
UM ESTUDO DA APLICAÇÃO DE ALGORITMOS BIO-INSPIRADOS AO PROBLEMA DE ESTIMAÇÃO DE DIREÇÃO DE CHEGADA

Levy Boccato*

lboccato@dca.fee.unicamp.br

Romis Ribeiro de Faissol Attux*

attux@dca.fee.unicamp.br

Rafael Krummenauer†

rkrummen@decom.fee.unicamp.br

Mauro Lopes†

mauro@decom.fee.unicamp.br

*Depto. de Engenharia de Computação e Automação Industrial, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, C. P. 6101, 13081-970, Campinas, SP - Brasil

†Depto. de Comunicações, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, C. P. 6101, 13081-970, Campinas, SP - Brasil

ABSTRACT

A study on the Application of Bio-inspired Algorithms to the Problem of Direction of Arrival Estimation

The classical solution to the problem of estimating the direction of arrival (DOA) of plane waves impinging on a sensor array is based on the application of the maximum likelihood method. This approach leads to the problem of optimizing a cost function which is non-linear, non-quadratic, multimodal and variant with respect to the signal-noise ratio (SNR). The methods proposed in the literature to solve this problem fail for a wide set of SNR values. This work presents the results obtained from a study on the application of natural computing algorithms to the DOA estimation problem. Computational simulations show that four of the analyzed algorithms find the global optimum for a broad range of SNR values with computational efforts lower than that associated with an exhaustive search.

KEYWORDS: Natural computing, direction of arrival, estimation, maximum likelihood, bio-inspired algorithms.

RESUMO

A solução clássica para o problema de estimação dos ângulos de chegada (DOA) de sinais incidindo em um arranjo de sensores é a aplicação do método de máxima verossimilhança. Este método leva ao problema de otimização de uma função custo não-linear, não-quadrática, multimodal e variante com a relação sinal-ruído (SNR). Os métodos propostos para tal tarefa, presentes na literatura, falham em uma ampla gama de valores de SNR. Este trabalho apresenta os resultados de um estudo sobre a aplicação de ferramentas pertencentes à computação natural ao problema de estimação DOA. Simulações demonstram que quatro dos algoritmos analisados alcançam o ótimo global para uma ampla faixa de valores de SNR, com esforços computacionais inferiores àquele exigido por uma busca exaustiva.

PALAVRAS-CHAVE: Computação natural, direção de chegada, estimação, máxima verossimilhança, ferramentas bio-inspiradas.

1 INTRODUÇÃO

Um arranjo é um conjunto de sensores (transdutores) dispostos segundo uma geometria no espaço tridimensional, com um ponto de referência em comum (Manikas, 2004). Tal estrutura tem sido empregada em uma série de aplica-

Artigo submetido em 27/11/2008 (Id.: 00926)

Revisado em 13/02/2009, 03/06/2009, 23/06/2009

Aceito sob recomendação do Editor Associado Prof. Ivan Nunes Da Silva

ções, dentre as quais destacam-se os sistemas de comunicações sem fio, sistemas de radar e sonar, além de sismologia (Haykin, 1985). A função dos sensores é coletar amostras da excitação que o ambiente exerce sobre o arranjo para que seja possível extrair informações espaço-temporais de sinais corrompidos por ruído. Uma informação de interesse direto é a localização da fonte emissora, a qual é medida através do ângulo de chegada ou DOA (*Direction Of Arrival*) (Van Trees, 2001).

Um dos principais métodos de estimação paramétrica utilizados no problema de estimação DOA é baseado no critério de máxima verossimilhança (em inglês, *Maximum Likelihood* (ML)). Segundo este método, a estimativa do parâmetro é obtida a partir da maximização da função de verossimilhança, a qual, no caso do problema de DOA, possui um caráter não-linear, não-quadrático e multimodal, especialmente em condições críticas de relação sinal-ruído (em inglês, *signal-noise ratio* (SNR)) (Van Trees, 2001) (Krummenauer, 2007). Como agravante, a superfície desta função muda consideravelmente à medida que alteramos a SNR, de forma que as posições de todos os pontos modais são alteradas. Uma vez que a SNR não é conhecida no problema DOA, qualquer método de busca deve ser eficaz neste contexto de variações desconhecidas e ainda para uma ampla gama de valores de SNR.

Dessa forma, é necessário, a rigor, realizar uma busca exaustiva¹ para se encontrar o ótimo global. No entanto, tal proposta é inviável devido a seu elevado custo computacional, especialmente quando o número de sinais incidentes no arranjo aumenta. Em virtude disto, algumas abordagens alternativas foram concebidas, dentre as quais destacam-se MODE (Stoica e Sharman, 1990), MODEX (Gershman e Stoica, 1999) e *modified* MODEX (Lopes et al., 2003). Não obstante, nenhuma delas consegue atingir os resultados oferecidos pela busca exaustiva.

Neste trabalho, propomos o uso de ferramentas pertencentes à computação natural como alternativa aos métodos clássicos de estimação DOA. Tais ferramentas são particularmente interessantes nos problemas para os quais os métodos matemáticos que garantem a obtenção da solução ótima são demasiadamente complexos.

Dentre os principais ramos da computação natural no contexto de otimização multimodal, a saber, a computação evolutiva, os sistemas imunológicos artificiais e a inteligência de enxame (de Castro, 2006), selecionamos as seguintes ferramentas bio-inspiradas: algoritmos genéticos (Holland, 1992), os métodos de *fitness sharing* (Sareni e Krahenbuhl, 1998) e *clearing* (Pétrowski, 1996), o algoritmo CLONALG (de Castro e Von Zuben, 2002) e a rede *opt-*

aiNet (de Castro e Timmis, 2002), além do algoritmo *particle swarm* (Kennedy e Eberhart, 1995).

Este conjunto representativo de ferramentas foi aplicado ao problema de estimação DOA e cada técnica teve seu desempenho avaliado. Um estudo comparativo foi realizado evidenciando as potencialidades e limitações de cada método e foram identificados alguns aspectos que certamente necessitam ser aprofundados em trabalhos futuros. Como breve resumo das conclusões deste estudo, constatou-se que as variações da superfície da função de verossimilhança com a SNR dificultam sobremaneira a escolha dos parâmetros dos algoritmos bio-inspirados, indicando que este tópico demanda estudos extensivos. Mesmo assim, foi possível estabelecer uma primeira seleção que capacitou os algoritmos *particle swarm*, *clearing*, CLONALG e *opt-aiNet* a localizarem o ótimo global da função de verossimilhança de modo a reproduzir o desempenho da busca exaustiva.

Na próxima seção, apresentamos os modelos de sinal e ruído utilizados para descrever o problema de DOA, assim como as principais suposições estatísticas. Na Seção 3, apresentamos o estimador ML para o problema de DOA. Em seguida, destacamos o caráter multimodal da função custo associada ao estimador, a qual necessita ser minimizada para a determinação das estimativas dos ângulos de chegada.

Os principais conceitos referentes aos algoritmos bio-inspirados empregados neste trabalho são apresentados na Seção 4, bem como as particularidades da aplicação de cada algoritmo ao problema de estimação DOA e o procedimento adotado para a adaptação dos principais parâmetros. Na Seção 5, é definido o cenário de teste utilizado na avaliação de desempenho das ferramentas bio-inspiradas, enquanto os resultados obtidos por cada algoritmo e as devidas análises se encontram na Seção 6. Seguem-se, por fim, conclusões, agradecimentos e referências bibliográficas.

2 MODELO DE SINAL

Suponha M ondas planas de banda estreita incidindo em um arranjo linear uniforme de N ($N > M$) sensores. Os sensores estão igualmente espaçados com $d = \lambda/2$, onde λ é o comprimento de onda. As ondas incidem com ângulos $\phi = [\phi_1 \dots \phi_M]^T$ ($[\cdot]^T$ denota o transposto), medidos em relação ao vetor normal ao eixo do arranjo. A Figura 1 apresenta a geometria do arranjo e, em destaque, a m -ésima onda incidente.

O problema de estimação DOA pode ser descrito pelo seguinte modelo matemático (Krim e Viberg, 1996):

$$\mathbf{y}(k) = \mathbf{A}\mathbf{x}(k) + \boldsymbol{\eta}(k), \quad k = 1, 2, \dots, K \quad (1)$$

onde $\mathbf{y}(k) = \mathbf{y}(t_k) \in \mathcal{C}^{N \times 1}$ representa o sinal recebido e

¹Salientamos que a busca exaustiva corresponde a uma varredura do espaço de busca usando uma grade uniforme.

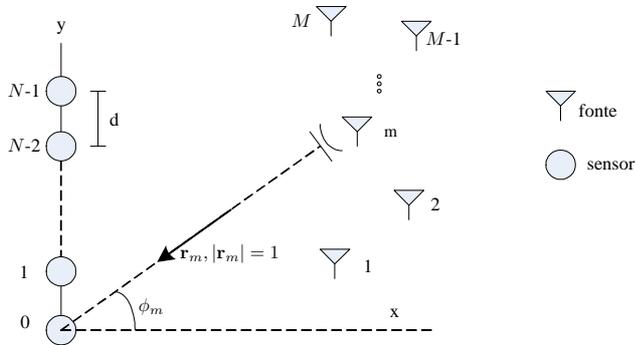


Figura 1: Arranjo linear uniforme de N sensores.

amostrado no instante t_k (*snapshot*), $\mathbf{x}(k) \in \mathbb{C}^{M \times 1}$ representa as amplitudes dos sinais recebidos, $\boldsymbol{\eta}(k) \in \mathbb{C}^{N \times 1}$ representa as amostras do ruído aditivo, K é o número total de *snapshots* considerados, e $\mathbf{A} \in \mathbb{C}^{N \times M}$ é a matriz de resposta do arranjo, definida pela estrutura de Vandermonde

$$\mathbf{A} = \begin{bmatrix} 1 & \dots & 1 \\ e^{j\omega_1} & \dots & e^{j\omega_M} \\ \vdots & & \vdots \\ e^{j(N-1)\omega_1} & \dots & e^{j(N-1)\omega_M} \end{bmatrix}, \quad (2)$$

sendo que $\omega_m = \pi \sin \phi_m$, $m = 1, \dots, M$. O objetivo é estimar os ângulos de chegada ϕ_m .

Admitimos que o número M de fontes é conhecido. Consideramos também que a matriz \mathbf{A} possui posto cheio em colunas, implicando em $N \geq M$. Outras hipóteses para o modelo da expressão (1) são apresentadas a seguir:

- 1: O ruído $\boldsymbol{\eta}(k)$ é um processo aleatório circular complexo Gaussiano estacionário de média zero. Além disso, é temporal e espacialmente decorrelacionado, com o momento de segunda-ordem

$$E\{\boldsymbol{\eta}(k)\boldsymbol{\eta}^H(i)\} = \sigma^2 \mathbf{I} \delta_{k,i}, \quad (3)$$

onde $\delta_{k,i}$ é o operador delta de Kronecker, $E\{\cdot\}$ é o operador de média estatística ou esperança, $\{\cdot\}^H$ denota conjugado transposto de uma matriz ou vetor, e \mathbf{I} denota a matriz identidade.

