_i$  também representará a realização de uma variável aleatória, que representaremos por  $Y_{(T:T)}$ . De maneira equivalente, poderíamos considerar o valor mínimo em cada realização, o que por sua vez representa a realização de uma outra variável aleatória. Generalizando este conceito, chegamos a um conjunto de variáveis aleatórias ordenadas,  $Y_{(1:T)} \leq Y_{(2:T)} \leq \dots \leq Y_{(T:T)}$ , onde  $Y_{(t:T)}$  corresponde à  $t$ -ésima *estatística de ordem*, e representa a variável aleatória associada ao  $t$ -ésimo menor valor das realizações de  $Y_1, Y_2, \dots, Y_T$ .

Note que o conceito de estatística de ordem também pode ser definido para uma única variável aleatória. De fato, cada realização de um conjunto de  $T$  variáveis aleatórias com distribuições de probabilidade idênticas pode ser vista como  $T$  realizações independentes de uma única variável aleatória. O interesse em se empregar o conceito de estatística de ordem na estimação da entropia reside em sua ligação com a função quantil de uma variável aleatória  $Q_Y(\cdot)$ , dada por:

$$E\{Y_{(t:T)}\} = Q_Y\left(\frac{t}{T+1}\right) \quad t=1, \dots, T \quad (7)$$

Por sua vez, a entropia de uma variável aleatória pode ser definida em termos da função quantil, e expressa por:

$$H(Y) = \int_0^1 \log Q'_Y(u) du \quad (8)$$

onde  $Q'_Y$  denota a derivada da função quantil. Empregando uma aproximação numérica para se calcular a integral em (8) obtemos

$$H(Y) = \sum_{l=2}^L \log \left[ \frac{Q_Y(u_l) - Q_Y(u_{l-1})}{u_l - u_{l-1}} \right] \frac{u_l - u_{l-1}}{u_L - u_1} \quad (9)$$

onde  $u_1, \dots, u_L$  é um conjunto de números crescentes no intervalo  $[0, 1]$ . Portanto, empregando as estatísticas de ordem podemos obter informação sobre a função quantil, que, por sua vez, pode ser utilizada para calcular a entropia  $H(Y)$ .

No caso em que há apenas uma realização de  $T$  amostras da variável aleatória  $Y$ , a estimação da função quantil pode ser feita, de maneira aproximada, com base na realização correspondente das estatísticas de ordem, ou seja

$$Y_{(t:T)} \approx Q_Y\left(\frac{t}{T+1}\right) \quad t=1, \dots, T \quad (10)$$

Para valores da função quantil que não pertencem ao conjunto  $t/(T+1)$ , é possível realizar uma interpolação linear, gerando um estimador de entropia simples e eficaz.

### B. Uma Visão do Problema de Otimização Resultante

O arcabouço conceitual estabelecido se baseia na idéia de aplicar ICA ao problema de separação de misturas PNL tendo por base uma função custo de informação mútua. Tal arcabouço, como discutido anteriormente, é, de uma perspectiva ideal, adequado ao problema em questão. No entanto, para que a eficiência ideal se reflita num adequado desempenho do sistema separador em aplicações práticas, é preciso resolver adequadamente o problema de otimização da estimativa da informação mútua relativamente aos parâmetros do separador. Supondo que o estimador baseado em estatísticas de ordem esteja fornecendo valores precisos para as entropias marginais, restarão dois problemas fundamentais a serem convenientemente abordados:

1) A função custo resultante deverá ter um caráter multimodal pronunciado, ou seja, existe a perspectiva concreta de que haja ótimos locais [2][3].

2) Tendo em vista que a função custo se baseia numa metodologia de estimação de estatísticas de ordem, torna-se complicado manipulá-la diretamente para obter informações como derivadas com respeito a parâmetros livres.

Esses problemas fundamentais indicam que métodos alternativos às abordagens clássicas de otimização não-linear são opções interessantes para o problema de separação de misturas PNL. Neste trabalho, adotaremos dois algoritmos bio-inspirados de estrutura relativamente simples como alternativas que são capazes de realizar otimização com grande potencial de busca global a partir, exclusivamente, do conhecimento de valores da função custo: um algoritmo imunológico e um algoritmo de enxame de partículas, que serão descritos na próxima seção. Nesse sentido, o trabalho corresponde a uma extensão do esforço reportado em [3] no sentido de buscar abordagens menos complexas e de realizar uma análise mais sólida em termos de robustez no que se refere à convergência para soluções sub-ótimas.

