Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e de Computação

Aplicação de Redes Complexas para a Definição de Vizinhança na Otimização por Enxame de Partículas

Dissertação de Mestrado apresentada à Faculdade de Engenharia Elétrica e de Computação como parte dos requisitos para obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: Engenharia de Computação.

Autor: Alan Godoy Souza Mello
Orientador: Fernando José Von Zuben

Campinas, SP
2010

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

Mello, Alan Godoy Souza
M489a Aplicação de redes complexas para a definição de vizinhança na otimização por enxame de partículas / Alan Godoy Souza Mello. –Campinas, SP: [s.n.], 2010.

Orientador: Fernando José Von Zuben.
Dissertação de Mestrado - Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Otimização. 2. Redes complexas. 3. Heurísticas. I. Von Zuben, Fernando José. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

Título em Inglês:	Application of complex networks on the definition of neighborhood in particle swarm optimization
Palavras-chave em Inglês:	Optimization, Complex networks, Heuristics
Área de concentração:	Engenharia de Computação
Titulação:	Mestre em Engenharia Elétrica
Banca Examinadora:	Leandro dos Santos Coelho, Coelho, Leandro dos Santos, Ricardo Suyama, Suyama, Ricardo
Data da defesa:	16/04/2010
Programa de Pós Graduação:	Engenharia Elétrica

COMISSÃO JULGADORA - TESE DE MESTRADO

Candidato: Alan Godoy Souza Mello

Data da Defesa: 16 de abril de 2010

Título da Tese: "Aplicação de Redes Complexas para a Definição de Vizinhança na Otimização por Enxame de Partículas"

Prof. Dr. Fernando José Von Zuben (Presidente): Fernando José Von Zuben

Prof. Dr. Leandro dos Santos Coelho:

Dr. Ricardo Suyama: Ricardo Suyama

Resumo

Este trabalho propõe uma variante de otimização por enxame de partículas, na qual as influências entre as partículas são definidas através de uma rede complexa dinâmica. O objetivo principal é melhorar a capacidade de exploração do mecanismo de busca, particularmente junto a problemas multimodais. Com a vizinhança entre as partículas sendo definida de forma dinâmica e apresentando propriedades típicas de redes complexas, pretende-se: (*i*) promover uma rápida difusão de informação pela rede; (*ii*) viabilizar a formação de comunidades; e (*iii*) permitir que a influência das partículas ao longo da busca seja ajustável. Para validar o algoritmo proposto, foi feita uma análise sobre qual sua sensibilidade à variação de características da topologia. Em seguida, o seu desempenho foi comparado ao de outras propostas de otimização por enxame de partículas, presentes na literatura, utilizando para isso sete funções de teste com alta dimensionalidade e diferentes graus de dificuldade. Os resultados obtidos mostraram-se competitivos, indicando que topologias dinâmicas e complexas conduzem a mecanismos de busca eficazes e flexíveis, capazes de lidar com diferentes cenários de otimização.

Palavras-chave: Otimização, meta-heurísticas, enxame de partículas, redes complexas.

Abstract

This work proposes a variant of particle swarm optimization, in which the influences among particles are defined by a dynamic complex network. The main purpose of this work is to improve the exploration capability of the search mechanism, particularly in multimodal problems. With the neighborhood of the particles being defined in a dynamic way and presenting typical properties of complex networks, the intention is: (*i*) to promote a rapid diffusion of information throughout the network; (*ii*) to enable the formation of communities; and (*iii*) to allow an adjustable influence of the particles along the search. To validate the proposed algorithm, an analysis of its sensitivity to alternative characteristics of the topology was performed. Further, its performance was compared to other particle swarm optimization proposals, available in the literature, on seven high-dimensional benchmark functions presenting distinct difficulty levels. The obtained results were competitive, indicating that dynamic complex topologies guide to effective and flexible search mechanisms, capable of dealing with distinct optimization scenarios.

Keywords: Optimization, metaheuristics, particle swarm, complex networks.

Agradecimentos

A Deus, primeiramente, por ter concedido tanto a capacidade como as oportunidades necessárias para chegar até aqui;

A meus pais, por me ensinarem o valor do conhecimento e pelos tantos sacrifícios feitos por mim;

A meu orientador, pelas oportunidades, pela inspiração, pelo incentivo e por indicar os caminhos a trilhar neste princípio de vida acadêmica;

A minhas irmãs, pelos seus ouvidos sempre dispostos e simplesmente por serem fantásticas como são;

A meus colegas de república, por me aguentarem, mesmo em momentos de mau-humor;

A meus colegas de faculdade, por mostrarem que, mesmo que você não imagine, sentado ao seu lado pode haver uma pessoa brilhante;

À FEEC, à Unicamp, e ao povo de São Paulo, por oferecerem a estrutura necessária para que este sonho se concretizasse, e aos quais espero retribuir;

À CAPES, pelo apoio financeiro.

À minha família

Sumário

Lista de Figuras **xiii**

Lista de Tabelas **xv**

Índice de Acrônimos **xix**

1 Introdução **1**

1.1 Estrutura da dissertação 3

2 Auto-organização **5**

2.1 Ingredientes básicos da auto-organização 7

2.1.1 Comunicação e comportamentos emergentes 9

2.2 Características de sistemas auto-organizados 11

2.3 Conclusão do capítulo 13

3 Otimização por Enxame de Partículas **15**

3.1 Otimização por Enxame de Partículas 15

3.2 Modificações 18

3.3 Aplicações 24

3.4 Conclusão do capítulo 26

4 Redes complexas **27**

4.1 Conceitos sobre grafos 29

4.2 Características 30

4.2.1 Graus, distribuições e correlações 30

4.2.2 Caminhos entre vértices e efeito *small-world* 32

4.2.3 Coeficiente de aglomeração e comunidades 34

4.2.4 Outras características 36

4.3 Modelos de crescimento 38

4.3.1 Modelo *small-world* . . . 38
4.3.2 Modelo livre-de-escala . . . 39
4.4 Conclusão do capítulo . . . 45

5 Aplicação de Redes Complexas na Definição de Vizinhança entre Partículas 47
5.1 Os algoritmos CNPSO . . . 47
5.2 Experimentos . . . 51
5.2.1 Funções de teste . . . 53
5.2.2 Ajustes da topologia . . . 54
5.2.3 Índice de desempenho . . . 55
5.3 Ajustes de parâmetros . . . 56
5.3.1 CNPSO e supressão de ajustes topológicos . . . 56
5.3.2 Propostas vistas na literatura . . . 64
5.3.3 Versões híbridas . . . 67
5.4 Comparação de resultados após ajustes de parâmetros . . . 72
5.5 Conclusão do capítulo . . . 75

6 Conclusões e Perspectivas Futuras 85

Referências bibliográficas 87

A Gráficos das funções de teste utilizadas 103

B Gráficos de convergência 107
B.0.1 Função Salomon . . . 108
B.0.2 Função Rastrigin . . . 112
B.0.3 Função Schwefel . . . 116
B.0.4 Função Rosenbrock . . . 120
B.0.5 Função Sphere . . . 124
B.0.6 Função Griewank . . . 128
B.0.7 Função Ackley . . . 132

C Valores e desvios-padrão 137

Lista de Figuras

2.1 Diagrama de bifurcação para a equação de diferença logística. 12

3.1 Algumas topologias utilizadas no PSO. 20
3.2 Exemplo da adaptação topológica no H-PSO. 21
3.3 Uma topologia do TRIBES. 23

4.1 Rede de amizade em uma escola americana. 35
4.2 Alguns motivos típicos e suas denominações. 37
4.3 Procedimento de reconexão aleatória no modelo de Watts-Strogatz. 38

A.1 Função Ackley. 103
A.2 Função Griewank. 104
A.3 Função Rastrigin. 104
A.4 Função Rosenbrock. 105
A.5 Função Salomon. 105
A.6 Função Schwefel. 106
A.7 Função Sphere. 106

Lista de Tabelas

5.1 Funções de teste utilizadas para avaliação dos algoritmos. 53
5.2 Espaços de busca e de inicialização, posição do mínimo global e valores a serem alcançados para cada uma das funções de teste. 54
5.3 Testes de modelos de geração da rede inicial do CNPSO – Resultados médios e indices de desempenho (d_f). 56
5.4 Testes de modelos de geração da rede inicial do CNPSO – Desvios-padrão. 57
5.5 Testes de modelos de geração da rede inicial do CNPSO – Taxas de sucesso. 57
5.6 Testes de modelos de geração da rede inicial do CNPSO – Avaliações de função esperadas até atingir o objetivo (AF_{esp}). 57
5.7 Testes de tamanhos do enxame e graus médios para a versão do CNPSO sem atualizações topológicas – Índices de desempenho (d_f). Ressalta-se que menores valores do índice de desempenho indicam melhores resultados. 58
5.8 Testes de tamanhos do enxame e graus médios para a versão do CNPSO sem atualizações topológicas – Taxas de sucesso. 58
5.9 Testes de tamanhos do enxame e graus médios para a versão do CNPSO sem atualizações topológicas – Médias das avaliações de funções esperadas até atingir o objetivo (AF_{esp}). 59
5.10 Testes de tamanhos do enxame e graus médios para a primeira versão do algoritmo - CNPSO-1 – Índices de desempenho (d_f). 59
5.11 Testes de tamanhos do enxame e graus médios para a primeira versão do algoritmo - CNPSO-1 – Taxas de sucesso. 60
5.12 Testes de tamanhos do enxame e graus médios para a primeira versão do algoritmo - CNPSO-1 – Médias das avaliações de funções esperadas até atingir o objetivo (AF_{esp}). 60
5.13 Testes de tamanhos do enxame e graus médios para a segunda versão do algoritmo - CNPSO+ – Índices de desempenho (d_f). 61
5.14 Testes de tamanhos do enxame e graus médios para a segunda versão do algoritmo - CNPSO+ – Taxas de sucesso. 61