- 2: O conjunto dos vetores de sinal $\mathbf{x}(k)$ (para $k = 1, 2, \dots, K$) corresponde a um processo aleatório circular complexo Gaussiano estacionário de média zero, decorrelacionado temporalmente e com matriz de covariância não-singular. Assim

$$E\{\mathbf{x}(k)\mathbf{x}^H(i)\} = \mathbf{C} \delta_{k,i}, \quad (4)$$

- 3: Os vetores de sinais $\mathbf{x}(k)$ e de ruído $\boldsymbol{\eta}(i)$ são decorrelacionados para todo k e i .

É possível trabalhar com dois modelos para as fontes de sinal (Stoica e Nehorai, 1990): modelo condicional (CM), o qual supõe que a sequência $\{\mathbf{x}(k)\}_{k=1}^K$ será sempre a mesma para todas as possíveis realizações dos dados $\{\mathbf{y}(k)\}_{k=1}^K$, e o modelo incondicional (UM), que admite $\{\mathbf{x}(k)\}_{k=1}^K$ variável de realização para realização. Tanto o modelo CM quanto UM assumem que $\{\boldsymbol{\eta}(k)\}_{k=1}^K$ varia de realização para realização.

Se considerarmos o modelo CM, a distribuição dos dados $\{\mathbf{y}(k)\}_{k=1}^K$ é dada por $\mathbf{y}(k) \sim \mathcal{N}(\mathbf{A}\mathbf{x}(k), \sigma^2 \mathbf{I})$, enquanto que para o modelo UM, temos $\mathbf{y}(k) \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_y)$, onde

$$\begin{aligned} \mathbf{R}_y &= E\{\mathbf{y}_k \mathbf{y}_k^H\} \\ &= \mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma^2 \mathbf{I}. \end{aligned}$$

Para um número finito de *snapshots*, a matriz de covariância teórica \mathbf{R}_y pode ser aproximada pela média amostral:

$$\hat{\mathbf{R}}_y = \frac{1}{K} \sum_{k=1}^K \mathbf{y}_k \mathbf{y}_k^H. \quad (5)$$

3 ESTIMADOR DE MÁXIMA VEROSSIMILHANÇA

O estimador ML é obtido por meio da maximização da função de verossimilhança (Kay, 1993). Em (Van Trees, 2001) e (Krummenauer, 2007), está demonstrado que, tendo em conta as hipóteses aqui adotadas, o estimador ML para o problema de DOA usando o modelo CM pode ser escrito como²:

$$\hat{\boldsymbol{\theta}}_{ML} = \arg \min_{\boldsymbol{\theta}} \text{Tr} \left\{ \mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{R}}_y \right\}, \quad (6)$$

onde $\text{Tr}\{\cdot\}$ indica o traço de uma matriz, $\mathbf{P}_{\mathbf{A}}^{\perp} = \mathbf{I} - \mathbf{A}\mathbf{A}^{\perp}$ é a matriz de projeção do subespaço de ruído, $\mathbf{P}_{\mathbf{A}} = \mathbf{A}\mathbf{A}^{\perp}$ é a matriz de projeção do subespaço de sinal e $\mathbf{A}^{\perp} = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$ é a pseudo-inversa de \mathbf{A} .

A função custo

$$J_{ML}(\boldsymbol{\theta}) = \text{Tr} \left\{ \mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{R}}_y \right\} \quad (7)$$

é não-linear, não-quadrática e possui diversos mínimos locais. Com o propósito de verificarmos as características desafiadoras do problema de minimização da função custo J_{ML} , apresentamos na Figura 2 as superfícies e curvas de nível

²Note que a variável $\boldsymbol{\theta}$ denota genericamente o parâmetro de direção de chegada, enquanto que ϕ representa os valores verdadeiros dos ângulos de chegada.

dessa função. As curvas apresentadas na Figura 2 foram geradas considerando duas fontes de faixa estreita e de amplitudes unitárias incidindo no arranjo linear uniforme de 10 sensores com ângulos de chegada $\phi_1 = 10^\circ$ e $\phi_2 = 15^\circ$ relativos à normal ao eixo do arranjo, para relações sinal-ruído de 0 dB e -15 dB. Nas duas situações, foram utilizados 100 *snapshots*. A matriz de covariância do sinal é dada por $\mathbf{C} = \mathbf{I}$, e a relação sinal-ruído, em dB, é definida como $\text{SNR} = 10 \log 1/\sigma^2$, onde σ^2 corresponde à variância do ruído.

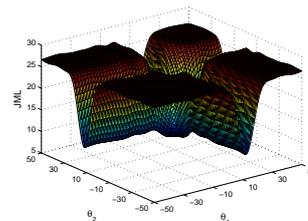
A Figura 2 evidencia o impacto da relação sinal-ruído na superfície da função custo J_{ML} . Para uma relação sinal-ruído de 0 dB (Figura 2(a)), a superfície tem dois mínimos globais simétricos (localizados aproximadamente em $(10^\circ, 15^\circ)$ e $(15^\circ, 10^\circ)$) e um caráter multimodal relativamente pouco pronunciado. No entanto, quando trabalhamos com -15 dB (Figura 2(c)), a superfície passa a apresentar diversos mínimos locais de amplitudes próximas à do mínimo global. Além disso, conforme reduzimos a relação sinal-ruído, o ponto de mínimo global se desloca, de modo que suas coordenadas não mais coincidem com os verdadeiros ângulos de chegada. Estas duas evidências servem para ilustrar os desafios que o estimador ML enfrenta no problema de DOA.

4 COMPUTAÇÃO NATURAL

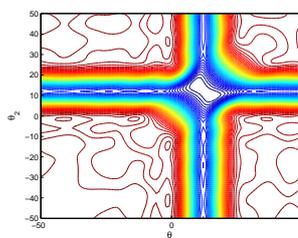
Conforme destacado na Seção 3, a estimativa ML só pode ser garantidamente determinada através de uma busca exaustiva pelo mínimo global da função custo J_{ML} . Porém, tal abordagem é computacionalmente inviável. Em virtude disto, métodos como MODE e MODEX foram propostos como alternativas à busca exaustiva.

O algoritmo MODE (Stoica e Sharman, 1990) emprega uma reparametrização do estimador de máxima verossimilhança original. Esta reparametrização gera um problema baseado no subespaço de sinal a partir da matriz de covariância dos *snapshots*. Sua solução, através de um algoritmo quadrático iterativo, produz um polinômio cujas raízes são as estimativas dos parâmetros DOA. O algoritmo MODEX (Gershman e Stoica, 1999) baseia-se no método MODE e apresenta uma performance superior para pequenos valores de relação sinal-ruído. Esta melhoria de desempenho é atingida através da geração de raízes extras e de um procedimento de máxima verossimilhança que seleciona a melhor combinação destas raízes como estimativas DOA. Todavia, é importante ressaltar que tais métodos não oferecem garantia de convergência para o mínimo global e seu esforço computacional cresce exponencialmente com o aumento de sensores e fontes.

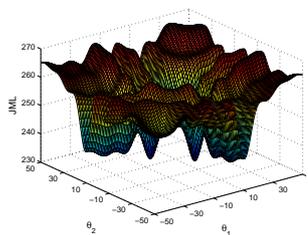
Seguindo um paradigma diferente, os algoritmos bioinspirados também constituem alternativas à busca exaustiva. De fato, ferramentas pertencentes à computação natural vêm



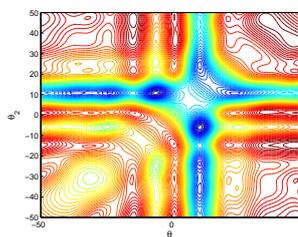
(a) SNR = 0 dB



(b) SNR = 0 dB



(c) SNR = -15 dB



(d) SNR = -15 dB

Figura 2: Análise das superfícies e curvas de nível da função ML Condicional para $M = 2$, $N = 10$, 100 *snapshots* e relações sinal-ruído de 0dB e -15 dB.

sendo aplicadas com notável sucesso na busca de máximos ou mínimos de funções multimodais, inclusive no contexto de processamento de sinais (Attux et al., 2003).

Uma característica particular do problema de estimação DOA com a qual as ferramentas bio-inspiradas terão de lidar é que a superfície da função custo J_{ML} é diferente para cada valor de SNR, conforme mostrado na Seção 3. Além disso, esta também se altera de experimento para experimento, principalmente para valores pequenos de SNR. Ou seja, considerando um único cenário de teste, os algoritmos de computação natural se deparam com diversos problemas de otimização muito distintos. Por outro lado, os algoritmos aqui tratados devem apresentar um ajuste fixo de seus parâmetros, independentemente do valor da SNR, uma vez que esta grandeza é desconhecida e que as ferramentas empregadas não possuem a habilidade de se ajustarem automaticamente a cada situação. Estas condições constituem grandes desafios para a aplicação destas ferramentas neste problema.

Nas Seções 4.1 a 4.4, apresentamos as ferramentas bio-inspiradas utilizadas neste trabalho e discutimos as particularidades da aplicação destas técnicas ao problema de estimação DOA. Quando possível, os valores estabelecidos para os parâmetros de cada algoritmo são mostrados. Do contrário, deixamos sua definição para a Seção 6.

4.1 Algoritmos genéticos

Algoritmos Genéticos (AGs) são técnicas de otimização inspiradas nos princípios fundamentais da moderna teoria da evolução (Bäck et al., 1997) (Goldberg, 1989). O funcionamento de tais algoritmos tem por base a sinergia entre mecanismos de recombinação e mutação e esquemas capazes de contemplar o efeito do meio sobre a seleção de indivíduos. Essa dinâmica peculiar confere aos AGs a possibilidade de realizar eficientemente tarefas de otimização em diversos contextos (Silva et al., 2007). Por exemplo, em (Coelho e Mariani, 2006), alguns algoritmos evolutivos, incluindo os AGs, se mostraram promissores quando utilizados na identificação e controle de processos industriais.

Os AGs operam sobre uma população de indivíduos, os quais representam soluções potenciais para o problema tratado. A cada geração, as soluções promissoras tendem a se reproduzir, enquanto que soluções relativamente “ruins” tendem a ser eliminadas. A medida da qualidade de uma solução candidata é fornecida por uma função de avaliação ou adaptabilidade (*fitness*), que simula o efeito da pressão exercida pelo ambiente sobre o indivíduo.