## IV. COMPUTAÇÃO NATURAL E OTIMIZAÇÃO: SISTEMAS IMUNOLÓGICOS ARTIFICIAIS E INTELIGÊNCIA DE ENXAME

Técnicas de otimização baseadas em computação natural têm se mostrado eficientes em termos de capacidade de busca global em problemas de difícil ou impossível tratamento por metodologias clássicas de otimização não-linear [7]. Neste trabalho, optamos por dois algoritmos com inspirações diferentes: um sistema imunológico artificial (SIA), que pode ser visto como uma extensão do algoritmo CLONALG [8] no sentido de lidar com codificação real, e uma implementação de uma técnica de enxame de partículas (*particle swarm*) [9], que tem por fundamento a noção de organização coletiva do conhecimento. O primeiro se baseia em princípios, teorias e modelos dos sistemas imunológicos naturais, embora estruturalmente apresente pontos de contato com ferramentas mais clássicas como algoritmos genéticos e estratégias evolutivas, enquanto o segundo método baseia-se na simulação do comportamento social. Ambos os métodos são estruturalmente mais simples que a rede imunológica artificial empregada num contexto similar em [3], o que constitui um

passo no sentido de buscar bases para um efetivo compromisso entre desempenho e complexidade computacional no que se refere ao problema de separação de misturas PNL.

#### A. Sistemas Imunológicos Artificiais

Sistemas Imunológicos Artificiais (SIAs) [7] são técnicas que buscam representar em computador características do sistema de defesa de organismos superiores contra invasores. Essas técnicas podem ter finalidades diversas, como o modelamento de sistemas biológicos e a solução de problemas de engenharia.

No contexto das aplicações práticas de sistemas imunológicos artificiais, merece grande destaque o uso dessas ferramentas em problemas de otimização de funções [4][8]. Em otimização, uma idéia importante é a de traçar um paralelo entre o problema a ser resolvido e o processo de otimização evolutiva que ocorre em nosso organismo numa escala microscópica e que permite que surjam anticorpos adaptados a novos antígenos. Nesse caso, tipicamente, considera-se que cada solução do problema de otimização corresponde à estrutura de um anticorpo e que a função custo a ser otimizada é uma medida da adequação dessa estrutura a um certo antígeno (que chamaremos de *fitness*).

Um algoritmo que emprega essas idéias é o CLONALG, proposto em [8]. Tal algoritmo se baseia fundamentalmente no princípio da seleção clonal, que, em termos simples, corresponde a um processo em que células de defesa são clonadas e sofrem modificações aleatórias, havendo uma tendência de células mais bem adaptadas ao reconhecimento do antígeno a ser proliferarem mais. Uma forma direta de trazer essa idéia para o âmbito de uma ferramenta de busca é supor que cada solução do problema dá origem a um certo número de réplicas (clones), que, por sua vez, sofrem um processo de mutação por meio da adição de valores aleatórios gaussianos, sendo a variância da distribuição empregada tanto maior quanto menos adaptado for o anticorpo. Ademais, adota-se, tendo por base a idéia de *edição de receptores*, a inserção periódica de novas soluções geradas aleatoriamente, o que permite ao algoritmo explorar novas regiões do espaço de busca. O resultado da interação entre os mecanismos de seleção clonal e a inserção de novos indivíduos é um algoritmo com capacidade de busca local e de busca global, adequado, portanto, a problemas de otimização multimodal como os que nos interessam em separação não-linear. Vale ressaltar que a opt-aiNet, o algoritmo imunológico empregado em [3], tem uma dinâmica similar à do CLONALG, sendo, no entanto, dotada de um mecanismo de controle automático de população que se baseia na idéia de *rede imunológica* [4]. Esse mecanismo pode, se bem ajustado, ser útil para permitir ao método uma maior adequação do tamanho de sua população, mas, por outro lado, torna o algoritmo estruturalmente complexo e o ajuste de seus parâmetros livres mais delicado que o do CLONALG. Nossa escolha se deveu, portanto, à busca por uma solução que combinasse eficiência e parcimônia estrutural.