5.15 Testes de tamanhos do enxame e graus médios para a segunda versão do algoritmo - CNPSO+ – Médias das avaliações de funções esperadas até atingir o objetivo (AF_{esp}). 61
5.16 Testes de quantidade de adaptações topológicas para as versões do CNPSO – Índices de desempenho (d_f). 63
5.17 Testes de quantidade de adaptações topológicas para as versões do CNPSO – Taxas de sucesso. 63
5.18 Testes de quantidade de adaptações topológicas para as versões do CNPSO – Médias das avaliações de funções esperadas até atingir o objetivo (AF_{esp}). 64
5.19 Testes de tamanhos do enxame para outros PSOs, vistos na literatura – Índices de desempenho (d_f). 65
5.20 Testes de tamanhos do enxame para outros PSOs, vistos na literatura – Taxas de sucesso. 65
5.21 Testes de tamanhos do enxame para outros PSOs, vistos na literatura – Médias das avaliações de funções esperadas até atingir o objetivo (AF_{esp}). 66
5.22 Testes de tamanhos do enxame e graus médios para a versão híbrida do FIPS com o CNPSO+ – Índices de desempenho (d_f). 67
5.23 Testes de tamanhos do enxame e graus médios para a versão híbrida do FIPS com o CNPSO+ – Taxas de sucesso. 68
5.24 Testes de tamanhos do enxame e graus médios para a versão híbrida do *Bare Bones Particle Swarm* com o CNPSO+ – Índices de desempenho (d_f). 69
5.25 Testes de tamanhos do enxame e graus médios para a versão híbrida do *Bare Bones Particle Swarm* com o CNPSO+ – Taxas de sucesso. 69
5.26 Testes de tamanhos do enxame e graus médios para a versão híbrida do AR-PSO com o CNPSO+ – Índices de desempenho (d_f). 70
5.27 Testes de tamanhos do enxame e graus médios para a versão híbrida do AR-PSO com o CNPSO+ – Taxas de sucesso. 70
5.28 Testes de quantidade de adaptações topológicas para as versões híbridas do CNPSO+ – Índices de desempenho (d_f). 71
5.29 Testes de quantidade de adaptações topológicas para as versões híbridas do CNPSO+ – Taxas de sucesso. 71
5.30 Comparação de desempenho dos algoritmos – Resultados médios das funções objetivo. Ressalta-se que, quanto mais intenso o vermelho no fundo de uma célula, melhor o resultado nela contido. 77
5.31 Comparação final de desempenho dos algoritmos – Desvios-padrão. 78
5.32 Comparação final de desempenho dos algoritmos – Taxas de sucesso. 79

5.33 Comparação final de desempenho dos algoritmos – Avaliações de função esperadas até atingir o objetivo (AF_{esp}). 80
5.34 Classificação dos algoritmos segundo o índice de desempenho proposto neste trabalho. 81
5.35 Classificação dos algoritmos segundo seus índices de desempenho (Mendes *et al.* [127]). 81
5.36 Classificação dos algoritmos segundo suas taxas de sucesso médias. 82
5.37 Classificação dos algoritmos segundo a quantidade de avaliações de função esperadas até atingir o objetivo (AF_{esp}). Aponta-se que os espaços sinalizados com "–" indicam que o algoritmo testado não foi capaz de obter ao menos uma execução bem sucedida em alguma das funções de teste. 83

Índice de Acrônimos

AH-PSO *Adaptive Hierarchical Particle Swarm Optimizer*.

AR-PSO *Attractive and Repulsive Particle Swarm Optimization*.

DMS-PSO *Dynamic Multi-Swarm Particle Swarm Optimizer*.

DNA Ácido desoxirribonucleico.

gbest *Global best*.

GRASP *Greedy Randomized Adaptive Search Procedure*.

H-PSO *Hierarchical Particle Swarm Optimizer*.

lbest *Local best*.

Modelo ER Modelo de grafo aleatório de Erdös e Rényi.

PSO Particle Swarm Optimization

VLSI *Very-large-scale integration*.

Capítulo 1

Introdução

Um processo de otimização consiste do ajuste dos parâmetros de um sistema, com a meta de levar alguma determinada métrica (denominada função objetivo) aos níveis mais baixos (ou mais altos) possíveis, respeitadas as restrições existentes. No planejamento de transporte coletivo, por exemplo, os itinerários dos veículos devem minimizar o combustível gasto, ao mesmo passo que deve ser suprida a demanda dos passageiros por acesso rápido a múltiplos destinos, sendo limitado a pequenas distâncias o caminho percorrido a pé. Outro exemplo de otimização pode ser visto na natureza, onde os animais devem executar as tarefas necessárias à perpetuação da espécie, como busca e digestão de alimentos, fuga de predadores e reprodução, com o menor gasto energético possível.

A resolução de problemas deste tipo com apoio computacional tem sido investigada pela Pesquisa Operacional desde a metade do século passado. Uma dificuldade muitas vezes encontrada, porém, é o fato de muitos dos problemas de otimização serem considerados NP-difíceis [74], ou seja, não há, ao menos atualmente, qualquer algoritmo que consiga resolvê-los de maneira geral e exata em tempo polinomial, o que faz com que o tempo de execução requerido seja inviável para instâncias grandes. Assim, é necessário abrir mão de algumas das garantias fornecidas por métodos exatos, valendo-se, por exemplo, de algoritmos de aproximação ou probabilísticos, que são técnicas que devem ser projetadas especificamente para um problema e que podem fornecer, respectivamente, ou soluções com um limite máximo de desvio em relação ao ótimo, ou a própria solução ótima, mas com probabilidade de o algoritmo falhar, sem retornar qualquer resultado. Quando não há uma técnica específica satisfatória, porém, é necessário o uso de métodos de resolução de problemas mais genéricos, as meta-heurísticas, que, no entanto, não fornecem garantias de convergência, de otimalidade e nem de custo para se chegar a uma proposta de solução. Alguns problemas de otimização clássicos, tais como o roteamento de veículos, o problema do caixeiro viajante e o problema da mochila, são exemplos de situações em que não há métodos específicos que forneçam bons resultados em tempo adequado. Assim, há considerável esforço devotado à pesquisa de meta-heurísticas eficientes, ha-

vendo uma diversidade de propostas como a busca tabu [77], o GRASP [71], o *simulated anealling* [107], os algoritmos genéticos [85], a otimização por colônia de formigas [49] e a otimização por enxame de partículas [101]. É ainda possível encontrar na literatura diversas variantes destes, tentando aprimorá-los, seja de forma mais genérica, seja quando aplicados a classes específicas de problemas. Uma classificação comum das meta-heurísticas leva em conta o tipo de espaço de busca associado ao problema de otimização. Existem meta-heurísticas específicas para espaços de busca discretos, como otimização por colônia de formigas, e meta-heurísticas específicas para espaços de busca contínuos, como otimização por enxame de partículas. No entanto, variantes para permitir o tratamento do outro tipo de espaço de busca já foram propostas junto às principais meta-heurísticas, de modo que elas podem ser consideradas em ambos os cenários de otimização. Neste trabalho, serão propostas variantes de otimização por enxame de partículas (ou PSO, do inglês *Particle Swarm Optimization*) voltadas para a otimização em espaços de busca contínuos.

Tal qual outras meta-heurísticas citadas anteriormente, o PSO é baseado em processos que ocorrem na natureza, fazendo, assim, parte de um ramo da inteligência artificial denominado computação natural. Esta técnica de otimização inspira-se na forma como o conhecimento se difunde em sociedades humanas, valendo-se de uma população de partículas interagindo na procura pelo ótimo de uma função objetivo. De forma similar a uma pessoa que está formando seu ponto de vista, cada partícula deve navegar pelo espaço de soluções estabelecendo sua trajetória através da combinação de sua própria experiência com a experiência de seus vizinhos, sendo atraída para regiões onde foram encontrados bons resultados anteriormente. Por sua simplicidade, o PSO apresenta como características principais baixa demanda de recursos computacionais e facilidade de implementação, o que, juntamente de uma convergência rápida, garante a aplicação deste método a uma ampla gama de problemas, obtendo soluções de boa qualidade em muitos deles. Como pode ser observado, a capacidade de otimização vista no PSO não é advinda do funcionamento das partículas (também chamadas de agentes) em si, que possuem uma dinâmica simples, ou da presença de um controlador centralizado, mas sim da combinação de uma forte comunicação entre as partículas com uma dinâmica de realimentação que garante a exploração de regiões promissoras já encontradas. Esta emergência de propriedades estruturais e comportamentais advindas de decisões tomadas de forma totalmente descentralizada pode ser vista em muitas outras situações, como no funcionamento do livre-mercado, nas interações de espécies em ecossistemas e em fenômenos de cristalização, sendo um exemplo de auto-organização.

A auto-organização é, assim, uma forma de um sistema chegar a um padrão sem quaisquer imposições externas, baseando-se na interação de seus elementos constituintes. Deste modo, para melhor entendê-la, torna-se crucial conhecer as topologias que regulam essas relações, algo que tem demandado grande esforço de pesquisa nos últimos anos. De início imaginava-se que, dada a diversidade

dos mecanismos envolvidos e a quantidade de elementos que eles contêm, estas redes seriam frutos de processos plenamente regidos pelo acaso. No entanto, estudos mais recentes indicaram que grande parte delas compartilha algumas características que não são explicadas completamente nem por modelos de redes aleatórias nem por grafos regulares. Tais redes são, assim, denominadas redes complexas e constituem uma classe à parte, na qual se encontram as redes sociais, as redes de computadores, as redes de interações gênicas e metabólicas, entre outras.

Apesar dessa correlação entre processos auto-organizáveis e redes complexas, pouco se aprofundou sobre quais os efeitos da aplicação destas topologias à definição de vizinhança na otimização por enxame de partículas. Tendo isso em vista, este trabalho propõe um conjunto de variantes do PSO nas quais as influências entre partículas são determinadas por redes complexas, tanto dinâmicas quanto estáticas, analisando seus desempenhos e comparando os resultados obtidos com aqueles de outras propostas vistas na literatura. Como a topologia dos mecanismos de inteligência de enxame vistos na natureza foi modelada pela evolução natural durante milhões de anos, estima-se que esta combinação possua potencial de gerar benefícios. Características como o agrupamento das partículas em comunidades e a rapidez na dispersão de informações, comuns a redes complexas, poderiam, por exemplo, garantir a capacidade do enxame refinar paralelamente e de forma eficiente múltiplas soluções, porém sem desperdiçar muito processamento em regiões de baixa qualidade. Além de implementar mais fielmente a metáfora na qual o PSO se baseia, busca-se aqui, também, analisar quais das propriedades destas redes são, de fato, positivas a esta técnica de otimização, e também sob quais condições e conjuntos de parâmetros estes possíveis ganhos são maximizados.