Seja $P(j) = \{x_1^j, \dots, x_n^j\}$ a população de soluções candidatas (indivíduos) na geração j . Em síntese, um algoritmo genético, após a definição de uma população inicial, realiza

as seguintes etapas em cada geração:

- cada solução x_i^j é avaliada através de uma função de *fitness* e tem-se uma medida de sua adaptação;
- alguns membros da população são selecionados para recombinação (*crossover*) e mutação, gerando novas soluções potenciais;
- através de um mecanismo de seleção, uma nova população é formada privilegiando a participação dos indivíduos mais adaptados.

Este processo se repete até que um número pré-determinado de gerações seja atingido ou até que um nível de adaptação esperado seja alcançado.

O projeto de um algoritmo genético para a solução de um problema em particular deve considerar os seguintes aspectos: representação genética para os indivíduos, definição da população inicial, função de avaliação (*fitness*) dos indivíduos, operadores genéticos e definição de grandezas numéricas como tamanho da população e probabilidades de aplicação dos operadores genéticos.

4.1.1 Algoritmos genéticos aplicados ao problema de estimação DOA

Adotamos a codificação real por ser a mais intuitiva para representar as estimativas dos ângulos de chegada. A construção da população inicial é feita de maneira aleatória, de modo que as estimativas dos ângulos são distribuídas uniformemente no intervalo $(-90^\circ, 90^\circ)$. Dentre os diversos operadores genéticos de mutação e *crossover* propostos na literatura, optamos por utilizar os mais difundidos no âmbito da codificação real: mutação gaussiana, com desvio padrão unitário, e *crossover* aritmético de dois indivíduos. Após algumas simulações, adotamos $p_m = 0,01$ e $p_c = 0,80$, sendo p_m e p_c as probabilidades de ocorrência dos operadores de mutação e *crossover*, respectivamente.

Quanto ao operador de seleção, utilizamos o método de seleção por torneio devido à facilidade que o mesmo oferece para ajustarmos a pressão seletiva através da definição do número de torneios (T_o) e do número de competidores (q). A esta justificativa se acrescenta o fato de, no problema de estimação DOA, estarmos interessados em manter a diversidade da população na tentativa de escapar das soluções sub-ótimas existentes, principalmente em situações críticas de relação sinal-ruído. Além disto, utilizamos uma abordagem salcionista que preserva o indivíduo com maior *fitness* para a próxima geração.

Em geral, os operadores genéticos trabalham no sentido de privilegiar os indivíduos de maior *fitness*. Por isso, é neces-

sário converter o problema de minimização da função custo J_{ML} em um problema de maximização de uma função de avaliação (*fitness*). Dentre as diversas formas de realizar esta conversão, definimos uma transformação monotônica e inversamente proporcional ao valor da função custo J_{ML} , expressa por $F_{fit} = \frac{1}{1+J_{ML}}$. Com isto, o algoritmo genético tentará localizar o máximo global de F_{fit} , que ocorre justamente quando J_{ML} é mínima.

Por fim, definimos o critério de parada como um número máximo de gerações (N_{gen}). Evidentemente existem outras opções mais elaboradas, como critérios baseados em medidas de estagnação do algoritmo. Entretanto, a primeira opção oferece a vantagem de não interferir diretamente no desempenho do algoritmo, já que o número máximo pode sempre ser alterado caso seu valor seja inadequado. Isto facilita o processo de adaptação de cada ferramenta, pois as possíveis falhas de desempenho verificadas em sua execução estão relacionadas ao conjunto de parâmetros e operadores escolhidos, e não ao critério de parada.

Deve ficar claro que o número máximo de gerações a ser escolhido deve ser suficiente para permitir que a população evolua satisfatoriamente, mas também não deve ser superestimado, dado que quanto maior sua magnitude, maior o esforço computacional requerido e o tempo de execução do algoritmo.

4.2 Técnicas de niching

Os algoritmos genéticos clássicos são ferramentas poderosas no contexto de otimização de funções. Entretanto, experimentos e análises mostram que estas técnicas não são particularmente eficientes em localizar os múltiplos ótimos de uma superfície multimodal, uma vez que não estão adaptadas para preservar outros pontos ótimos durante a busca (Mahfoud, 1995).

Inspirados no conceito de nichos ecológicos, os quais representam regiões do ambiente que suportam diferentes tipos de vida compartilhando os recursos disponíveis, vários métodos foram desenvolvidos com o propósito de manter a diversidade na população e permitir a investigação de múltiplos ótimos de modo paralelo. Tais mecanismos, conhecidos como métodos de *niching*, criam e mantêm subpopulações nas vizinhanças das soluções ótimas (Mahfoud, 1995).

Um nicho é caracterizado por uma porção limitada de recursos disponíveis para os indivíduos que nele se encontram. Cada indivíduo em um nicho tem direito a uma fração dos recursos disponíveis - quanto maior o tamanho da subpopulação, menor a fração. No contexto de otimização de funções multimodais, os recursos disponíveis em um nicho estão associados à medida de *fitness* do indivíduo. As técnicas de

niching atuam sobre esta medida com a intenção de impedir que todos convirjam para a mesma região do espaço de busca, mantendo assim a diversidade da população.

O método de *fitness sharing* (Goldberg e Richardson, 1987) reduz o *fitness* f_i dos indivíduos da população por um fator m_i , de modo que a nova medida de *fitness* f_i^* é dada por $f_i^* = f_i/m_i$. No cálculo de m_i , dois parâmetros são levados em consideração: o limiar de similaridade (σ_s), que determina o raio de cada nicho, e a distância entre dois indivíduos, de forma que m_i corresponde, aproximadamente, ao número de indivíduos com os quais o *fitness* f_i é compartilhado (Sareni e Krahenbuhl, 1998). Esta técnica deve ser usada em conjunto com operadores de seleção de baixa pressão seletiva e o operador de *crossover* deve ser construído a fim de evitar o cruzamento entre indivíduos que pertencem a nichos diferentes.

Apesar de melhorar o desempenho dos algoritmos genéticos em problemas de otimização multimodal e de conseguir manter a diversidade da população, o método de *fitness sharing* apresenta algumas limitações, principalmente por causa do parâmetro σ_s . A escolha do valor de σ_s requer um conhecimento *a priori* da distância aproximada entre os ótimos, o que muitas vezes não está disponível. Além disso, σ_s é, tipicamente, o mesmo para todos os nichos, de forma que supõe-se que os ótimos são aproximadamente equidistantes no espaço de busca. Por estas razões, o método de *fitness sharing* pode não funcionar adequadamente.

O método de *clearing* (Pétrowski, 1996) difere do de *fitness sharing* na medida em que os recursos de um determinado nicho não são compartilhados, mas totalmente atribuídos aos melhores indivíduos. Basicamente, o método de *clearing* preserva o *fitness* dos melhores indivíduos de um nicho enquanto zera o *fitness* de todos os outros indivíduos da mesma subpopulação. O número máximo de indivíduos aceitos em cada nicho é denominado capacidade do nicho (κ).

Assim como no caso de *fitness sharing*, uma medida de similaridade entre os indivíduos é utilizada para determinar se estes pertencem à mesma subpopulação. Considera-se que dois indivíduos pertencem ao mesmo nicho se esta medida de similaridade (usualmente dada pela distância entre os indivíduos) é menor do que um limiar σ_s . Novamente, a definição do valor de σ_s requer um conhecimento *a priori* da distância entre os ótimos e supõe, em geral, que os mesmos são aproximadamente equidistantes.

4.2.1 Fitness sharing aplicado ao problema de estimativa DOA

Adotamos para o método de *fitness sharing* as mesmas opções usadas no caso do algoritmo genético no que se refere

à codificação, à população inicial, à função de *fitness* e aos operadores genéticos, agora com probabilidades $p_m = 0,01$ e $p_c = 1,00$. Adicionalmente, utilizamos uma abordagem elitista que preserva os indivíduos com *fitness* maior que a média para a próxima geração (Sareni e Krahenbuhl, 1998).

O principal parâmetro, e também o mais crítico do método de *fitness sharing*, é o limiar de similaridade (σ_s). Seja σ_s^* o valor ideal do limiar de similaridade para um problema de otimização multimodal. Se o valor escolhido de σ_s for muito menor que σ_s^* , haverá a formação de um número maior de nichos, já que pontos pertencentes ao mesmo pico da superfície a ser otimizada serão interpretados como pertencentes a nichos distintos. Porém, permite-se uma maior concentração de pontos numa mesma região do espaço e, como consequência disto, aumentamos as chances de convergência para pontos sub-ótimos e perda de diversidade.

Por outro lado, no caso em que σ_s for muito maior que σ_s^* , pontos explorando diferentes regiões do espaço de busca (possivelmente picos) serão interpretados como pertencentes ao mesmo nicho e terão que compartilhar os recursos, de modo que sofrerão uma redução no *fitness* proporcional ao número de indivíduos existentes no nicho. Isto também pode comprometer o desempenho do algoritmo, pois aumentamos a probabilidade de descobrir e destruir nichos ao longo das gerações, e, portanto, podemos perder o nicho que englobava a solução ótima (Cioppa et al., 2004).

Destacamos que algumas peculiaridades do problema de estimação DOA tornaram a tarefa de determinar o limiar de similaridade ainda mais difícil. Conforme mostrado na Seção 3, a superfície da função custo e, conseqüentemente, da função de *fitness*, é alterada drasticamente com a mudança da relação sinal-ruído, de modo que o número de picos e a distância entre eles variam radicalmente. Uma vez que temos de definir um único valor de σ_s independentemente da SNR, é evidente que, em algumas circunstâncias, esta escolha será inadequada e comprometerá o desempenho do algoritmo. Diante disto, buscamos uma métrica de compromisso entre todos os fatores citados e, após diversas simulações para diferentes SNRs, arbitramos $\sigma_s = 0,75$.

Uma dificuldade adicional encontrada pelo método de *fitness sharing* se deve à própria natureza da superfície da função de *fitness*. Por exemplo, para uma SNR de 15 dB, se o ótimo global apresenta *fitness* igual a A , as regiões menos promissoras apresentam aproximadamente *fitness* igual a $A/2$. Desta maneira, quando houver 3 indivíduos compartilhando recursos na região do ótimo global, o novo valor de *fitness* de cada um deles é, aproximadamente, $A/3$, que é inferior ao *fitness* das soluções sub-ótimas. Com isto, pode ocorrer a perda dos indivíduos que mais se aproximam do ótimo global e a manutenção daqueles situados em regiões nada promissoras.

soras.