No algoritmo CLONALG, busca-se a maximização do *fitness*. Porém, estamos interessados na minimização da informação mútua. Além disso, é desejável que o *fitness*

assuma sempre valores positivos. Deste modo, o *fitness* pode ser definido no seguinte modo:

$$J_{Fit} = \frac{1}{\sum_i H(y_i) - \log|\det \mathbf{W}| - E\left\{\log \prod_i |g'_i(x_i)|\right\}} \quad (11)$$

O denominador desta expressão vem de (6), porém, é desconsiderada a entropia conjunta das misturas, já que esta medida não depende dos parâmetros do sistema separador. As entropias marginais são calculadas de acordo com a abordagem de estatísticas de ordem, discutida na seção III.

A seguir, na Tabela I, apresentamos um pseudo-código que ilustra o funcionamento do CLONALG.

TABELA I  
Descrição do algoritmo CLONALG

<ol style="list-style-type: none"> <li>1. Inicialize aleatoriamente a população (<math>N</math>);</li> <li>2. Determine o <i>fitness</i> de cada indivíduo: <math>J_{Fit}</math>;</li> <li>3. Enquanto o número máximo de gerações não é atingido, faça;             <ol style="list-style-type: none"> <li>3.1. Crie <math>N_c</math> clones para cada indivíduo;</li> <li>3.2. Mantenha o indivíduo original e aplique um processo de mutação para cada clone seguindo a equação:                 <math display="block">c' = c + \alpha Y(0,1) \quad (12)</math> <math display="block">\alpha = \frac{1}{\beta} \exp(-J_{Fit}) \quad (13)</math> </li> <li>3.3. Avalie o <i>fitness</i> de cada novo indivíduo e mantenha na população somente a melhor solução de cada grupo formado pelo indivíduo e seus clones mutados.</li> <li>3.4. A cada <math>t</math> iterações, elimine os <math>m</math> elementos da população com os menores valores de <i>fitness</i> e os substitua por indivíduos gerados aleatoriamente;</li> </ol> </li> <li>4. Retorne ao passo 3.</li> </ol>
---

O processo de mutação está representado no passo 3.2, em que as cópias do vetor de parâmetros estão sujeitas a uma modificação proporcional ao valor de *fitness*. Nessa etapa, juntamente com o passo 3.3, o algoritmo estabelece um eficiente mecanismo de busca local, apesar de haver também um certo potencial de busca global na mutação. Em 3.4, há um aumento determinante no potencial de busca global, causado pela inserção de indivíduos aleatórios na população, o que aumenta a diversidade populacional e permite a exploração de novas porções no espaço de busca.

#### B. Métodos de Otimização Baseados em Enxames de Partículas

O estudo do comportamento de animais que vivem em sociedade é uma riquíssima fonte de idéias e estratégias relevantes para a solução de diversos tipos de problemas de interesse prático [7]. Neste trabalho, em que lidamos com problemas de otimização de natureza contínua, destacamos as técnicas de *enxames de partículas* (*particle swarm*), que buscam representar o deslocamento de um grupo de indivíduos tendo por base o conhecimento individual

adquirido sobre o ambiente e também o conhecimento construído pelo grupo como um todo.

Na implementação clássica de métodos de otimização baseados em enxames de partículas (PSO, do inglês *particle swarm optimization*), cada solução do problema de otimização abordado corresponde a uma partícula que ocupa uma certa posição no espaço de busca. Essa partícula tem a capacidade de armazenar a região de melhor valor de custo em que esteve, e, ademais, é capaz de ter acesso à melhor posição já visitada pelas partículas com as quais tem uma relação de vizinhança topológica [9]. A melhor posição pessoal e a melhor posição relativamente aos vizinhos definem duas direções de busca que são combinadas para formar a expressão de atualização de cada partícula. Uma escolha cuidadosa da relação de vizinhança permite que o método seja capaz de explorar convenientemente o espaço de busca, enquanto, ao mesmo tempo, é possível dispor de um certo grau de refinamento das soluções já existentes. Uma análise do pseudo-código a seguir, que detalha a implementação do algoritmo, revelará ao leitor que se trata de um método de estrutura simples, embora, para um número elevado de indivíduos na população, pode haver uma certa demanda por recursos para que se implementem determinadas relações de vizinhança. Outro fator que chama a atenção é o pequeno número de parâmetros que devem ser ajustados no algoritmo, o que o torna flexível à atuação em cenários desconhecidos. Na Tabela II, apresentamos um pseudo-código que ilustra o funcionamento da técnica. Vale ressaltar que, neste trabalho, optamos por uma relação em que todas as partículas são vizinhas.