1.1 Estrutura da dissertação

No capítulo 2 são introduzidos conceitos sobre auto-organização, que dão suporte conceitual e permitem uma melhor compreensão dos capítulos seguintes. A otimização por enxame de partículas, suas variantes e aplicações são apresentadas em detalhes no capítulo 3. Em seguida, no capítulo 4, é feita uma explicação aprofundada sobre redes complexas, com a apresentação de suas características mais importantes e principais métodos de crescimento propostos até o momento. No capítulo 5, são propostas as versões do PSO utilizando vizinhança complexa, as quais representam as principais contribuições deste trabalho, e são apresentados os procedimentos experimentais adotados e os resultados obtidos, com suas respectivas análises. Por fim, no capítulo 6 são apresentados comentários conclusivos, seguidos da indicação de direções a serem melhor exploradas futuramente nesta área.

Capítulo 2

Auto-organização

Observando revoadas de aves no céu, é possível notar a alta coordenação do vôo de seus membros em um balé aéreo que, à primeira vista, sugeriria a existência de alguma forma de controle centralizado. Dentre os insetos, é possível observar esse tipo de coordenação em gafanhotos e grilos da espécie *Anabrus simplex*, que são capazes de definir qual rumo uma nuvem com milhares de integrantes deve seguir [19, 171], comportamento similar ao observado em formigas-correição, que organizam expedições para captura de presas envolvendo mais de duzentas mil operárias [29]. Ao contrário do esperado, porém, esses comportamentos globais surgem de maneira descentralizada a partir da forte interação dos elementos do sistema, pertencendo a uma ampla classe de fenômenos denominados auto-organizados.

O mais notável efeito de um processo auto-organizável é o aumento da habilidade coletiva em realizar tarefas superando, assim, a mera soma das capacidades de seus membros. Formigas-tecelãs (*Oecophylla smaragdina*), por exemplo, formam cadeias com seus corpos, sendo capazes de constituir pontes para que operárias atravessem grandes espaços, como fendas no chão, e de juntar várias folhas para a formação de ninhos [88, 89]. Múltiplas dessas cadeias podem se unir, caso seja necessária uma força ainda maior ou caminhos mais amplos.

Em colônias de insetos sociais é comum a separação da colônia em diferentes grupos, responsáveis por subconjuntos das tarefas disponíveis. Tais grupos podem ser estabelecidos pela idade de seus integrantes, por sua morfologia ou aleatoriamente. Em formigas do gênero *Pheidole*, por exemplo, há a existência de duas classes morfológicas, definidas geneticamente: as operárias menores zelam pela manutenção no ninho e das larvas, enquanto que as operárias maiores são responsáveis pela proteção da colônia e por cortar grandes presas. A despeito dessa especialização, a divisão do trabalho é adaptável, de forma que quando há menos operárias menores que o necessário, em poucas horas a colônia pode se reoganizar, com algumas operárias maiores sendo incubidas das tarefas de manutenção [181]. Tal situação exemplifica como pode ser estabelecido um comportamento auto-organizável

mesmo entre indivíduos com funcionamentos diferentes.

As idéias modernas de auto-organização surgiram com os estudos da Cibernética, nos anos 40 [59], tendo se estendido por diversas disciplinas como a Matemática, a Física, a Economia e a Biologia. Pode-se definir a auto-organização como o conjunto de mecanismos dinâmicos através dos quais os múltiplos elementos de um sistema conseguem estabelecer uma ordem em nível global sem sua imposição por quaisquer forças externas [23, 26]. Tal comportamento à primeira vista pode parecer violar os princípios da entropia e leis da termodinâmica, porém ele torna-se possível uma vez que os sistemas auto-organizados são capazes de extrair energia do ambiente para conseguir manter-se distantes do equilíbrio.

Neste momento é importante diferenciar os significados de ordem e organização. A organização é definida como o conjunto de relações, ou restrições mútuas, que governam os elementos de um sistema, enquanto a ordem é dada pela existência de regularidades ou padrões espaciais ou temporais em um conjunto, sendo a presença de ordem uma característica de certas organizações. Como exemplo, tome-se a descrição feita por Ashby [7] da evolução no tempo das variáveis x_i que definem o estado de um sistema valendo-se de um conjunto de equações diferenciais:

$$\frac{dx_i}{dt} = f_i(x_1, x_2, ..., x_n), \tag{2.1}$$

Nessa situação, a organização do sistema é dada pelo conjunto de funções f_i, enquanto que sua ordem é definida pelos valores das variáveis x_i.

Regularmente, mas não necessariamente, a ordem auto-organizada é emergente no sistema, ou seja, é uma propriedade qualitativamente nova que surge das múltiplas interações dos seus elementos, não podendo ser compreendida através da simples combinação das contribuições individuais [35]. Relativamente ao funcionamento geral do sistema, as regras utilizadas pelos agentes são extremamente simples, sendo executadas sempre sobre informações locais, sejam elas obtidas do ambiente ou de outros agentes, sem quaisquer referências ao padrão global. Muitas vezes sequer o comportamento dos componentes do sistema pode ser reconhecido como uma parte do padrão global, sendo seus efeitos colaterais que contribuem para chegar à ordem [166]. Propriedades globais complexas podem surgir mesmo quando as sub-unidades são similares, não apresentam especialização e não possuem grande quantidade de regras de interação. Tais características levam a uma arquitetura baseada em um modelo descentralizado de fluxo de informação entre os agentes, fazendo com que a principal estrutura de controle do sistema não seja um agente central localizado no topo de uma hierarquia mas sim o próprio conjunto (heterarquia).

Apesar de a auto-organização estar presente em numerosos sistemas, na natureza é usual observar ocorrências de ordem que prescindem deste modelo, sendo alternativas comuns à auto-organização o estabelecimento de ordem através de um líder bem-informado, a utilização de plantas ou modelos

mentais para a definição do padrão a ser alcançado, a execução de receitas detalhadas e o uso de modelos pré-existentes [35]. Tais alternativas, porém, apresentam alguns inconvenientes, como demanda de alta capacidade intelectual por parte dos agentes do sistema, a exigência de canais de comunicação eficientes ligando todos os elementos, a falta de habilidade em adaptar-se a novas situações e a dificuldade em encontrar modelos disponíveis que apresentem as características desejadas. A despeito de algumas deficiências da auto-organização e idéias relacionadas em sistemas biológicos [58, 167], esse modelo produz regras econômicas no que tangem a exigências fisiológicas e comportamentais para implementá-la. Desta forma, a auto-organização tende a ser favorecida pela seleção natural quando torna-se custosa a codificação genética de outras propostas.

É importante notar que a ordem pode surgir de uma composição de interações internas e forças externas, não excluindo o envolvimento de outros mecanismos. Em colônias de abelhas, a despeito da auto-organização em vários aspectos, a mistura de feromônios emitida pela rainha é um importante regulador da organização do grupo [183]. Mesmo em sistemas que em um nível são considerados auto-organizados, quando analisados em uma situação diferente podem não ser classificados da mesma forma. O funcionamento integrado dos neurônios pode ser visto como um comportamento emergente, ao mesmo passo que o cérebro pode ser visto como um líder centralizado em relação aos outros sistemas fisiológicos.

Dada sua descentralização, sistemas auto-organizados costumam ser robustos, especialmente contra a perda de elementos. Além disso propriedades como a estigmergia, que será abordada mais adiante, provêm flexibilidade ao fazer com que modificações no ambientes decorrentes de fatores externos ao sistema sejam respondidas da mesma forma que seriam se fossem resultado de uma ação interna. Para o projeto de agentes artificiais tais propriedades são de extrema valia, de forma que atualmente existem várias técnicas inspiradas em processos auto-organizados, abrangendo, por exemplo, mecanismos de otimização [50, 101], classificação de padrões [109] e coordenação de robôs [51].

2.1 Ingredientes básicos da auto-organização

Ao executar suas regras baseando-se apenas em informações locais, os componentes de um sistema auto-organizado realizam uma ponderação sobre um conhecimento enviesado do ambiente fazendo com que, a despeito da semelhança dos elementos formantes, as respostas individuais sejam muito diversificadas levando a padrões globais quase imprevisíveis. Sendo assim, mostra-se uma tarefa árdua a compreensão de como esses componentes interagem produzindo um comportamento complexo. Um primeiro passo para alcançar tal entendimento é a análise dos ingredientes básicos

presentes em um processo auto-organizável, a saber, realimentação[1] positiva e negativa, amplificação de flutuações e múltiplas interações [26].

Algumas aves, como as gaivotas [110], costumam fazer seus ninhos em grandes colônias, em um arranjo que lhes provê alguns benefícios, como facilidade na detecção de predadores e na localização de fontes de alimentos. Esses conjuntos formam-se através de regras comportamentais muito simples, segundo as quais indivíduos que estão buscando locais para formar ninhos sentem-se atraídos por regiões onde já há ninhos de outros pássaros. Esse mecanismo pelo qual a ocorrência de um estímulo provoca respostas que causam seu aumento é chamado de amplificação, ou realimentação positiva. A realimentação positiva está presente na maioria dos processos auto-organizados, usualmente promovendo alterações no sistema. Pode-se observar a presença dessa dinâmica também no crescimento da população humana. Nos últimos séculos a população tem se reproduzido mais do que os níveis necessários para a manutenção de seu tamanho, de forma que um número maior de nascimentos ocorre a cada geração, o que aumenta ainda mais o tamanho da população. Outros casos em que verifica-se a realimentação positiva são a dispersão de infecções e o processo de imitação de comportamento, como bocejos e risos [151].