Algumas alternativas para a função de *fitness* foram testadas, mas estas também apresentaram o mesmo problema para certos valores de SNR. Sendo assim, é necessário conduzir uma investigação acerca de outras propostas para a função de *fitness* que atenuem este efeito indesejável ou então utilizar um esquema de *fitness scaling* (Darwen e Yao, 1995). Por fim, definimos o critério de parada como um número máximo de iterações (N_{gen}).

4.2.2 Clearing aplicado ao problema de estimação DOA

Assim como no caso do algoritmo genético e do método de *fitness sharing*, utilizamos no método de *clearing*: codificação real, distribuição uniforme da população inicial, $F_{fit} = \frac{1}{1+J_{ML}}$, mutação gaussiana com $p_m = 0,02$, *crossover* aritmético com $p_c = 1,00$ e seleção por torneio.

Adotamos $\kappa = 1$ para o número de indivíduos de cada nicho, uma vez que também utilizamos uma abordagem elitista, a qual preserva os indivíduos com *fitness* maior que o *fitness* médio (Pétrowski, 1996). Conforme discutido na Seção 4.2.1, o ajuste de σ_s não é uma tarefa fácil, e novamente o fizemos de maneira heurística. Após algumas simulações, estabelecemos $\sigma_s = 1,40$. Por fim, o critério de parada escolhido foi um número máximo de iterações (N_{gen}).

4.3 Sistemas imunológicos artificiais

A classe dos sistemas imunológicos artificiais (SIAs) compreende técnicas cujo *modus operandi* se vincula a elementos conceituais derivados de teorias que versam sobre o funcionamento do sistema de defesa dos mamíferos. Um dos principais elementos teóricos desenvolvidos no estudo dos sistemas imunológicos é o princípio de seleção clonal (Ada e Nossal, 1987), o qual é utilizado para explicar os principais aspectos de uma resposta imunológica adaptável a um estímulo antigênico. A idéia é que apenas as células que conseguem reconhecer os antígenos são selecionadas para proliferação e, em seguida, submetidas ao processo de maturação de afinidade.

O modelo computacional do processo de maturação de afinidade envolve a realização de duas tarefas: gerar clones dos anticorpos de maneira proporcional à afinidade com os antígenos e aplicar mutação sobre os clones com taxa inversamente proporcional à afinidade, de modo que os clones com maior afinidade tendem a sofrer mutações mais suaves e aqueles com menor afinidade tendem a sofrer mutações mais abruptas. A taxa de mutação de um clone do i -ésimo indivíduo da população é dada por $pm_i = e^{-\varphi f_i}$, onde φ determina a velocidade de queda da taxa de mutação e f_i é o

fitness do i -ésimo indivíduo normalizado no intervalo (0,1) .

O algoritmo CLONALG (de Castro e Von Zuben, 2002) une os princípios de seleção clonal e maturação de afinidade aplicando-os em problemas de reconhecimento de padrões (algo que nosso sistema imunológico tem de fazer constantemente) e otimização. A seguir, apresentamos uma descrição básica do algoritmo CLONALG adaptado para esta última tarefa:

1. os anticorpos (indivíduos da população) são codificados em cromossomos binários;
2. a própria função de *fitness* fornece a medida de afinidade anticorpo-antígeno;
3. selecionamos os n indivíduos de maior afinidade. Nos casos em que desejamos localizar múltiplos ótimos, usualmente fazemos n igual ao tamanho da população;
4. os n indivíduos selecionados são clonados independentemente e de forma proporcional à afinidade. Se desejamos localizar múltiplos ótimos, a clonagem proporcional à afinidade não é mais necessária e são criados $N_c = \lfloor \beta N \rfloor$ clones para cada um dos n indivíduos, onde β está associado à velocidade de convergência do algoritmo;
5. o conjunto de todos os clones é submetido ao processo de maturação de afinidade: quanto maior a afinidade, menor a taxa de mutação;
6. o melhor clone de cada um dos n indivíduos originais é selecionado;
7. finalmente, substituem-se os d indivíduos com menor afinidade por novos indivíduos gerados aleatoriamente.

Interessa-nos também uma rede imunológica artificial destinada à otimização de funções de variáveis reais - a *opt-aiNet* (de Castro e Timmis, 2002). Essa rede, que une as idéias de seleção clonal e maturação de afinidade (de Castro e Von Zuben, 2002) à noção de rede imunológica (Jerne, 1974), caracteriza-se por um elegante compromisso entre exploração (*exploration*) e exploração (*exploitation*) do espaço de busca, o que abre perspectivas interessantes em domínios multimodais.

A rede *opt-aiNet* realiza os mesmos processos de seleção clonal e maturação por afinidade que o algoritmo CLONALG. Além disso, trabalha com codificação real e introduz um controle dinâmico do tamanho da população através de mecanismos de supressão de indivíduos similares. Com isto, a diversidade da população é mantida e consegue-se explorar adequadamente o espaço de busca. Entretanto, a rede requer

o ajuste de dois parâmetros: o limiar de similaridade entre indivíduos σ_s (como nas técnicas de *niching* discutidas na Seção 4.2) e uma medida da variação de *fitness* médio entre as gerações, útil para apontar o momento em que as etapas de supressão e inserção de indivíduos devem ser realizadas.

4.3.1 CLONALG aplicado ao problema de estimação DOA

Optamos pelo uso do algoritmo em sua forma canônica (de Castro e Von Zuben, 2002), de modo que os indivíduos da população são representados por cromossomos binários. Para o problema de estimação DOA, uma vez que cada indivíduo deve representar uma solução completa, ou seja, deve explicitar os M ângulos de chegada pertencentes ao intervalo $(-90^\circ, 90^\circ)$, o cromossomo é composto de M trechos de comprimento l , sendo cada trecho referente a um ângulo de chegada.

O valor de cada ângulo de chegada pode ser obtido a partir de um vetor binário \mathbf{v} de comprimento l através da seguinte relação $\theta = \theta_{min} + V(\theta_{max} - \theta_{min})/(2^l - 1)$, onde $V = \sum_{i=1}^l \mathbf{v}(i)2^{(l-i)}$ corresponde ao número inteiro representado pelo vetor \mathbf{v} , θ_{min} representa o limitante inferior para o ângulo de chegada e θ_{max} é o limitante superior.

O comprimento da cadeia binária que representa um ângulo de chegada foi escolhido de forma a obtermos uma precisão numérica adequada no domínio real. O valor adotado foi $l = 30$. Adotamos, à semelhança dos algoritmos anteriores, um esquema de inicialização aleatória da população e a função de *fitness* dada por $F_{fit} = \frac{1}{1+J_{ML}}$.

A velocidade de queda da mutação, especificada pelo parâmetro φ , foi ajustada de maneira heurística tendo por base a seguinte constatação: se φ for muito elevado, a taxa de mutação, dada pela expressão $pm_i = e^{-\varphi F_{fit}^*(i)}$, onde $F_{fit}^*(i)$ é o *fitness* normalizado no intervalo (0,1) do i -ésimo indivíduo, tende a cair rapidamente conforme o *fitness* do indivíduo aumenta. Isto significa que os melhores indivíduos apresentarão taxas muito baixas de mutação, de forma que a exploração local realizada por seus clones será muito lenta e restrita, o que não é interessante, especialmente no início do processo de busca, quando os melhores indivíduos da população possivelmente ainda são soluções medianas ou pontos sub-ótimos. Por outro lado, se a velocidade de queda da mutação for muito pequena, a taxa de mutação tende a permanecer alta mesmo para os indivíduos com maior *fitness*. Isto também não é interessante, pois o progresso já alcançado pelo indivíduo pode não ser aproveitado por seus clones, uma vez que estes sofrem mutações bem acentuadas. Após alguns testes, concluímos que $\varphi = 1,75$ era um valor adequado.

Durante o processo de busca conduzido pelo algoritmo

CLONALG, a fim de manter a diversidade da população e ampliar a exploração do espaço de busca, é inserida, após um certo número de iterações, uma porcentagem de indivíduos aleatórios. Novamente certos limites foram considerados: se a porcentagem for muito grande, boa parte do progresso já atingido pelo algoritmo será perdida; por outro lado, se for muito pequena, a introdução de diversidade será bem reduzida e o algoritmo pode não conseguir explorar outras regiões do espaço de busca. Por meio de algumas simulações preliminares, concluímos que é possível manter um desempenho satisfatório se, a cada 5 iterações, forem gerados de forma aleatória 20% dos indivíduos da população. Por fim, o critério de parada adotado novamente foi um número máximo de iterações (N_{gen}).

4.3.2 Opt-aiNet aplicada ao problema de estimação DOA

Uma vez que a própria natureza da rede está plenamente adaptada à codificação real, decidimos representar a população desta maneira. A população inicial foi gerada aleatoriamente de modo similar ao apresentado na Seção 4.1.1, e a função de avaliação (*fitness*) foi construída de maneira semelhante à apresentada na Seção 4.3.1.

O passo seguinte envolveu o ajuste dos principais parâmetros da rede. Uma vez que a rede *opt-aiNet* é capaz de dinamicamente ajustar o tamanho da população através de mecanismos de inserção e poda, é preciso apenas definir o número inicial de indivíduos (N_i). Entretanto, devido a este comportamento dinâmico da rede, não é possível prever o número de vezes que a função de *fitness* será avaliada somente a partir do conjunto de parâmetros, como ocorre com os outros algoritmos bio-inspirados escolhidos para este trabalho. Por isso, para permitir uma comparação justa de desempenho e custo computacional, estabelecemos como critério de parada um número máximo de avaliações da função de *fitness*, de modo que o processo de evolução da rede é interrompido quando tal valor é atingido.

A velocidade de queda da taxa de mutação, definida pelo parâmetro φ , foi ajustada considerando-se os mesmos cuidados apresentados na Seção 4.3.1 e, por fim, adotamos $\varphi = 1$.

A rede *opt-aiNet* também define um limiar de similaridade entre os indivíduos (σ_s) e, à semelhança dos métodos de *nicheing*, devemos buscar uma solução de compromisso. Por meio de algumas simulações, definimos $\sigma_s = 1,20$.

As etapas de supressão e inserção de novos indivíduos são executadas pela rede *opt-aiNet* quando uma estagnação no processo de evolução é identificada. Essa medida de estagnação é obtida por meio do cômputo da variação de *fitness*

médio entre gerações de acordo com a expressão

$$\Delta Fit_m^j = 1 - \frac{Fit_m^j}{Fit_m^{j-1}},$$

onde j é a geração atual. Desta maneira, quando ΔFit_m é menor que um limiar δ_s , temos a evidência de que o processo está estagnado e por isso realizamos as etapas de supressão e inserção. Além disso, usualmente avaliamos a variação de *fitness* médio após um período de iterações T_i , pois é necessário também deixar que os indivíduos explorem o espaço de busca.