TABELA II  
Descrição do Algoritmo de Enxame de Partículas

<p>1. Inicializar as partículas, escolher os valores de velocidade máxima e mínima (<math>V_{max}</math> e <math>V_{min}</math>) e calcular o <i>fitness</i> (custo) inicial de todas as partículas.</p> <p>2. Para <math>i = 1</math> até <math>N</math>, fazer:</p> <p>2.1 se <math>f(\mathbf{x}_i) &gt; f(\mathbf{p}_i)</math> então <math>\mathbf{p}_i = \mathbf{x}_i</math></p> <p>2.2 <math>g = i</math></p> <p>2.3 se <math>f(\mathbf{p}_j) &lt; f(\mathbf{p}_g)</math>, para <math>j =</math> índices dos vizinhos, então <math>g = j</math></p> <p>2.4 Atualizar velocidade e posição:</p> $\Delta \mathbf{x}_i \leftarrow \Delta \mathbf{x}_i + \varphi_1 \otimes (\mathbf{p}_i - \mathbf{x}_i) + \varphi_2 \otimes (\mathbf{p}_g - \mathbf{x}_i) \quad (14)$ $\Delta \mathbf{x}_i \in [V_{min}, V_{max}] \quad (15)$ $\mathbf{x}_i \leftarrow \mathbf{x}_i + \Delta \mathbf{x}_i \quad (16)$ <p>3. Repetir procedimento a partir do passo 2 até que o critério de parada seja satisfeito.</p>
--

Da equação (14), em que ocorre a atualização da velocidade, temos que:  $\varphi_1$  e  $\varphi_2$  representam vetores aleatórios com valores positivos derivados de uma distribuição uniforme entre 0 e  $AC_1$  e entre 0 e  $AC_2$ , respectivamente, sendo  $AC_1$  e  $AC_2$  chamadas de constantes de aceleração. O segundo termo do segundo membro da equação (14) é proporcional à diferença entre a melhor posição ( $\mathbf{p}_i$ ) já visitada por um indivíduo e a sua posição atual ( $\mathbf{x}_i$ ) e o último termo é proporcional à diferença entre a melhor posição ocupada por um vizinho ( $\mathbf{p}_g$ ) e a posição atual. O símbolo  $\otimes$  representa a multiplicação de vetores elemento a elemento. Para limitar a

mudança de posição de um indivíduo, definem-se dois valores,  $V_{min}$  e  $V_{max}$ , que não podem ser excedidos por nenhum elemento de  $\Delta \mathbf{x}$  (vide (15)).

Tendo apresentado as técnicas, estamos aptos a estudar o desempenho delas no problema PNL descrito anteriormente.

## V. RESULTADOS

Para avaliar o desempenho dos algoritmos bio-inspirados, realizamos simulações em dois cenários de misturas PNL distintos, sendo que a tarefa é realizar separação cega de fontes de modo a se obter um desempenho (em termos de erro quadrático médio) tão bom quanto possível sem que se empreguem excessivamente os recursos computacionais disponíveis.

Foram considerados cenários de separação de duas e três fontes, sendo que, em nosso estudo, optamos por lidar com sinais uniformemente distribuídos no intervalo  $[-1,1]$ . Seguindo, em linhas gerais, os cenários estudados em [3], definimos o sistema misturador PNL no primeiro cenário como mostrado em (17) e no segundo como em (18):

$$\mathbf{A} = \begin{bmatrix} 1 & 0,6 \\ 0,6 & 1 \end{bmatrix} \quad e \quad \begin{cases} f_1(e_1) = \tanh(e_1) \\ f_2(e_2) = \sqrt[5]{e_2} \end{cases} \quad (17)$$

$$\mathbf{A} = \begin{bmatrix} 1 & 0,6 & 0,5 \\ 0,5 & 1 & 0,4 \\ 0,4 & 0,6 & 1 \end{bmatrix} \quad e \quad \begin{cases} f_1(e_1) = 2\sqrt[3]{e_1} \\ f_2(e_2) = 2\sqrt[3]{e_2} \\ f_3(e_3) = 2\sqrt[3]{e_3} \end{cases} \quad (18)$$

Em ambos os casos, o sistema separador é constituído de uma matriz quadrada  $\mathbf{W}$  e de funções não-lineares polinomiais de quinto grau do tipo  $y = ax^5 + bx^3 + cx$ . Para que as não-linearidades sejam sempre monotônicas, os coeficientes de cada polinômio foram restritos a serem positivos.