Mecanismos de realimentação positiva, quando atuando de forma isolada, são responsáveis por gerar comportamentos explosivos ou implosivos, algo potencialmente destrutivo que põe em risco a existência do próprio sistema. Assim, é necessária uma regulação capaz de prover estabilidade a processos auto-organizados. A realimentação negativa cumpre esse papel, ao promover modificações no sentido contrário às alterações verificadas, de forma que, quando uma pequena perturbação é aplicada ao sistema, seja ativada uma resposta visando neutralizá-la. Em sistemas fisiológicos, por exemplo, essa realimentação é usada para evitar flutuações indesejadas e manter a homeostase (estabilidade). Um caso disso é a regulação dos níveis de açúcar no sangue [133]. O aumento do açúcar no sangue dispara uma rápida resposta do pâncreas, liberando insulina. Essa insulina é responsável, então, por converter a glicose em glicogênio – um polissacarídeo capaz de armazenar energia – que se deposita no fígado e nos músculos. A regulação da temperatura corporal em mamíferos homeotermos também é realizada através de realimentação negativa: sensores na região do hipotálamo, no cérebro, monitoram a temperatura do sangue arterial do corpo [28], de forma que se esta se encontra em um determinado limite o corpo sente-se confortável. Caso contrário, se alguém encontrar-se em um ambiente frio, por exemplo, são disparados mecanismos comportamentais e fisiológicos visando o aumento da temperatura corporal.

A realimentação negativa pode tomar a forma de saturação, exaustão ou competição, não precisando ser codificada nas regras de funcionamento dos elementos do sistema, surgindo diretamente

[1] Realimentação é o mecanismo pelo qual a ocorrência de um estímulo gera uma resposta que influencia o próprio estímulo original, podendo ser entendida como o retorno de parte da saída do sistema para sua própria entrada. Tal processo é regularmente utilizado na engenharia quando objetiva-se a manutenção da estabilidade em um sistema.

de restrições físicas. Considere a exploração de recursos por abelhas: a quantidade de operárias em uma colméia extraindo recursos de determinada fonte é limitada superiormente pela quantidade de indivíduos existentes na colônia.

Muitas vezes os componentes de processos auto-organizáveis apresentam, especialmente em sua fase inicial, certa dose de aleatoriedade, a qual muitas vezes mostra-se crucial. Desta forma, ao acaso podem surgir em meios homogêneos algumas flutuações que, amplificadas pela realimentação positiva, permitem que novas soluções sejam descobertas ou dão origem a complexas estruturas espaço-temporais. Cupins, ao serem inseridos em um novo ambiente, começam a construir uma série de pilares e arcos, constituídos de grãos de areia e excremento [79]. Em uma primeira fase é observado um comportamento descoordenado, onde os operários depositam porções de areia em regiões aleatórias do espaço, muitas vezes removendo grãos depositados por outros indivíduos ou desmanchando seu próprio trabalho. Ocasionalmente, porém, surgem pequenos acúmulos de grãos, o que leva a uma súbita mudança de comportamento, sendo iniciada uma fase de trabalho mais coerente e coordenado. Uma vez que os cupins tendem a depositar grãos em regiões onde já exista um montículo formado, há a atração para que os demais operários comecem a construir em um mesmo local.

Algo comum que pode ser observado, direta ou indiretamente, em todas as ocorrências de auto-organização anteriormente citadas é a existência de múltiplas interações dos componentes do sistema. No caso de grupos de animais, especialmente, essas interações baseiam-se na troca de informações, permitindo que indivíduos se valham dos resultados de suas ações, assim como das de seus companheiros. A seguir, na subseção 2.1.1, serão abordadas mais profundamente as formas de comunicação em sistemas auto-organizáveis.

2.1.1 Comunicação e comportamentos emergentes

Conforme afirmado anteriormente, a comunicação entre os elementos do sistema é determinante para fenômenos auto-organizáveis. Sua forma específica pode variar, baseando-se em interações diretas ou indiretas e podendo ser composta de estímulos designados especificamente para transmitir informações ou que transmitem informações meramente por efeito colateral de outro comportamento [117, 158].

As interações diretas são evidentes a um observador externo, como contato visual, sonoro ou químico de um membro com seus companheiros. O mecanismo de exploração em abelhas, por exemplo, é baseado em contato direto: ao encontrar uma fonte de alimentos, uma abelha retornará à colônia e, dependendo de sua avaliação do recurso, ela pode realizar uma dança de recrutamento que codifica a localização da fonte e sua qualidade [176]. Danças que representam fontes mais próximas ou com melhores recursos atraem mais indivíduos e tendem a ser executadas mais vezes, de forma que a cada retorno à colônia mais abelhas sejam recrutadas para explorar as melhores fontes. Por outro lado,

por mais que uma fonte possua recursos com alta qualidade, cada dança pode atrair uma audiência limitada uma vez que, para que as abelhas possam compreender corretamente as informações ali indicadas, é necessário que elas estabeleçam contato próximo com a recrutadora. Como uma abelha não compara as informações contidas em múltiplas danças antes de decidir qual local explorar [159], a colônia é capaz de continuamente obter informações sobre locais inicialmente considerado de menor qualidade e, assim, redirecionar rapidamente os locais a serem explorados caso ocorram modificações no ambiente.

Na interação indireta a comunicação é mediada pelo próprio ambiente, com um indivíduo realizando determinada modificação no meio e outro companheiro seu reagindo de forma diferente ao novo ambiente. Esse modelo de colaboração através do meio físico é chamado de estigmergia [79, 80]. Em contraposição ao processo de recrutamento em abelhas, as formigas conduzem a colônia a fontes de alimentos através de trilhas [16]. Nesse caso, as formigas não mantêm contato direto mas seguem probabilisticamente caminhos marcados por feromônios, depositados por suas companheiras ao retornar de alguma fonte de alimento. Quanto maior a intensidade do feromônio, maior a probabilidade de uma formiga seguir a trilha. Combinado a esse comportamento, a deposição de feromônios em quantidade proporcional à qualidade da fonte aumenta as chances de que a colônia convirja para a exploração da melhor fonte de alimentos localizada até o momento. Apesar de usualmente seguirem as trilhas, há uma probabilidade não-nula de uma formiga se perder, o que aumenta a capacidade da colônia em encontrar novos recursos. Essa dinâmica, porém, apresenta como inconveniente a facilidade com que o grupo fica preso a uma determinada solução [170]. Quando uma formiga localiza uma fonte de melhor qualidade depois de a colônia ter convergido para uma solução, por exemplo, dificilmente ela conseguirá estabelecer uma trilha tão forte a ponto de recrutar colegas suficientes para redirecionar a exploração para o novo local. Outra face desse problema ocorre quando uma determinada fonte encontra-se extenuada, uma vez que as formigas continuarão a ser redirecionadas para ela tornando-se necessário adicionar novas interações ao sistema, como a existência de feromônios negativos.

Os casos das abelhas e das formigas são exemplos de situações em que a comunicação ocorre propositalmente. A trilha provocada por cervos ao caminhar por uma floresta, por outro lado, possui um caráter não intencional, sendo apenas um subproduto de múltiplos animais caminhando por um mesmo caminho. Mesmo assim, ela pode fornecer informação valiosa acerca de fontes de recursos, gerando comportamentos emergentes que muitas vezes são tidos como de origem desconhecida, uma vez que esse tipo de interação é deveras mais discreto e, logo, facilmente negligenciado [158].

O uso apenas de informações locais em sistemas auto-organizáveis não ocorre somente no relacionado à observação do ambiente, de forma que também seus elementos interagem diretamente apenas com um subconjunto do grupo, seus vizinhos. Em cardumes de peixes, por exemplo, a capacidade de

obter informações sobre o estado do grupo é extremamente limitada. Por um lado o tempo de resposta exigido para que um movimento coordenado seja atingido torna impossível que um peixe obtenha e processe informações sobre todos os seus companheiros, por outro lado a própria disposição espacial do cardume, compacto e com muitos indivíduos, impossibilita que um indivíduo consiga visualizar o posicionamento de todo o conjunto. A coordenação do cardume surge do uso de regras simples aplicadas pelos componentes do sistema a informações obtidas apenas de seus vizinhos [145]. A realimentação positiva funciona da mesma forma que ocorre com gaivotas, um indivíduo sendo atraído pela presença de outros indivíduos. A realimentação negativa, por sua vez, evita colisões fazendo com que cada peixe busque manter uma certa distância de seus companheiros, através de um ajuste constante de sua velocidade e direção à percepção da atividade de seus vizinhos. Qualquer atraso no processo de sincronização pode resultar em colisões e desgarramento do cardume.

2.2 Características de sistemas auto-organizados

Algumas propriedades costumam evidenciar a existência da auto-organização [26]. O surgimento de padrões espaço-temporais complexos, tais como trilhas de exploração, organização social e arquiteturas de ninhos, é uma delas. Cupins, por exemplo, possuem a capacidade de construir ninhos elaborados, dotados de sistemas de ventilação, câmaras centrais onde são mantidas as larvas, uma câmara com paredes reforçadas para a rainha e jardins onde são cultivados fungos, além de túneis conectando a colônia a suas fontes de alimentos [119, 120], tudo isso em um meio a princípio uniforme que é sucessivamente transformado pela colônia, em um sistema em que cada configuração do ambiente ativa um novo comportamento em um operário que, por sua vez, realizará uma diferente modificação na configuração atual.

Outra característica notável da auto-organização é a ocorrência de bifurcações, mudanças abruptas no comportamento do sistema ao serem realizadas pequenas modificações em seus parâmetros. Para exemplificar o fenômeno da bifurcação, tome-se o modelo matemático proposto por Robert May [125, 126] para a evolução do tamanho da população de um organismo hipotético cujas gerações não se sobreponham. Neste modelo o crescimento da população é determinado pela equação de diferença logística: $N_{t+1} = rN_t(1 - N_t)$. A única variável da equação, N_t, representa o tamanho da população na geração t, e pode variar entre 0 – a extinção da população – e 1 – a quantidade máxima de indivíduos suportada pelo ambiente –, enquanto que o parâmetro r representa a taxa de reprodução da espécie. A variação desse parâmetro gera diferentes comportamentos da população, como pode ser observado na figura 2.1. Para valores de r tal que $0 \leq r < 1$, temos que, independentemente de seu tamanho inicial N_0, após algumas iterações a população encontrar-se-á extinta enquanto que para $1 \leq r < 3$ a população tenderá a um valor específico, denominado atrator, e que também independe

de N_0. Aumentar r para além de 3 faz com que o padrão modifique-se radicalmente: inicialmente ocorrerá uma bifurcação e o tamanho final da população oscilará entre dois atratores, até que para $r > 3,4$ o sistema sofrerá outra transição a partir da qual o tamanho passará a oscilar entre 4 valores. A partir de $r > 3,57$ ocorrerão séries de bifurcações e o sistema entrará em um caos determinístico em $r > 4$, com a população evoluindo sem um padrão regular.