Alguns cuidados devem ser tomados durante a atribuição de valores aos parâmetros δ_s e T_i . Se δ_s for muito pequeno (ou T_i for muito grande), podemos retardar excessivamente as etapas de supressão e inserção e, conseqüentemente, manter a rede em situações de lento progresso. Por outro lado, um valor muito grande de δ_s (ou muito pequeno de T_i) faz com que o critério de variação mínima de *fitness* médio entre gerações seja freqüentemente satisfeito e, como conseqüência, as etapas de supressão e inserção sejam executadas com maior freqüência. Desta maneira, a população tende a crescer rapidamente, porém com pouco refinamento local, o que também pode não ser interessante.

Tendo isto em mente, realizamos algumas simulações e adotamos $T_i = 20$ e $\delta_s = 0,005$. Por fim, estabelecemos que o número de indivíduos aleatórios a serem introduzidos na população durante a etapa de inserção deve ser 30% do tamanho atual da população.

4.4 Particle Swarm

As técnicas de otimização baseadas em *particle swarm* (enxame de partículas) (Kennedy e Eberhart, 1995) surgiram a partir da idéia de reproduzir em computador a forma pela qual as sociedades coletivas processam o conhecimento (Kennedy, 1997). Assim como todas as outras abordagens de inteligência de enxame, o *particle swarm* baseia-se em uma população de indivíduos capazes de interagir entre si e com o meio ambiente.

O *particle swarm* utiliza os princípios de uma teoria sociocognitiva muito simples: 1) cada indivíduo de uma população possui sua própria experiência e é capaz de avaliar a qualidade desta experiência; 2) como há interação social entre os indivíduos, eles também possuem conhecimentos sobre os desempenhos de seus vizinhos.

Em termos matemáticos, os principais componentes do algoritmo *particle swarm* podem ser representados da seguinte forma:

- a posição de cada partícula (indivíduo) é dada por x_i no

espaço dos atributos \mathbb{R}^N ;

- a partícula se moverá com uma “velocidade vetorial” ν_i de modo que a próxima posição é dada por: $\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + \nu_i(t+1)$

O vetor velocidade ν_i indica a direção do movimento da partícula e pode ser calculado da seguinte forma:

$$\nu_i(t+1) = \nu_i(t) + \alpha(\mathbf{p}_i(t) - \mathbf{x}_i(t)) + \beta(\mathbf{p}_g(t) - \mathbf{x}_i(t)),$$

onde $\nu_i(t)$ corresponde à direção do último movimento da i -ésima partícula, $\mathbf{p}_i(t)$ representa a melhor posição já visitada pela i -ésima partícula, $\mathbf{p}_g(t)$ corresponde à melhor posição já visitada pelos vizinhos da i -ésima partícula e os parâmetros α e β são escalares limitados a um intervalo finito. A parcela $(\mathbf{p}_g(t) - \mathbf{x}_i(t))$ explicita a interação social entre a partícula e seus vizinhos ao passo que o termo $(\mathbf{p}_i(t) - \mathbf{x}_i(t))$ representa a parcela cognitiva.

A vizinhança de cada partícula corresponde a uma vizinhança topológica, e não à vizinhança no espaço dos atributos. A partir de um arranjo definido *a priori*, cada partícula terá acesso ao desempenho e posição de seus vizinhos estabelecidos pela topologia adotada. Existem diversas formas de se definir a vizinhança, dentre as quais destacamos a vizinhança em anel (cada partícula tem um vizinho à direita e à esquerda, fechando um anel) e a vizinhança totalmente conectada (todas as partículas são vizinhas entre si) (de Castro, 2006).

4.4.1 Particle Swarm aplicado ao problema de estimação DOA

Dado o comportamento da função custo J_{ML} em relação à SNR, conforme destacado na Seção 3, é interessante mantermos adequadamente a diversidade da população na tentativa de escapar de soluções sub-ótimas. Após algumas simulações preliminares, concluímos que a vizinhança em anel é adequada para realizar esta tarefa (de Castro, 2006).

Poderíamos utilizar a mesma função de *fitness* definida para os casos anteriores. Entretanto, adotamos uma outra transformação monotônica e inversamente dependente do valor da função custo J_{ML} , dada pela relação $F_{fit} = e^{-J_{ML}}$, uma vez que, considerando um conjunto de simulações preliminares, o algoritmo obteve um desempenho ligeiramente superior quando esta última função de *fitness* foi utilizada.

Um parâmetro ajustável do algoritmo *particle swarm* é o limite do vetor velocidade de cada partícula (lim_v). Se o valor de lim_v for muito pequeno, o tamanho do passo dado por cada partícula na direção de movimento estabelecida será menor e, portanto, a tendência é que a convergência seja mais lenta. Em contrapartida, se o valor de lim_v for muito

elevado, pode-se prejudicar a capacidade de refinamento das soluções quando estas se localizam na vizinhança dos pontos ótimos, dado que serão permitidos passos maiores na direção de movimento escolhida. Assim sendo, após algumas simulações preliminares, adotamos $|lim_v| = 1,5$.

A população inicial é gerada de modo aleatório com distribuição uniforme de cada atributo das partículas no intervalo $(-90^\circ, 90^\circ)$. Por fim, o critério de parada adotado foi um número máximo de iterações (N_{gen}).

5 CENÁRIO DE TESTE

O desempenho dos algoritmos bio-inspirados na estimação dos ângulos de chegada será analisado considerando-se o mesmo cenário usado para gerar as curvas da Figura 3. A única exceção diz respeito à relação sinal-ruído, a qual assumirá valores entre ± 15 dB. Enfatizamos que este cenário foi o escolhido por ser a referência de teste para métodos de estimação DOA e por estar presente nas diversas publicações da área, como em (Stoica e Sharman, 1990), (Gershman e Stoica, 1999), (Krim e Viberg, 1996) e (Lopes et al., 2003).

Conforme ressaltado na Seção 3, o mínimo global da superfície da função custo J_{ML} define a estimativa ML. Uma vez que trabalhamos com apenas 2 ângulos de chegada neste cenário, é possível realizarmos uma busca exaustiva bidimensional pelo ótimo global sem exceder a capacidade computacional disponível.

O método de busca exaustiva corresponde a uma varredura do espaço de busca com o auxílio de uma grade uniforme de precisão p . Uma vez definido o valor de p , o número total de pontos existentes na grade, que equivale ao número de vezes que a função custo considerada será avaliada, é expresso por

$$N_P = \left\lfloor \frac{x_{max} - x_{min}}{p} + 1 \right\rfloor^{dim},$$

onde dim corresponde à dimensão do espaço de busca e (x_{min}, x_{max}) representa a faixa de excursão das variáveis.

Os algoritmos bio-inspirados também devem buscar os mínimos da superfície da função custo J_{ML} e, portanto, tentam alcançar o mesmo desempenho da busca exaustiva. Entretanto, devem executar esta tarefa com um custo computacional menor. Caso isto não se verifique, a aplicação destas ferramentas no problema de estimação DOA não se justifica.

A principal medida de desempenho que utilizaremos é a raiz do erro quadrático médio (RMSE), dada pela expressão:

$$RMSE = \sqrt{\frac{\sum_{i=1}^M \sum_{j=1}^P (\phi_i - \theta_{ij})^2}{MP}}, \quad (8)$$

onde ϕ_i é o ângulo de chegada do i -ésimo sinal incidente,

θ_{ij} é a estimativa do i -ésimo ângulo de chegada no j -ésimo experimento, M é o número de fontes e P representa o número de experimentos. Em todas as simulações utilizamos $P = 1000$.

O limite de Cramer-Rao (Stoica e Nehorai, 1990) representa o limite teórico de desempenho em termos de variância para qualquer estimador não-polarizado. Contudo, como já destacado, a referência de desempenho para os algoritmos bio-inspirados é a busca exaustiva. Assim sendo, foi obtida a curva de RMSE para a busca exaustiva em função da SNR considerando o cenário de teste definido anteriormente.

O valor do limite de Cramer-Rao na máxima SNR considerada, o qual denotamos pelo símbolo ϵ , nos dá um indicativo da distância média existente entre o ótimo global da função custo e o ponto associado aos verdadeiros ângulos de chegada. Uma vez que desejamos minimizar os erros introduzidos pela discretização do espaço, dada pela grade uniforme, o método de busca exaustiva deve utilizar uma grade de precisão $p \ll \epsilon$. Tendo isto em mente, adotamos $p = 0,01$. Sendo assim, uma vez que temos $M = 2$, o que corresponde à dimensão do espaço de busca, e que $(-90^\circ, 90^\circ)$ indica a faixa de possíveis valores de cada ângulo de chegada, o número de vezes que a busca exaustiva avalia a função custo em cada experimento é aproximadamente igual a $N_P = 3,2 \cdot 10^8$.

A Figura 3 apresenta o limite de Cramer-Rao e a curva de RMSE para a busca exaustiva em função da relação sinal-ruído.

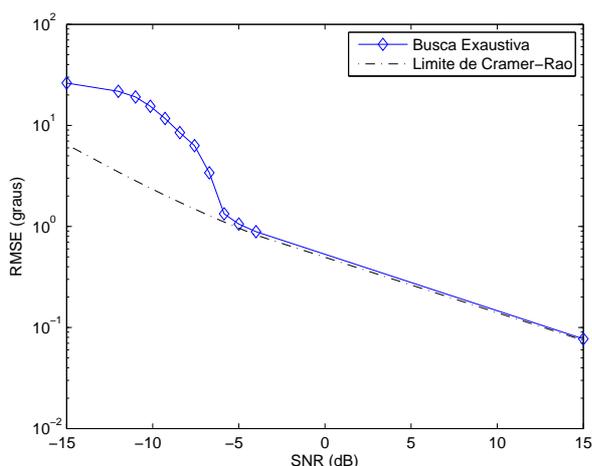


Figura 3: Limite de Cramer-Rao e RMSE para busca exaustiva.

Podemos observar na Figura 3 que existe um valor de SNR abaixo do qual a curva de RMSE da busca exaustiva deixa de acompanhar o limite de Cramer-Rao. Este particular valor é denominado de SNR de limiar, e o efeito correspon-

dente é denominado efeito de limiar (Rife e Boorstyn, 1974) (Forster et al., 2004). A qualidade da estimação dos ângulos de chegada realizada por cada algoritmo bio-inspirado também será analisada através da medida de RMSE, para então ser comparada com a referência de desempenho, dada pela busca exaustiva. Quando pertinente, os histogramas das estimativas dos ângulos de chegada também serão analisados, uma vez que permitem visualizações distintas daquela oferecida pela medida de RMSE ao evidenciar a distribuição das estimativas ao longo dos experimentos.