Para realizar os experimentos para o cenário de duas fontes, os parâmetros do algoritmo CLONALG foram ajustados da seguinte maneira:  $N = 10$ ,  $N_c = 7$  e  $\beta = 60$ , sendo o número máximo de gerações igual a 10000. Os parâmetros do algoritmo de enxame de partícula foram  $N = 60$ ,  $AC_1$  e  $AC_2 = 2.05$ ,  $V_{max} = 0.1$  e  $V_{min} = -0.1$ , com uma quantidade de gerações foi igual a 4490. Em ambos os casos, os parâmetros foram escolhidos de modo a permitir uma comparação adequada com [3].

Foram realizados dois experimentos diferentes, nos quais os algoritmos foram executados 25 vezes com as mesmas realizações das duas fontes e, no segundo, os algoritmos foram executados 50 vezes com realizações diferentes. Na Tabela III, apresentamos os valores de erro quadrático médio (EQM) obtidos para ambas as fontes com cada método nas duas modalidades de teste. Ressaltamos que os valores de EQM foram calculados após a normalização da potência de todos os sinais, de modo que se tenha uma base mais efetiva de comparação. Verificamos, primeiramente, que ambas as técnicas são capazes de atingir um nível médio de erro bastante reduzido, o que atesta a sua eficiência em termos de busca no âmbito do problema PNL. Percebe-se, ainda, que o desempenho de ambas as técnicas foi razoavelmente equivalente. Na Fig. 2, apresentamos uma representação da distribuição das fontes, das misturas e das saídas do separador

para uma realização típica com o algoritmo CLONALG. Percebemos que o separador é capaz de contrabalançar o efeito nocivo do processo de mistura e obter uma distribuição próxima à ideal: uniforme e com sinais independentes. Nesse sentido, a análise apresentada atesta a robustez dos métodos de busca e mostra que as duas técnicas analisadas são capazes de obter um desempenho não inferior ao da abordagem de [3] (embora no artigo não se tenha realizado um levantamento dos erros quadráticos médios para um cenário de duas fontes, restringindo-se a análise ao estudo da distribuição dos sinais).

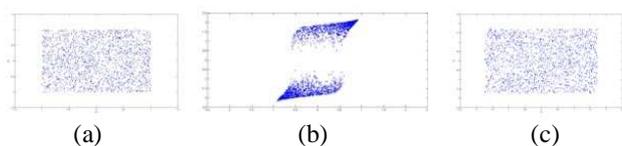


Fig. 2. Primeiro Cenário – Clonalg – fontes (a), misturas (b) e sinais separados (c)

TABELA III  
EQM – Primeiro Cenário

Quantidade de Simulações	EQM ( $\times 10^{-2}$ )	$y_1$	$y_2$
25	Clonalg	0,11	0,65
	Swarm	0,73	0,52
50	Clonalg	0,77	0,43
	Swarm	0,62	0,47

Para realizar os experimentos para o segundo cenário, com três fontes, os parâmetros do algoritmo CLONALG foram ajustados da seguinte maneira:  $N = 50$ ,  $N_c = 15$  e  $\beta = 40$ , sendo a quantidade máxima de gerações igual a 10000. Os parâmetros do algoritmo *swarm* foram  $N = 200$ ,  $AC_1$  e  $AC_2 = 2.05$ ,  $V_{max} = 1$  e  $V_{min} = -1$ , sendo a quantidade máxima de gerações igual a 13400. Os parâmetros foram ajustados por meio de uma série de experimentos preliminares, tendo em vista a idéia de obter um bom compromisso entre parcimônia e precisão das estimativas.