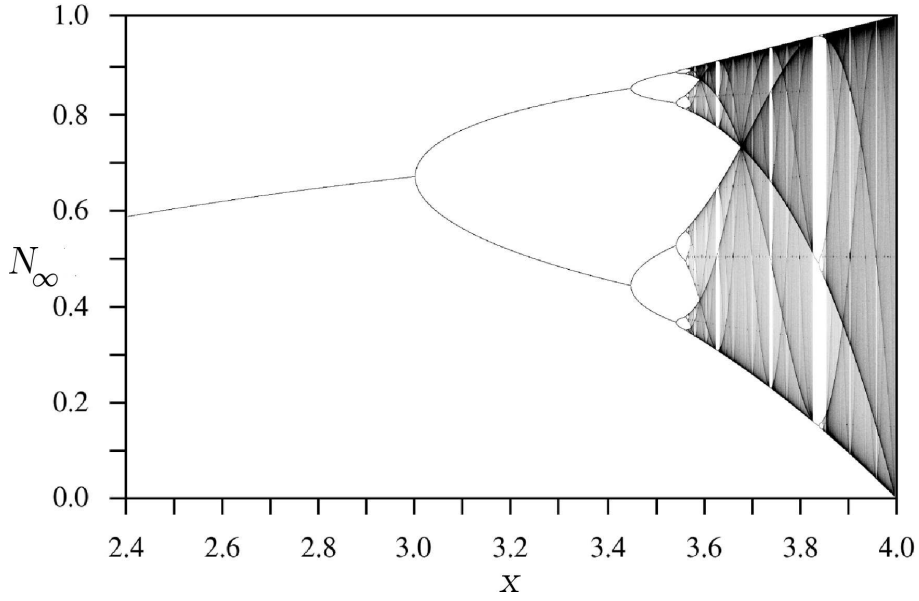


Fig. 2.1: Diagrama de bifurcação para a equação de diferença logística.

Uma terceira característica de sistemas auto-organizáveis é a coexistência de múltiplos estados estáveis, ou multi-estabilidade. Dada a presença de mecanismos de ampliação de flutuações, qualquer pequeno desvio pode ser significativamente aumentado, fazendo o sistema convergir para diferentes estados estáveis, dependendo das condições de contorno dadas.

A combinação da existência de bifurcações com a multi-estabilidade pode implicar relevantes efeitos evolutivos para sistemas auto-organizáveis, permitindo que estes forneçam diferentes respostas frente a pequenos ajustes de parâmetros, sejam estes parâmetros internos ou ambientais. Suponha que a seleção natural ajustou um determinado padrão de um sistema próximo a um ponto de bifurcação: em tal situação uma pequena alteração no ambiente ou nos indivíduos pode induzir grandes modificações no funcionamento global do grupo. Nesse caso, ao invés de um conjunto de complexos modelos comportamentais embutidos em cada indivíduo para determinar o funcionamento do sistema em diferentes cenários, os elementos podem valer-se de um mesmo procedimento para situações levemente diferentes, encarregando as próprias não-linearidades do sistema em prover uma resposta diversa. Um exemplo pode ser visto no trabalho de Deneubourg et al. [47], no qual foram analisados modelos em que um mesmo conjunto de regras comportamentais e fisiológicas eram utilizadas por

várias espécies de formigas-correição, porém aplicadas a diferentes condições ambientais. O estudo apontou que diferentes padrões de expedições surgiam apenas variando a distribuição inicial de alimentos, demonstrando que diferenças no funcionamento de grupos distintos e em ambientes diversos não são necessariamente resultado de diferenças genotípicas geradas evolutivamente para adaptação ao meio.

O uso de uma mesma dada regra para várias situações permite a codificação compacta de comportamentos globais refinados e flexíveis. Isso possibilita que o DNA possua menos informação que a necessária para a descrição de um indivíduo adulto, uma vez que é preciso apenas que eles contenham as regras que geram tal informação [124]. Essa economia pode ser observada no sistema imunológico: um ser humano é capaz de sintetizar mais de 100 milhões de proteínas de anticorpos, mesmo possuindo apenas cerca de 100 mil diferentes genes.

Ao contrário do que a ocorrência de bifurcações possa indicar, a maior parte dos sistemas auto-organizáveis são extremamente robustos, sendo estáveis por um amplo espectro de parâmetros e pouco sucetíveis a perturbações. Apesar de a seleção natural ajustar certos parâmetros para perto de limites de bifurcação, a maioria deles opera longe destes pontos [35]. Assim, tais comportamentos tornam-se mais estáveis, sendo capazes de reparar-se uma vez que o padrão estabelecido é um atrator do sistema.

2.3 Conclusão do capítulo

Conforme apresentado, a auto-organização é uma forma de um sistema atingir a ordem através apenas de relações internas a este e na ausência de referências a quaisquer padrões globais, estando presente em diversas áreas, como a física, a economia e a biologia. Esta dinâmica demanda poucos recursos dos componentes do sistema, assim como possui baixas exigências de comunicação, uma vez que baseia-se na comunicação local entre os indivíduos e com o ambiente. A ordem auto-organizada mostra-se extremamente robusta e flexível, permitindo que diferentes respostas sejam dadas a cenários diversos, mesmo sem uma codificação explícita nas regras do sistema.

Algo importante a ressaltar é que as interações não-lineares simples entre os elementos podem levar a complexos padrões em nível de grupo, padrões estes muitas vezes inesperados mesmo tendo-se em mãos o conhecimento completo sobre o funcionamento dos indivíduos e suas interações. Tal dificuldade deve ser levada em conta no projeto de sistemas artificiais que se valham da auto-organização, caso dos métodos de inteligência de enxame.

Capítulo 3

Otimização por Enxame de Partículas

Conforme mostrado no capítulo 2, na natureza podem-se observar alguns exemplos marcantes de auto-organização, onde a cooperação entre diversos animais permite a realização de tarefas complexas, como a habilidade que formigas possuem de encontrar fontes de suprimentos mais próximas à colônia [17], e de, assim como abelhas [18], chegar a um consenso acerca de um novo local para o estabelecimento de um ninho, com base na qualidade das localizações disponíveis [121]. Segundo o sociobiólogo Edward Osborne Wilson [182], a ocorrência de tais comportamentos em múltiplas espécies sugere que os benefícios do compartilhamento social das informações obtidas até o momento entre indivíduos de um mesmo enxame ou cardume, por exemplo, superam os custos decorrentes da competição por alimento, constituindo-se uma relevante vantagem evolutiva. Com essa hipótese em mente, torna-se tentadora a idéia de utilizar, em certos ramos da Engenharia, processos e modelos inspirados nesses comportamentos, valendo-se dos milhões de anos através dos quais eles têm sido testados na natureza. Na computação, esse pensamento se concretiza através da inteligência de enxame, que tem como principal idéia o projeto de sistemas multi-agentes cujos comportamentos sejam baseados no funcionamento de sociedades de animais, projetando unidades simples de processamento de informação que, através de intensa comunicação entre si e interação com o "ambiente", cooperam para a obtenção de um objetivo comum. Dentro dessa disciplina da inteligência artificial, podemos explicitar algumas frentes de maior destaque, como a robótica de enxame [51], as técnicas de otimização e classificação por colônia de formigas [50, 161], além dos sistemas de otimização por enxame de partículas, sendo esta última a técnica que será abordada neste capítulo.

3.1 Otimização por Enxame de Partículas

A otimização por enxame de partículas (ou *Particle Swarm Optimization* – PSO) foi proposta inicialmente em 1995 por James Kennedy e Russell Eberhart [61, 101], como um modelo para otimi-

zação de funções contínuas não-lineares. O PSO constitui-se de uma heurística populacional derivada de modelos do comportamento coletivo de animais e insetos sociais na coordenação de seus movimentos em tarefas como busca e obtenção de alimentos [82, 155]. Tais modelos foram simplificados, perdendo a exigência de manter uma distância mínima de seus vizinhos e transformando a arquitetura de comunicação, a qual era inicialmente baseada em proximidade espacial e foi alterada para usar uma topologia bem definida por um grafo. Dessa forma, o PSO acaba por possuir, atualmente, mais semelhanças com modelos de influência mútua entre seres humanos em suas formas de pensar e agir [8, 96].

Cada enxame do PSO é constituído por uma quantidade fixa de agentes, chamados de partículas, e estruturado de forma que cada agente seja capaz de comunicar-se com um subconjunto de seus pares – chamados de vizinhos –, o qual pode ser definido estática ou dinamicamente. Cada partícula movimenta-se pelo espaço de soluções com uma determinada velocidade, avaliando a cada iteração a solução correspondente à posição ocupada. Essa velocidade deve sofrer influência das experiências passadas próprias (fator cognitivo) e de seus vizinhos (fator social), influências essas que são usualmente implementadas como dois atratores, o primeiro localizado na melhor posição já avaliada pela partícula em questão e o segundo localizado na melhor posição visitada pelas partículas adjacentes. Como resultado dessa dinâmica, ocorre que o enxame, inicialmente disperso por todo o espaço de busca, passe a focar em determinadas regiões que apresentaram soluções mais promissoras com o decorrer da execução do algoritmo. Isso permite que o PSO seja capaz de refinar boas soluções já encontradas mas sem perder a capacidade de realizar uma busca mais ampla no espaço. Observa-se que, ao contrário de alguns métodos tradicionais da inteligência artificial, um enxame de partículas avalia uma solução inteira por vez, não fazendo avaliações da contribuição de cada parte para a solução como um todo. Note, também, que uma partícula isolada possui uma limitada capacidade de otimização, sendo altamente propensa a ficar presa em ótimos locais. Dessa forma, o poder de otimização do enxame não é apenas a soma das capacidades individuais de cada partícula, emergindo das interações das partículas.