Conforme já destacado, além da comparação de desempenho via RMSE, é preciso analisar os custos computacionais associados a cada algoritmo, incluindo a busca exaustiva. Tal análise pode ser feita com base no número de vezes que cada algoritmo avalia a função custo J_{ML} . No caso da busca exaustiva, tal número é igual a $N_P = 3,2 \cdot 10^8$, como calculado anteriormente. Já no caso dos algoritmos bio-inspirados, o número de vezes que a função de *fitness* é avaliada depende dos valores atribuídos aos parâmetros de cada algoritmo. Isto impõe uma dificuldade adicional à etapa de ajuste de parâmetros: os valores escolhidos não podem fazer com que o número de avaliações da função de *fitness* fique próximo ao da busca exaustiva. Entretanto, devem ser adequados o suficiente para tornar cada algoritmo capaz de localizar o ótimo global e, assim, obter um bom desempenho em termos de RMSE.

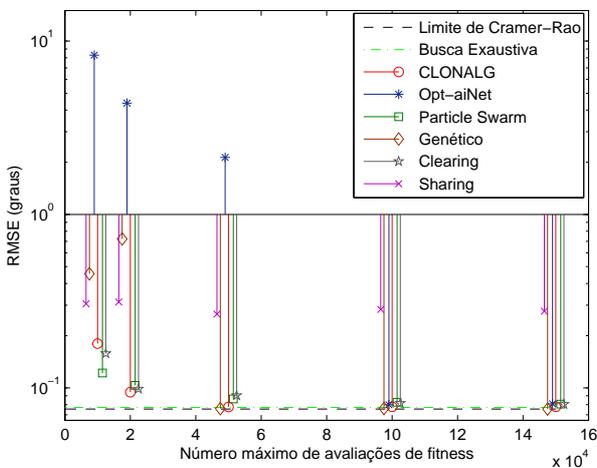
Diante disso, é necessário que seja analisado o impacto do número de avaliações da função de *fitness* no desempenho de cada algoritmo. Por isso, na próxima seção serão apresentadas curvas com os valores de RMSE em função do número de avaliações para alguns valores de relação sinal-ruído. Nestas simulações, o número máximo de avaliações da função de *fitness* terá os seguintes valores: {10.000, 20.000, 50.000, 100.000, 150.000}.

Para apresentar a última forma de avaliação de desempenho, é preciso considerar que, no âmbito da computação natural, uma grande preocupação é a questão da manutenção da diversidade da população de soluções candidatas. Se no decorrer das gerações a diversidade cai muito rapidamente, isto pode levar a uma convergência prematura para um ponto sub-ótimo. Por outro lado, se a diversidade é demasiadamente elevada e não há um esquema de preservação de soluções promissoras, isto também pode comprometer o desempenho do algoritmo.

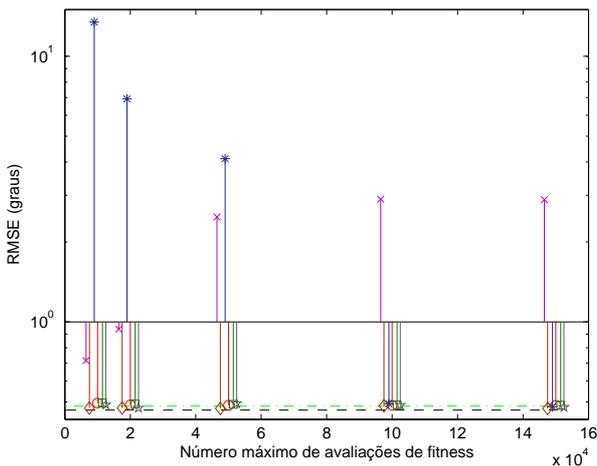
Uma maneira muito comum de se avaliar a capacidade de manutenção da diversidade de um algoritmo é a observação da separação entre as curvas de evolução do *fitness* médio da população e do *fitness* do melhor indivíduo. Estas curvas também serão analisadas. A seguir, apresentamos e discutimos os resultados obtidos.

6 RESULTADOS E ANÁLISE

Considerando o cenário de teste detalhado nas Seções 3 e 5, foi possível obter as curvas com os valores de RMSE em função do número de avaliações da função de *fitness* para cada um dos algoritmos bio-inspirados, considerando alguns valores de SNR. As Figuras 4 e 5 apresentam os resultados obtidos. Pedimos ao leitor que atente à escala logarítmica utilizada para o eixo de valores de RMSE.



(a) SNR = 15 dB

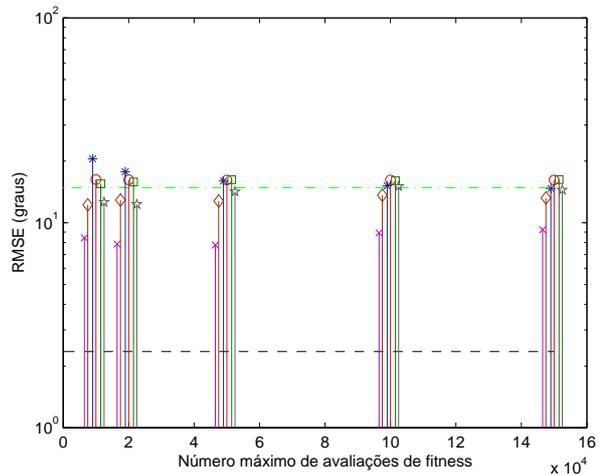


(b) SNR = 0 dB

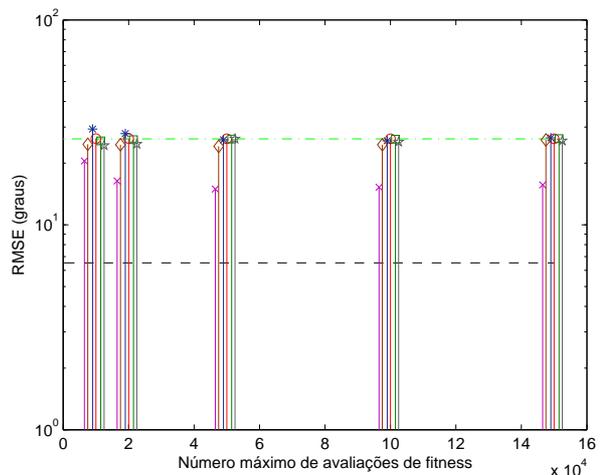
Figura 4: Curvas de RMSE em função do número de avaliações da função de *fitness*.

É interessante notar na Figura 4 que os valores de RMSE obtidos com os algoritmos bio-inspirados se aproximam do valor obtido com a busca exaustiva conforme aumentamos o número máximo de avaliações da função de *fitness*. A exceção é o método de *fitness sharing*, cujo desempenho,

para alguns valores de SNR, praticamente não variou com o número de avaliações. Também é interessante observarmos que, para um mesmo número máximo de avaliações da função de *fitness*, o desempenho de cada algoritmo pode ser bem diferente. Por exemplo, considerando 10.000 avaliações da função de *fitness* na Figura 4(a), é possível perceber que o algoritmo *particle swarm* foi aquele que mais se aproximou da busca exaustiva e que a *opt-aiNet* obteve um valor de RMSE bastante superior. Entretanto, quando consideramos 100.000 avaliações da função de *fitness*, ambos obtiveram desempenhos similares. Isto serve para ilustrar a influência que o número de avaliações da função de *fitness* exerce no desempenho de cada algoritmo.



(a) SNR = -10 dB



(b) SNR = -15 dB

Figura 5: Curvas de RMSE em função do número de avaliações da função de *fitness*.

Em suma, é possível notar que, a partir de 100.000 avaliações

da função de *fitness*, os valores de RMSE obtidos para os algoritmos *clearing*, CLONALG, *opt-aiNet* e *particle swarm* são praticamente iguais aos da busca exaustiva para todos os valores de SNR considerados. As pequenas discrepâncias existentes decorrem, fundamentalmente, do número de simulações utilizado e também do conjunto de valores definido para os parâmetros de cada ferramenta. Ou seja, avaliando um número muito menor de pontos do espaço de busca, o que representa um custo computacional menor, estes algoritmos praticamente reproduziram o desempenho da busca exaustiva.

A Tabela 1 apresenta os valores atribuídos aos principais parâmetros de cada algoritmo³ considerando um número máximo de avaliações da função de *fitness* igual a 150.000. Escolhemos este valor uma vez que é bastante inferior ao número de avaliações requerido pela busca exaustiva e também pelo fato de os algoritmos bio-inspirados terem obtido desempenhos ligeiramente superiores nesta situação.

Algoritmo	Parâmetro	Valor
<i>clearing</i>	Número de indivíduos	100
	Número de gerações N_{gen}	750
	Número de torneios T_o	8
	Número de competidores q	7
CLONALG	Número de indivíduos	50
	Número de clones N_c	10
	Número de gerações N_{gen}	275
<i>opt-aiNet</i>	Número inicial de indivíduos N_i	50
	Número de clones N_c	5
<i>particle swarm</i>	Número de indivíduos	150
	Número de gerações N_{gen}	1000

Tabela 1: Conjunto de parâmetros adotado.

Empregando os conjuntos de parâmetros apresentados na Tabela 1, foi possível obter a curva de RMSE em função da relação sinal-ruído para estes algoritmos bio-inspirados, para os métodos MODE, MODEX e busca exaustiva, além do limite de Cramer-Rao. A Figura 6 apresenta os resultados.

Observando a Figura 6, é possível perceber que, ao longo de toda a faixa de valores de SNR, os algoritmos *clearing*, CLONALG, *opt-aiNet* e *particle swarm* acompanharam o desempenho da busca exaustiva. Isto significa que estas ferramentas conseguiram implementar o método de estimação ML para o problema de estimação DOA. Observe ainda que estas ferramentas superaram os métodos MODE e MODEX na região de limiar.

Com o propósito de ilustrarmos o comportamento destas ferramentas em um contexto mais desafiador e também de

³São apresentados os parâmetros cujos valores não foram explicitados na Seção 4, uma vez que são determinantes na definição do número de avaliações da função de *fitness*.

verificar a escalabilidade do problema de estimação DOA, apresentamos na Tabela 2 o valor de RMSE obtido pelos algoritmos CLONALG e *particle swarm* considerando 1000 experimentos em um cenário com três ângulos, a saber, $\{10^\circ, 15^\circ, 40^\circ\}$ para a SNR de 20 dB. O conjunto de parâmetros adotado é equivalente ao apresentado na Tabela 1, o que significa que o número máximo de avaliações da função de *fitness* permanece igual a 150.000. Neste cenário, é absolutamente inviável utilizarmos uma busca exaustiva, já que, caso mantivéssemos a precisão da busca em $p = 0,01$, o número de pontos existentes na grade uniforme seria aproximadamente igual a $5,83 \cdot 10^{12}$. Por isso, sabendo que para valores altos de SNR, o desempenho da busca exaustiva tende assintoticamente ao limite de Cramer-Rao, incluímos também na Tabela 2 este último valor, o qual servirá de referência para o desempenho dos algoritmos bio-inspirados.