Na Tabela IV, são apresentados os valores de EQM associados à estimação de cada fonte para os dois cenários experimentais. Neste caso, podemos considerar que o CLONALG teve o melhor desempenho, tendo atingido níveis de EQM aproximadamente uma ordem de grandeza menores que os reportados em [3]. Apesar disso, o *swarm* teve um desempenho bastante razoável, com erros médios, mesmo no pior caso, comparáveis aos obtidos em [3]. Vale ainda ressaltar que, especialmente no caso com 50 simulações, os valores mais elevados de EQM médio obtidos com o *swarm* são devidos à não convergência em 8 casos: nas demais simulações, o nível de EQM obtido foi comparável ao EQM médio associado ao CLONALG. De qualquer forma, podemos considerar a técnica imunológica como a mais robusta no cenário em questão. Embora não seja trivial realizar uma análise desse tipo quando se lida com técnicas de caráter marcadamente estocástico, julgamos que o melhor desempenho do CLONALG nesse caso com um espaço de busca de maior dimensão se deva à sinergia entre um eficiente mecanismo de busca global (a mutação) e um esquema de introdução de diversidade simples e eficiente. Podemos, no

entanto, afirmar, que ambas as metodologias bio-inspiradas se mostraram bastante adequadas ao problema de separação formulado com base em estatísticas de ordem e numa medida de informação mútua. Tal constatação e as conseqüências por ela trazidas no sentido de que métodos desse tipo sejam considerados com atenção em domínios não-lineares e/ou relacionados à teoria de informação são, em nossa opinião, aspectos dignos de nota do esforço aqui reportado.

TABELA IV  
EQM – Segundo Cenário

Quantidade de Simulações	EQM ( $\times 10^{-2}$ )	$y_1$	$y_2$	$y_3$
25	Clonalg	0,42	0,54	0,77
	Swarm	2,12	0,95	0,66
50	Clonalg	0,97	4,10	4,69
	Swarm	23,71	18,62	32,18

## VI. CONCLUSÕES

Neste trabalho, analisou-se a aplicabilidade de duas propostas pertencentes à área de computação natural no âmbito do problema de separação de misturas com não-linearidade posterior. Os resultados obtidos em dois cenários diferentes e sob duas metodologias experimentais distintas revelaram que os métodos, particularmente o algoritmo imunológico CLONALG, são ferramentas de busca robustas e eficientes no que se refere ao problema em questão. Mostrou-se ainda que os dois algoritmos têm desempenho equivalente ou superior ao da metodologia bio-inspirada proposta em [3]. Os resultados obtidos indicam que algoritmos desse tipo são opções promissoras quando se lida com processamento não-linear e baseado em teoria de informação.

## REFERÊNCIAS

- [1] Hyvärinen, A., Karhunen, J., Oja, E., *Independent Component Analysis*, John Wiley & Sons (2001).
- [2] Taleb, A., Jutten, C., "Source Separation in Postnonlinear Mixtures", *IEEE Trans. Signal Processing*, Vol. 47, No. 10, pp. 2807-2820 (1999).
- [3] Duarte, L.T., Suyama, R., Attux, R., Von Zuben, F.J., Romano, J.M.T., "Blind Source Separation of Post-Nonlinear Mixtures Using Evolutionary Computation and Order Statistics", *Springer Lecture Notes in Computer Science*, vol. 3889, pp. 66-73 (2006).
- [4] de Castro, L. N., Timmis, J., "An Artificial Immune Network for Multimodal Function Optimization", *Proceedings of the IEEE Congress on Evolutionary Computation, EUA* (2002).
- [5] Comon, P., "Independent Component Analysis, a New Concept?", *Signal Processing*, Vol. 36, No. 3, pp. 287-314 (1994).
- [6] Pham, D.-T., "Blind Separation of Instantaneous Mixtures of Sources Based on Order Statistics", *IEEE Trans. Signal Processing*, Vol. 48, No. 2, pp. 363-375 (2000).
- [7] de Castro, L. N., *Fundamentals of Natural Computing: Basic Concepts, Algorithms and Applications*, CRC Press (2006).
- [8] de Castro, L. N., Von Zuben, F. J., "Learning and Optimization Using the Clonal Selection Principle", *IEEE Trans. on Evolutionary Computation*, Vol. 6, No. 3, pp. 239-251 (2002).
- [9] Kennedy, J., Eberhart, R. C., "Particle Swarm Optimization", *Proceedings of the IEEE International Conference on Neural Networks, Austrália* (1995).