A despeito de ser uma meta-heurística relativamente recente, a otimização por enxame de partículas conseguiu granjear um relevante espaço [150] devido à sua simplicidade e poder de processamento. Ao contrário de outras técnicas, o PSO não requer informações de gradiente da função objetivo, necessitando apenas de seus valores nos pontos avaliados e de operadores matemáticos simples. Sendo sistemas auto-organizados, os paradigmas de inteligência de enxame levam a processos altamente distribuídos, com baixa dependência de indivíduos isolados, de forma que o PSO é um modelo de otimização relativamente robusto contra a atração exercida por regiões de baixa qualidade sobre alguns agentes. O fato deste mecanismo centrar-se na ação de agentes baseados em regras muito simples, leva a uma facilidade em sua implementação e também a baixas exigências de me-

mória e poder de processamento [64, 144]. Estas baixas exigências permitem que cada iteração seja executada em pouco tempo, levando a uma rápida convergência, decorrente também da intensa troca de informação entre os agentes.

Entrada: ω, c_1, c_2, N e vizinhança
Inicializar a posição x_i, a velocidade v_i e a melhor posição pessoal p_i das N partículas;
enquanto *o critério de parada não é satisfeito* **faça**
  para $j \leftarrow 1$ **até** N **faça**
    $g \leftarrow \arg\min_{i \in vizinhos(j)} f(p_i)$;
    $v_j \leftarrow \omega v_j + c_1 r_1 (p_j - x_j) + c_2 r_2 (p_g - x_j)$;
    $x_j \leftarrow x_j + v_j$;
    se x_j *está dentro do espaço de busca e* $f(x_j) < f(p_j)$ **então**
      $p_j \leftarrow x_j$;
    fim
  fim
fim

Algoritmo 1: Otimização por Enxame de Partículas para minimização

No pseudo-código do PSO canônico, exibido em sua versão para minimização no algoritmo 1, r_1 e r_2 são duas funções que retornam valores aleatórios uniformemente amostrados no espaço $[0, 1]$, e c_1 e c_2 são coeficientes de aceleração que definem se o enxame atuará de forma mais coletiva ou mais individualizada e N é o número de partículas no enxame. O primeiro elemento da equação de atualização da velocidade, ωv_j, pode ser interpretado como o momento linear da partícula reduzido por uma força de atrito, funcionando de forma análoga à mutação nos algoritmos de computação evolutiva, com a diferença de que no PSO essa mutação seria direcional. É importante apontar que o fator ω não estava presente na versão original do algoritmo, sendo introduzido posteriormente como método de evitar que as partículas ganhassem velocidades tão altas que tornassem impossível a convergência [163]. Um método alternativo para evitar um comportamento explosivo do enxame, mas que pode ser usado em conjunto com o fator de inércia, é a definição de limites de velocidade $[-V_{max}, V_{max}]$ em cada dimensão para as partículas. Os arranjos paramétricos ideais dependem do problema abordado. No entanto, valores usuais são $\omega = 0,7298$ e $c_1 = c_2 = 1,49618$, com uma população entre 20 e 50 partículas [150].

No PSO, verifica-se que o balanço entre exploração e busca local, crucial para um bom algoritmo de otimização, é realizado através de um consenso entre um indivíduo e seus vizinhos [100], sendo o tamanho do passo limitado em cada dimensão pela soma do fator de inércia da partícula às suas distâncias aos atratores pessoal e social. Dessa forma, se um conjunto de partículas conectadas obtiver sucesso em uma determinada região, ele se concentrará nesta porção do espaço e permitirá uma fase de maior explotação; caso contrário, se o grupo encontrar-se disperso pelo espaço de busca, a partícula

deverá vagar sem uma tendência definida.

3.2 Modificações

Apesar de sua capacidade de busca, a otimização por enxame de partículas apresenta algumas deficiências. Como exemplo pode ser citada a tendência do enxame a prematuramente perder diversidade e convergir a um ótimo local [150], provocada em parte pela própria rapidez do algoritmo. Diversas alternativas foram apresentadas no intuito de melhorar o PSO, propondo modificações em seus três elementos básicos: sua topologia, seu modelo de influência e sua regra de atualização [100]. Algumas das principais idéias são apresentadas a seguir.

Modelo de influência

A escolha de quais as informações, dada uma vizinhança, serão utilizadas para atualizar a posição da partícula é definida pelo modelo de influência, ou regra de interação. No PSO canônico utiliza-se um modelo denominado *best-of-neighborhood*, onde apenas a própria partícula e o vizinho com melhor resultado são usados como informantes. Mendes [127] propõe uma modificação, denominada *Fully Informed Particle Swarm* (FIPS), onde uma partícula utiliza as informações disponibilizadas por todos os vizinhos, generalizando a equação da velocidade para:

$$v_j \leftarrow \omega v_j + \frac{\sum_{i \in vizinhos(j)} W(i)\phi_i(p_i - x_j)}{\sum_{i \in vizinhos(j)} W(i)\phi_i} \tag{3.1}$$

$$\phi_i = r[0, \frac{c_{max}}{|vizinhos(j)|}] \tag{3.2}$$

A função $W(i)$ é uma função de ponderação levando em conta algum aspecto da partícula que possa ser considerado relevante e $r[a, b]$ é um número aleatório entre a e b, obtido a partir de uma distribuição uniforme. Experimentos indicaram que o FIPS é capaz de prover resultados satisfatórios em menos iterações, mesmo utilizando um valor de retorno fixo para a função W [127].

O agrupamento das partículas, utilizando um algoritmo de k-medianas, foi proposto por Kennedy [98]. No trabalho em questão, foi feita a análise do efeito da substituição das influências pessoal e social pelos centróides dos *clusters* ao qual pertencem a partícula e o melhor vizinho, respectivamente. A versão que provou-se a mais efetiva foi aquela em que apenas o fator individual foi substituído, obtendo bons resultados. No entanto, é importante notar que a execução do algoritmo de agrupamento adiciona um tempo significativo ao mecanismo de busca.

Regra de atualização

A regra de atualização define como as influências de uma partícula devem ser combinadas para calcular o seu próximo passo. Uma modificação bem conhecida da regra de atualização é o uso de um coeficiente de constrição [39], o qual busca evitar a explosão das velocidades do enxame. A equação da velocidade é, então, modificada para:

$$v_j \leftarrow \chi(v_j + \phi_1 r_1(p_j - x_j) + \phi_2 r_2(p_g - x_j)), \tag{3.3}$$

onde $\phi = \phi_1 + \phi_2 > 4$ e

$$\chi \leftarrow \frac{2}{\phi - 2 + \sqrt{\phi^2 - 4\phi}}. \tag{3.4}$$

Quando este método é utilizado, normalmente $\phi = 4,1$ e $\phi_1 = \phi_2$. Tal modelo é algebricamente equivalente à versão do PSO com fator de inércia, com $\omega = \chi$, $c_1 = \chi\phi_1$ e $c_2 = \chi\phi_2$.

Outra modificação, a qual busca fornecer um controle mais profundo da alternância das fases de exploração e de busca local, é a utilização de um fator de inércia ω variante no tempo [62, 163]. Uma forma de variação usualmente encontrada na literatura [62, 163] é o decréscimo linear de tal parâmetro de $0,9$ até $0,4$.

Alguns trabalhos propõem a supressão do uso do termo de velocidade. Um caso é o *Bare Bones Particle Swarm*, proposto por Kennedy [99], onde a nova posição da partícula é definida através de uma distribuição gaussiana N centrada na média entre a influência pessoal e a influência social, com desvio definido pela distância entre tais pontos:

$$x_j \leftarrow N(\frac{p_j + p_g}{2}, p_j - p_g) \tag{3.5}$$

Esta regra de atualização foi inspirada a partir da análise da distribuição das posições exploradas pelas partículas em torno de suas influências no PSO canônico, a qual foi verificada ter a forma de sino, assim como uma distribuição normal. Outras distribuições, que não a gaussiana, também podem ser utilizadas [156].

Para tentar superar o problema de queda prematura da diversidade, o AR-PSO [157] propõe alternar fases de atração e de repulsão entre as partículas. Assim, se a métrica de diversidade do enxame cai abaixo de um limite, as partículas passam a repelir-se de suas influências, voltando a seu comportamento normal apenas quando tal diversidade atinge um limitante superior. Os dados obtidos no trabalho em questão indicam que o AR-PSO é capaz de obter resultados melhores que o PSO, apesar de, assim como este, também ter uma rápida convergência.

Topologia

No PSO, a topologia define a vizinhança das partículas, sendo usada para controlar o fluxo das informações no enxame e podendo ser estática ou dinâmica. A variação de topologia gera diferentes comportamentos das partículas, de forma que enxames densamente conectados costumam convergir rapidamente, obtendo bons resultados em problemas mais simples, enquanto que vizinhanças esparsas permitem que as subpopulações convirjam de forma independente, trabalhando melhor em problemas complexos [127].

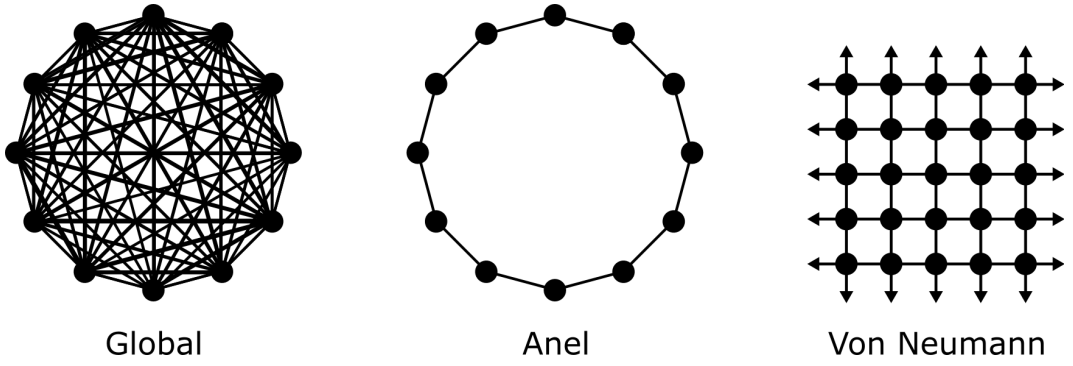


Fig. 3.1: Algumas topologias utilizadas no PSO.