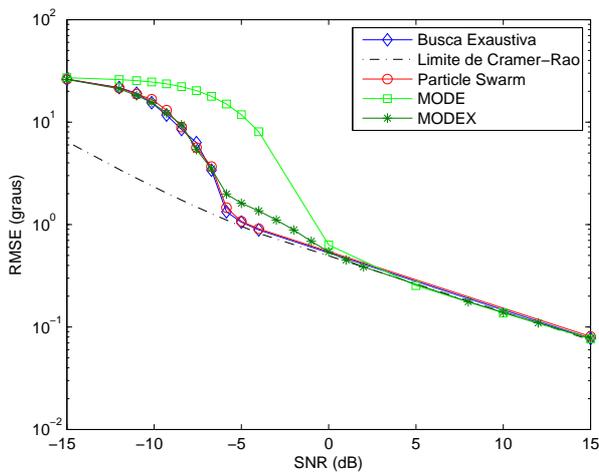
Algoritmo	RMSE
<i>particle swarm</i>	0,0687
CLONALG	0,0773
Cramer-Rao	0,0377

Tabela 2: RMSE para a SNR de 20 dB no cenário com 3 fontes.

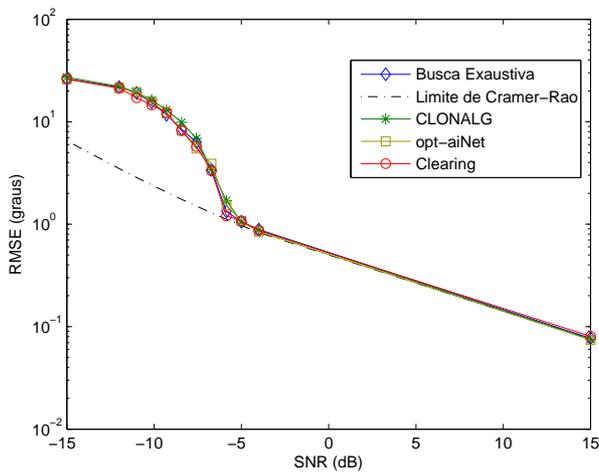
Podemos perceber que, mesmo sem reajustarmos os parâmetros, o que certamente melhor adaptaria os algoritmos ao cenário escolhido, o desempenho destas ferramentas se mantém próximo ao limite de Cramer-Rao. Note também que tal desempenho foi obtido considerando um número de avaliações da função de *fitness* muito inferior ao exigido pela busca exaustiva. Este fato ressalta o potencial da aplicação destes algoritmos bio-inspirados no problema de estimação DOA como alternativa aos métodos existentes.

Os dois algoritmos restantes, a saber, genético e *fitness sharing*, não foram capazes de reproduzir plenamente o desempenho alcançado pela busca exaustiva, conforme mostra a Figura 4. Por causa disto, faremos também uma análise particular para cada um deles evidenciando as causas desta performance inferior. Em síntese, podemos observar na Figura 4 a ocorrência de duas situações: 1) para SNRs altas, alguns valores de RMSE estão acima da busca exaustiva. 2) para SNRs baixas, especialmente abaixo do limiar, alguns valores de RMSE estão abaixo da busca exaustiva.

A primeira situação está presente na Figura 4(a) para o algoritmo *fitness sharing*. O fato de termos valores de RMSE acima da busca exaustiva não necessariamente significa que o algoritmo não foi capaz de localizar o ótimo global da superfície de *fitness* na maioria das simulações. Na verdade, basta que falhe em alguns poucos experimentos e forneça estimativas um pouco distantes dos valores verdadeiros dos ângulos de chegada para que a medida de erro se eleve. Para que isto fique claro, apresentamos na Figura 7 os histogra-



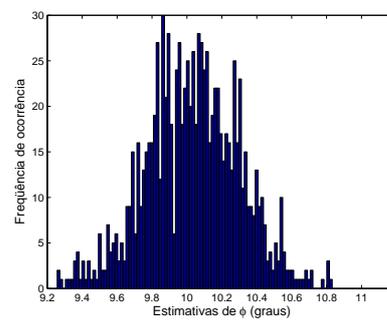
(a) RMSE - *particle swarm*



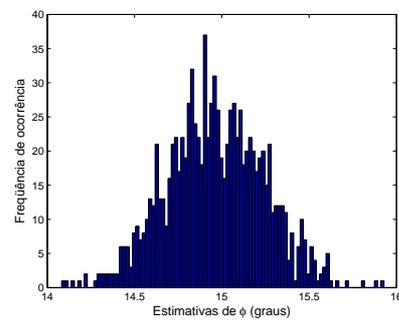
(b) RMSE - *clearing*, *CLONALG* e *opt-aiNet*

Figura 6: Curvas de RMSE.

mas das estimativas dos ângulos de chegada referentes ao algoritmo *fitness sharing* considerando a SNR de 15 dB e 150.000 avaliações da função de *fitness*. Também apresentamos na Figura 8 os histogramas das estimativas dos ângulos de chegada referentes ao algoritmo *particle swarm*⁴, os quais servem de referência para aqueles mostrados na Figura 7, uma vez que esta ferramenta obteve desempenho similar ao da busca exaustiva. Pedimos que o leitor atente para as escalas dos eixos de cada histograma.



(a) Estimativas de $\phi = 10^\circ$ para SNR = 15 dB



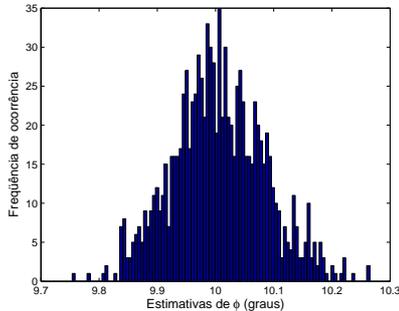
(b) Estimativas de $\phi = 15^\circ$ para SNR = 15 dB

Figura 7: Histogramas das estimativas - *fitness sharing*.

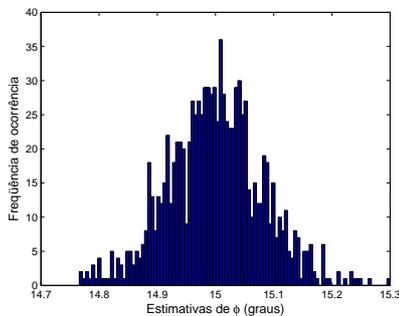
Observe na Figura 7 que o método de *fitness sharing* foi capaz de concentrar a maior parte das estimativas nas proximidades dos valores verdadeiros dos ângulos de chegada. Porém, em alguns experimentos, o algoritmo não foi capaz de refinar as soluções candidatas, de forma que algumas estimativas se afastaram do ótimo global cerca de 1° , o que, conseqüentemente, elevou o valor de RMSE. É interessante comparar estes histogramas com aqueles mostrados na Figura 8, referentes ao algoritmo *particle swarm*.

A segunda situação destacada anteriormente e que caracteriza um desempenho indesejado ocorre tanto com o algoritmo genético quanto com o método de *fitness sharing*. Podemos ver, na Figura 5(a), que existem valores de RMSE inferiores aos da busca exaustiva. Este fato é uma forte evi-

⁴Poderíamos também ter optado por outro algoritmo dentre aqueles que atingiram o mesmo desempenho da busca exaustiva.



(a) Estimativas de $\phi = 10^\circ$ para SNR = 15 dB



(b) Estimativas de $\phi = 15^\circ$ para SNR = 15 dB

Figura 8: Histogramas das estimativas - *particle swarm*.

dência de que ambas ferramentas não conseguiram localizar o ótimo global em alguns experimentos, mas sim pontos sub-ótimos localizados em regiões mais próximas aos valores verdadeiros dos ângulos de chegada que o ótimo global. Como consequência disto, foram obtidos valores de RMSE menores que os da busca exaustiva.

Com o propósito de elucidarmos este fato, exibimos na Tabela 3 a média μ e o desvio padrão σ do conjunto de estimativas obtido pelo algoritmo genético e pelo algoritmo *particle swarm* para SNR = -10 dB, considerando 150.000 avaliações da função de *fitness*.

Algoritmo	$\phi = 10^\circ$		$\phi = 15^\circ$	
	μ	σ	μ	σ
Genético	4,543	15,532	16,830	8,670
<i>particle swarm</i>	3,301	18,327	18,261	11,728

Tabela 3: Média e desvio padrão das estimativas - SNR = -10 dB.

Através dos resultados mostrados na Tabela 3, concluímos que houve uma maior concentração das estimativas obtidas pelo algoritmo genético em torno do valor verdadeiro do ângulo de chegada, já que o desvio padrão foi menor e a média esteve mais próxima do valor real se comparado com o *particle swarm*. O mesmo pode ser observado para o método de

fitness sharing, conforme mostra a Tabela 4.

Algoritmo	$\phi = 10^\circ$		$\phi = 15^\circ$	
	μ	σ	μ	σ
<i>fitness sharing</i>	-3,487	13,120	10,837	10,894
<i>particle swarm</i>	-8,600	25,660	23,123	17,810

Tabela 4: Média e desvio padrão das estimativas - SNR = -15 dB.

Ambas constatações ratificam o fato de que estes métodos não conseguiram localizar o ótimo global e acabaram convergindo para pontos sub-ótimos localizados em regiões mais próximas aos valores verdadeiros dos ângulos de chegada que o ótimo global. Com isto, os respectivos valores de RMSE foram inferiores àqueles da busca exaustiva.

Para concluir a análise do desempenho dos algoritmos bioinspirados, apresentamos na Figura 9 as curvas de evolução do *fitness* do melhor indivíduo e do *fitness* médio considerando uma execução típica de cada ferramenta para uma SNR de 0 dB.

Podemos notar que, em especial, os algoritmos *fitness sharing*, *clearing*, CLONALG, *opt-aiNet* e *particle swarm* preservaram uma boa separação entre as curvas, o que indica a existência de diversidade na população, fator que é bastante interessante em contextos multimodais. Por outro lado, observe que as curvas de *fitness* apresentadas na Figura 9(a), referentes ao algoritmo genético, praticamente se sobrepõem a partir de um certo ponto, o que indica que a diversidade da população chegou próxima a zero. Isto é indesejado, pois torna difícil a exploração de novas regiões do espaço e aumenta as chances de convergência para pontos sub-ótimos.

A solução deste conflito requer um estudo mais aprofundado acerca dos operadores genéticos e do impacto de cada parâmetro no desempenho do algoritmo, por pelo menos duas razões: 1) tal conflito não ocorre para todas as SNRs; 2) a simples redução da pressão seletiva pode fazer com que o algoritmo não localize o ótimo em algumas situações por não privilegiar as regiões mais promissoras.

7 CONCLUSÃO

O problema de estimação de direção de chegada pode ser formulado, mediante o critério ML, como um problema de minimização de uma função custo J_{ML} de caráter não-linear, não-quadrático, multimodal e variante com a relação sinal-ruído. Neste trabalho, aplicamos um conjunto representativo de algoritmos vinculados à computação natural para realizar a otimização desejada como alternativa aos métodos clássicos, como MODE e MODEX.

Simulações no contexto de um cenário de teste consagrado na literatura sobre o problema DOA mostraram que os al-

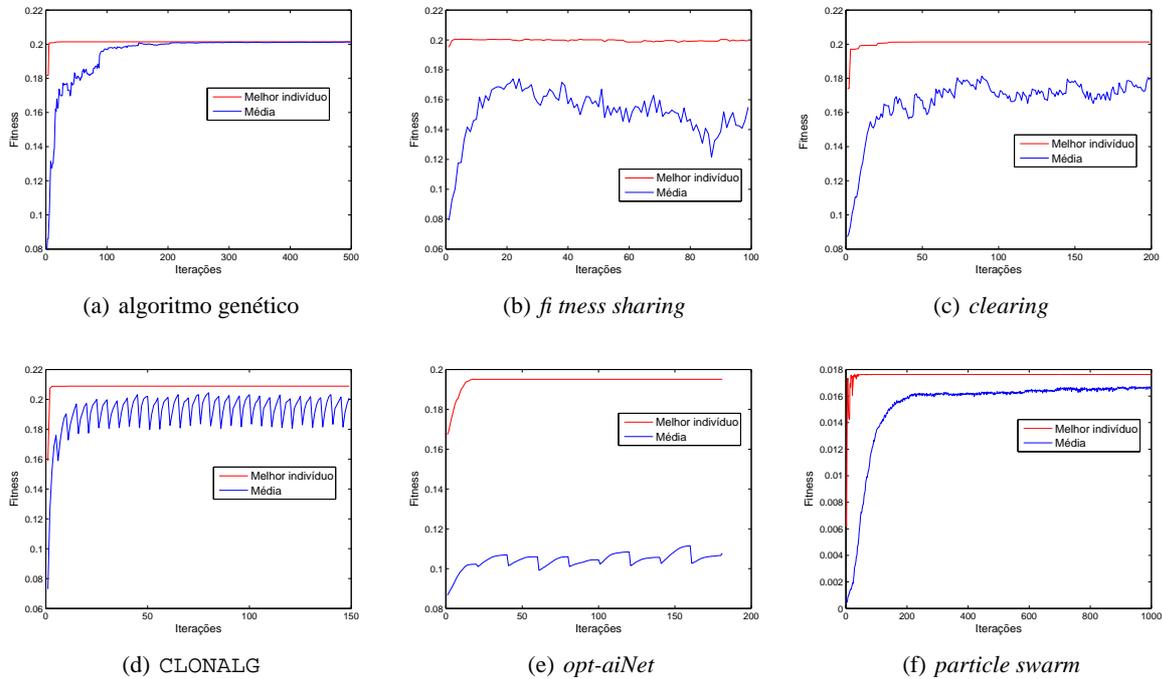


Figura 9: Curvas de evolução de f_i *mess* médio e f_i *mess* do melhor indivíduo.

algoritmos *particle swarm*, *clearing*, CLONALG e *opt-aiNet* determinam com precisão o mínimo global da função custo J_{ML} para uma ampla faixa de valores de relação sinal-ruído. Neste sentido, o desempenho destes algoritmos praticamente se iguala àquele oferecido pelo método de busca exaustiva. Vale ressaltar que todas estas técnicas apresentaram esforços computacionais inferiores ao de uma metodologia convencional de busca exaustiva, o que dá suporte à viabilidade da aplicação das mesmas.

No que tange à comparação com os algoritmos clássicos para a estimação DOA, as simulações também mostraram que estas ferramentas superam os métodos MODE e MODEX, pois estes últimos falham em determinar o mínimo global para um conjunto de valores de SNR, especialmente abaixo da região de limiar, conforme já evidenciado na literatura.

Este trabalho trouxe uma contribuição inicial para a análise da aplicação de ferramentas de computação natural no problema de estimação DOA. No entanto, deixa ainda muitos aspectos a serem considerados. Para obter uma caracterização definitiva das potencialidades dos algoritmos bio-inspirados no contexto do problema DOA, este estudo será aprofundado em várias direções, tais como: 1) análise da influência dos principais parâmetros no desempenho de cada algoritmo; 2) exploração de mecanismos que incorporem conhecimentos específicos do problema DOA na conformação dos algoritmos bio-inspirados; 3) tentativa de estimar a relação sinal-

ruído e aplicação desta estimativa para ajustes automáticos dos parâmetros visando a otimização do esforço computacional requerido por cada ferramenta; 4) análise do desempenho destes algoritmos considerando outros cenários de teste, inclusive com mais fontes e um número reduzido de sensores.

AGRADECIMENTOS

Este trabalho teve grande apoio da FAPESP, processo nº 07/57040-3, o qual se insere no projeto temático LINEARIDADE EM SINAIS, CIRCUITOS E SISTEMAS, processo nº 03/09858-6, e também do CNPq e CAPES. Agradecemos também aos revisores anônimos, pelos comentários e sugestões.

REFERÊNCIAS

- Ada, G. L. e Nossal, G. J. V. (1987). The clonal selection theory, *Scientific American* pp. 50–57.
- Attux, R. R. F., Loiola, M. B., Suyama, R., de Castro, L. N., Von Zuben, F. J. e Romano, J. M. T. (2003). Blind search for optimal wiener equalizers using an artificial immune network model, *EURASIP Journal of Applied Signal Processing* **2003**(6): 740–747.
- Bäck, T., Fogel, D. e Michalewicz, Z. (1997). *Handbook of Evolutionary Computation*, Institute of Physics Pu-

blishing and Oxford University Press.

- Cioppa, A. D., Stefano, C. D. e Marcelli, A. (2004). On the role of population size and niche radius in fitness sharing, *IEEE Transactions on Evolutionary Computation* **8**(6): 580–592.
- Coelho, L. S. e Mariani, V. C. (2006). Sistema híbrido neuro-evolutivo aplicado ao controle de um processo multivariável, *SBA Controle & Automação* **17**: 32–48.
- Darwen, P. e Yao, X. (1995). A dilemma for fitness sharing with a scaling function, *Evolutionary Computation, Proceedings of IEEE International Conference on*, Piscataway, NJ, pp. 166–171.
- de Castro, L. N. (2006). *Fundamentals of Natural Computing: Basic Concepts, Algorithms and Applications*, Chapman & Hall/CRC.
- de Castro, L. N. e Timmis, J. (2002). An artificial immune network for multimodal function optimization, *IEEE International Conference on Evolutionary Computation* **1**: 674–699.
- de Castro, L. N. e Von Zuben, F. J. (2002). Learning and optimization using the clonal selection principle, *IEEE Transactions on Evolutionary Computation* **6**(3): 239–251.
- Forster, P., Larzabal, P. e Boyer, E. (2004). Threshold performance analysis of maximum likelihood DOA estimation, *Signal Processing, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on]* **52**(11): 3183–3191.
- Gershman, A. e Stoica, P. (1999). MODE with extra-roots (MODEX): a new DOA estimation algorithm with an improved threshold performance, *IEEE International Conference on Acoustics, Speech, and Signal Processing* **5**: 2833–2836.
- Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley.
- Goldberg, D. E. e Richardson, J. (1987). Genetic algorithms with sharing for multimodal function optimization, *2nd Int. Conf. Genetic Algorithms*, pp. 41–49.
- Haykin, S. (1985). *Array Signal Processing*, Prentice Hall, Englewood Cliffs, NJ.
- Holland, J. (1992). *Adaptation in Natural and Artificial Systems*, 2nd edn, The MIT Press.
- Jerne, N. K. (1974). Towards a network theory of the immune system, *Ann. Immunol. (Inst. Pasteur)* pp. 373–389.
- Kay, S. M. (1993). *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*, Prentice Hall Signal Processing Series, Englewood Cliffs, NJ.
- Kennedy, J. (1997). The particle swarm: social adaptation of knowledge, *IEEE International Conference on Evolutionary Computation* pp. 303–308.
- Kennedy, J. e Eberhart, R. (1995). Particle swarm optimization, *IEEE International Conference on Neural Networks* **4**: 1942–1948.
- Krim, H. e Viberg, M. (1996). Two decades of array signal processing research: the parametric approach, *IEEE Signal Processing Magazine* **13**(4): 67–94.
- Krummenauer, R. (2007). *Filragem ótima na estimação de direção de chegada de ondas planas usando arranjo de sensores*, Master's thesis, School of Electrical and Computer Engineering - UNICAMP, Campinas-SP-Brazil.
- Lopes, A., Bonatti, I. S., Peres, P. L. D. e Alves, C. A. (2003). Improving the MODEX algorithm for direction estimation, *Signal Processing* **83**(9): 2047–2051.
- Mahfoud, S. W. (1995). *Niching Methods for Genetic Algorithms*, PhD thesis, University of Illinois at Urbana-Champaign.
- Manikas, A. (2004). *Differential Geometry in Array Processing*, Imperial College Press.
- Pérowski, A. (1996). A clearing procedure as a niching method for genetic algorithms, *Evolutionary Computation, Proceedings of IEEE International Conference on*, pp. 798–803.
- Rife, D. e Boorstyn, R. (1974). Single tone parameter estimation from discrete-time observations, *IEEE Transactions on Information Theory* **20**(5): 591–598.
- Sareni, B. e Krahenbuhl, L. (1998). Fitness sharing and niching methods revisited, *IEEE Transactions on Evolutionary Computation* **2**: 97–106.
- Silva, V. V. R., Khatib, W. e Fleming, P. J. (2007). Control system for a gas turbine engine using evolutionary computing for multidisciplinary optimization, *SBA Controle & Automação* **18**(4): 471–478.
- Stoica, P. e Nehorai, A. (1990). Performance study of conditional and unconditional direction-of-arrival estimation, *IEEE Transactions on Acoustics, Speech, and Signal Processing* **38**(10): 1783–1795.
- Stoica, P. e Sharman, K. C. (1990). Novel eigenanalysis method for direction estimation, *IEE Proceedings part F (Radar and Signal Processing)* **137**(1): 19–26.

Van Trees, H. L. (2001). *Optimum Array Processing. Part IV of Detection, Estimation and Modulation Theory*, John Wiley and Sons, New York, USA.