Inspiradas no movimento de pássaros, as versões iniciais do PSO usavam uma vizinhança definida pela proximidade espacial, em uma topologia dinâmica. Entretanto, esse método mostrou-se por demais custoso computacionalmente, sem oferecer boas qualidades de convergência. Desta forma, o algoritmo padrão utiliza atualmente uma topologia estática denominada global ou *gbest* [101], onde o enxame é totalmente conectado, de forma que é necessário manter apenas um atrator para todas as partículas do enxame. Outra topologia amplamente difundida é a *lbest* [61], ou em anel, segundo a qual cada indivíduo conecta-se a apenas outros dois elementos, formando um ciclo.

Alguns estudos dedicaram-se à avaliação dos efeitos de diferentes topologias estáticas sobre o desempenho do PSO canônico. Kennedy [97] analisou topologias em anel e em estrela (um nó central, conectado a todos os outros) com variadas intensidades do efeito *small world* [179], comparando também com vizinhanças aleatórias [66] e *gbest*. Foram verificadas diferenças significativas de acordo com a topologia utilizada. Porém tais variações eram altamente dependentes da função a ser otimizada, não sendo possível obter resultados conclusivos.

Uma análise mais ampla foi conduzida por Kennedy e Mendes [103, 104], tendo avaliado topologias *gbest*, em anel, em pirâmide e von Neumann (exibida na figura 3.1), além de redes aleatórias e de uma rede projetada manualmente com agrupamentos e heterogeneidades. Foi também estudado o efeito da inclusão, ou não, de uma partícula em sua própria vizinhança. Os resultados experimentais não apontaram nenhuma topologia que superasse todas as outras em todas as métricas utilizadas.

Porém, foi possível observar que a topologia von Neumann mostrou um desempenho geral superior. Verificou-se, também, que a presença de uma partícula em sua própria vizinhança não causa grande impacto sobre o resultado da otimização.

A escolha do modelo de vizinhança possui diferentes impactos em cada dinâmica do PSO [104]. No PSO canônico, como afirmado anteriormente, a topologia *gbest* favorece uma rápida convergência, isso porque todas as partículas sofrem influência direta daquela que obteve melhor resultado até o momento. No FIPS, porém, uma vizinhança global tende a prejudicar a chegada a um resultado uma vez que cada membro do enxame não será influenciado apenas por seus melhores vizinhos, mas também por outras partículas que se encontram em regiões muito distintas, incluindo alguns locais de baixa qualidade. Dado seu funcionamento, um enxame do FIPS tende a obter melhores desempenhos em vizinhanças menos conectadas, sendo significativamente sensível a modificações na topologia. Em todo o caso, o uso de uma vizinhança von Neumann proporciona bons resultados ao FIPS, superando aqueles obtidos pelo PSO canônico.

Dadas as diferenças evidenciadas anteriormente entre enxames com mais ou menos conexões, Suganthan [168] propõe um modelo em que as conexões são estabelecidas com base na distância euclidiana entre as partículas no espaço de busca, sendo ligados os pares de partículas cujas distâncias se encontrem dentro de um determinado limiar. A cada iteração esse limiar é aumentado, de forma que a vizinhança, inicialmente esparsa, torna-se em determinado momento uma topologia *gbest*. Apesar de os dados experimentais apontarem alguns resultados positivos em comparação com o PSO original, essa metodologia é extremamente custosa computacionalmente, uma vez que a topologia é refeita a cada iteração, exigindo o cálculo da distância entre todas as partículas.

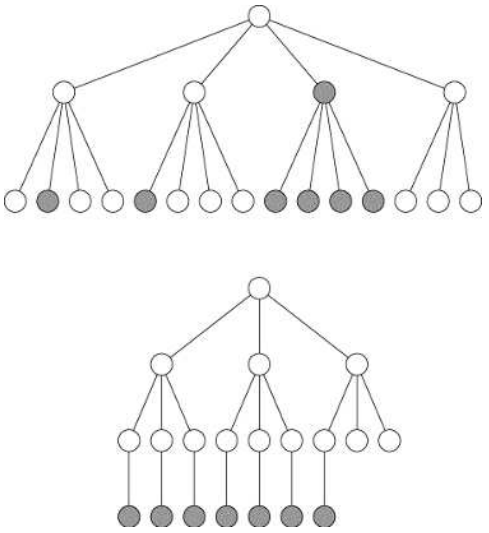


Fig. 3.2: Exemplo da adaptação topológica no H-PSO, de $d = 4$ para $d = 3$ [90].

Uma abordagem hierárquica é proposta por Janson e Middendorf [90]. No H-PSO o enxame

organiza-se em uma árvore regular onde, a cada iteração, uma partícula deve subir um nível caso ela possua um histórico melhor que seu pai, em um arranjo onde partículas com melhores resultados ganham maior influência. O grafo de vizinhança é direcionado, com cada partícula sendo influenciada apenas pelos melhores resultados pessoal e da partícula imediatamente superior na hierarquia, exceto para o nó raiz, o qual não sofre influências externas. Ainda na proposta do H-PSO, foi sugerida uma versão adaptativa (AH-PSO), na qual a árvore sofre modificações para que, com o decorrer do algoritmo, esta ganhe níveis, ao mesmo passo em que sua ramificação d vai diminuindo (ver figura 3.2). Os testes realizados indicaram melhoras na maioria das funções testadas, tanto para o H-PSO quanto para o AH-PSO. Apesar de a versão adaptativa ter obtido resultados melhores, esta mostrou-se muito sensível a parâmetros, enquanto um determinado arranjo do H-PSO foi capaz de obter bons resultados para todas as funções de teste.

O *Fitness-Distance-Ratio based PSO* (FDR-PSO) [147] propõe o uso concomitante de duas vizinhanças, uma global fixa (*gbest*) e outra dinâmica. A vizinhança dinâmica é definida independentemente para cada dimensão, conectando cada partícula i à vizinha j que maximiza a razão entre a diferença de valores de função objetivo e a distância da projeção das partículas na dimensão d analisada, conforme a equação:

$$FDR(i, j, d) = \frac{f(x_i) - f(p_j)}{|p_{jd} - x_{id}|} \tag{3.6}$$

A equação da velocidade é, então, modificada para considerar a influência da nova vizinha p_{fdr}:

$$v_j \leftarrow \omega v_j + c_1 r_1 (p_j - x_j) + c_2 r_2 (p_g - x_j) + c_3 r_3 (p_{fdr} - x_j) \tag{3.7}$$

O FDR-PSO busca, assim, reduzir a concentração de partículas em uma única região, permitindo uma maior exploração do espaço. Os testes realizados com o algoritmo buscaram avaliar também o efeito da supressão e da definição de diferentes pesos para cada uma das influências presentes na equação de atualização da velocidade. Foram reportadas melhorias significativas em relação ao PSO canônico em todas as versões testadas do FDR-PSO, especialmente naquelas em que todas as influências estavam presentes, com maior peso para aquela advinda da topologia dinâmica. Porém, é preciso notar que, a exemplo do que ocorre em outras versões que consideram a distância entre as partículas, a modificação proposta apresenta custo quadrático no tamanho do enxame para a definição da nova topologia.

A estruturação do enxame em sub-grupos também foi alvo de alguns trabalhos [38, 114]. O *Dynamic Multi-Swarm Particle Swarm Optimizer* (DMS-PSO) [114] particiona aleatoriamente as partículas em pequenos grupos totalmente conectados e com tamanhos iguais, os quais são refeitos após uma determinada quantidade de iterações. Tal dinâmica busca permitir que cada sub-enxame convirja

a um ótimo local diferente, modificando a topologia assim que esta não puder fornecer novas informações às partículas. No final da busca, após identificadas algumas regiões promissoras, o resultado é refinado através do agrupamento das partículas em um único enxame. Os resultados indicam que sub-enxames pequenos, com três partículas, obtêm bons resultados em diversos problemas de teste, em comparação com outras versões do PSO, superando o FDR-PSO na maioria dos casos.

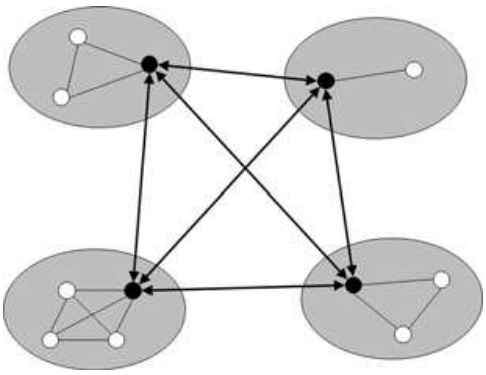


Fig. 3.3: Uma topologia do TRIBES. Cada tribo encontra-se delimitada pelas áreas em cinza. As partículas em preto são os xamãs da tribo [43].

Clerc [38] propõe o TRIBES, uma versão adaptativa do PSO onde o tamanho e a estrutura da população evoluem no decorrer do algoritmo, de acordo com o seu desempenho. O enxame é dividido em "tribos" de tamanhos diferentes e variáveis, as quais possuem uma estrutura interna totalmente conectada. Todas as tribos possuem comunicação direta entre si, a qual é realizada pelas suas melhores partículas (chamadas "xamãs"). Um exemplo de topologia do TRIBES pode ser visto na figura 3.3. A cada $NL/2$ iterações (sendo NL a quantidade de conexões na rede), a topologia sofre atualizações, com a adição e remoção de partículas. Essa dinâmica baseia-se na separação das tribos entre boas e ruins: uma tribo é considerada ruim se nenhuma de suas partículas conseguiu alguma melhoria na última iteração. Caso contrário, ela pode ser chamada boa ou ruim, com 50% de probabilidade. Cada tribo boa terá eliminada sua partícula com pior histórico de busca, exceto em tribos com apenas um elemento, onde a partícula só é removida se seu desempenho for pior que o desempenho do melhor xamã externo. As tribos ruins, por sua vez, gerarão uma quantidade de novos elementos, de acordo com a equação a seguir, determinada empiricamente [37]:

$$NB_{particles} = max(2, \lfloor \frac{9,5 + 0,124(D-1)}{tribeNb} \rfloor), \tag{3.8}$$

onde D é a dimensão do espaço de busca e $tribeNb$ é o número de tribos no enxame. As partículas geradas são, então, agregadas em uma nova tribo, sendo seu xamã conectado aos demais. Os novos elementos são criados aleatoriamente por todo o espaço de busca ou em uma porção limitada deste, definida de acordo com a tribo geradora. O enxame é iniciado com apenas uma partícula em uma

única tribo, crescendo e se organizando, de acordo com as regras estabelecidas. Desta forma, não há no TRIBES a necessidade de quaisquer ajustes paramétricos relativos à topologia. No mesmo trabalho também foram propostas algumas novas regras de atualização, com a supressão do termo de velocidade e a seleção da nova posição a ser visitada segundo uma distribuição aleatória uniforme. As modificações mostraram bons resultados nos experimentos realizados por Clerc [38].

Outras propostas

Uma última proposta a ser abordada é a adaptação do PSO para a otimização de problemas binários discretos, feita por Kennedy e Eberhart [102]. Nesta versão a fórmula clássica do cálculo da velocidade da partícula mantém-se intacta, mas o cálculo da nova posição da partícula é modificado para que $x_i^j \leftarrow 1$ com probabilidade $S(v_j^i)$, onde S é uma transformação logística. Desta forma, a partícula navega nas arestas do hipercubo limitado por $[0, 1]$ em cada dimensão. É importante notar que há algumas modificações substanciais nas idéias originais do PSO, não sendo mais a população de partículas composta de potenciais soluções, mas sim de um conjunto de probabilidades.

Além das modificações apresentadas, diversas outras idéias foram propostas na literatura, como a adaptação do fator ω utilizando um sistema *fuzzy* [62], o uso de mutação gaussiana [84] e versões para problemas dinâmicos [22] e multi-objetivo [154].

3.3 Aplicações

As aplicações da otimização por enxame de partículas podem ser divididas, basicamente, em três grupos: o primeiro busca explorar as capacidades do algoritmo, adaptando-o quando necessário, o segundo trabalha com a adaptação de problemas para o uso do PSO, enquanto que o terceiro, menos comum, utiliza a metáfora da inteligência de enxame para tentar compreender o comportamento dos indivíduos em grupos [11].

Um dos primeiros trabalhos envolvendo o uso do PSO para a resolução de problemas práticos o utilizou para o treino de um *perceptron* com múltiplas camadas, em substituição ao algoritmo de *back propagation* [64], apresentando bons resultados e rápida convergência quando confrontado com outros métodos de otimização. Uma comparação de resultados obtidos por redes neurais para o reconhecimento de faces foi realizada em [169], de forma que o enxame de partículas mostrou resultados competitivos, porém sendo superado por técnicas de classificação mais específicas.

Outra frente em que o PSO já foi aplicado é o ajuste de sistemas de recomendação – ferramentas que fazem recomendações aos usuários, com base em seu histórico e em propriedades de produtos, comumente usados em sites de comércio eletrônico ou de músicas online. Ujjin e Bentley [173]

compararam o desempenho do enxame de partículas com o de algoritmos genéticos e de uma técnica não-adaptativa baseada no algoritmo de Pearson, de forma que o PSO forneceu resultados mais precisos na maioria dos casos, sendo significativamente mais rápido quando comparado algoritmo genético. Outras aplicações populares do PSO na engenharia são o desenho e modelagem de antenas [94] e o projeto de sistemas de controle [40, 73, 129, 187], havendo usos do algoritmo também no projeto de redes [93], de circuitos VLSI [63, 146], na área de robótica [83, 112] e na evolução de programas de computador [140]. Problemas clássicos da otimização combinatória, como o caixeiro viajante [141], o problema da mochila [116], o problema de satisfabilidade [1] e o roteamento de veículos [190], também já foram objeto do uso da otimização por enxame de partículas.

Dentro dos estudos do PSO, outras áreas bastante exploradas são a biologia e a medicina, sendo o algoritmo usado para a inferência de redes gênicas [186], a reconstrução de árvores filogenéticas [118] e o diagnóstico de doenças [60, 185]. O enxame de partículas também foi aplicado na cinesiologia para a otimização do modelo biomecânico de um chute em uma bola de futebol [105]. No trabalho em questão foram simulados 17 grupos musculares envolvidos, resultando em um problema de busca em 56 dimensões que inclui restrições, como a penalização quando os dedos do pé atingem o chão. Os resultados obtidos foram comparáveis àqueles de pesquisas realizadas por fabricantes de materiais esportivos.

Uma versão híbrida do PSO com técnicas estatísticas foi utilizada por Ko e Lin [108] no ajuste dos parâmetros de um mecanismo de regressão para a realização de predições de séries financeiras. Uma melhoria de desempenho foi verificada com o uso deste algoritmo, quando comparado com versões puramente estatísticas ou baseadas em algoritmos genéticos.

Uma aplicação inusitada de enxames de partículas é a improvisação musical [21]. No trabalho em questão o enxame é visto como uma partitura a ser interpretada, com cada partícula representando uma nota e uma duração, de acordo com sua posição no espaço.

O uso do PSO como metáfora social foi feito no estudo da influência da inércia organizacional [27], onde foi modelada a adaptação de estratégias em grandes companhias frente a um cenário dinâmico, concluindo que a combinação de alguma inércia com a adoção de um certo elitismo, o qual garanta que apenas movimentos de melhoria sejam realizados, é capaz de previnir mudanças organizacionais desnecessárias.

Uma lista mais completa de aplicações pode ser vista em [149], onde foram listados e analisados mais de 650 trabalhos envolvendo o uso da otimização por enxame de partículas.

3.4 Conclusão do capítulo

A otimização por enxame de partículas, técnica originalmente inspirada no movimento de revoadas de pássaros em busca por alimentos, apresenta uma série de características extremamente desejáveis em mecanismos de otimização, como baixas exigências computacionais, robustez, flexibilidade e rápida convergência. A despeito de ser uma técnica recente, quando comparado a meta-heurísticas similares, o PSO conseguiu alcançar uma grande relevância, sendo seu uso explorado em uma ampla gama de problemas, da otimização combinatória à composição musical.

Exemplo do uso da auto-organização na computação, o PSO tem sua força advinda de uma intensa interação de agentes muito simples e de uma dinâmica de realimentação que garante uma atração das partículas por regiões promissoras do espaço, ao mesmo passo que o fator de inércia provê uma certa divergência, que garante uma exploração mais efetiva. Apesar de possuir algumas deficiências conhecidas, como a convergência prematura, as quais podem ser particularmente impactantes na otimização de certos problemas, diversos trabalhos buscaram adaptações que sanassem tais limitações, esforço esse que é facilitado pela simplicidade do algoritmo.

Capítulo 4

Redes complexas

Imagine uma fonte despejando um fluxo contínuo de areia sobre uma superfície plana. De início, a pilha terá pouca altura, com a ocorrência de pequenos deslizamentos locais à medida que a altura e a inclinação começam a aumentar. Conforme a pilha atinge proporções maiores que sua base, deslizamentos mais intensos passarão a ocorrer, envolvendo não mais apenas uma determinada vizinhança, mas atingindo toda a pilha e de forma generalizada. Tal deslizamento persiste até o surgimento de uma forma de organização, na qual o sistema chega a um estado estacionário com a pilha possuindo uma inclinação constante. Esta situação aparentemente simples, é um exemplo típico de sistema complexo [36, 48] onde, apesar de ser possível conhecer relativamente bem propriedades como massa e formato dos grãos de areia, não é possível modelar o comportamento global do sistema e prever a ocorrência e a extensão dos deslizamentos, sendo tais situações emergentes de grandes quantidades de interações de um grão com seus vizinhos. A dificuldade de inferir o comportamento global do sistema a partir da decomposição e análise de suas partes é uma das características de sistemas complexos, podendo ser observada em algumas da principais linhas de pesquisa atuais na área, como os estudos da emergência da consciência a partir das interações de neurônios e destes com o ambiente, do papel do DNA na coordenação dos processos em células vivas e de como as interações de agentes influenciam o mercado financeiro, além dos diversos exemplos comentados no capítulo 2.

É difícil estabelecer uma definição suficientemente ampla sobre o que é um sistema complexo. Porém, de maneira simplificada, pode-se dizer que um sistema complexo é um sistema com grande quantidade de elementos capazes de interagir entre si e com o ambiente, evoluindo com o tempo. Essa interação pode se dar apenas em um âmbito local, de um elemento com seus vizinhos, ou envolvendo componentes distantes no sistema, podendo este ser composto por partes idênticas ou diferentes entre si. A principal característica que permeia todos os sistemas complexos é sua auto-organização e a emergência de comportamentos, com o todo sendo mais do que a soma de suas partes [5].

Um fator intrínseco a sistemas complexos é sua adaptabilidade e a plasticidade das relações entre

os elementos. Assim, o número de seus componentes não é um fator crítico, podendo tais sistemas perderem algumas partes sem prejuízo de sua funcionalidade. Uma situação fundamentalmente diferente ocorre em um sistema complicado como um avião. Neste também há grande quantidade de elementos constituintes – em um Boeing 747-400 há algo como $3 * 10^6$ peças –, porém cada parte possui funções e interações bem-definidas, podendo o conjunto exibir uma quantidade bem limitada de respostas a mudanças ambientais, de modo que um defeito em uma peça chave pode comprometer todo o sistema [142].

Atualmente, o ferramental utilizado no estudo de sistemas complexos é suportado por três campos principais, a saber: dinâmica não-linear, física estatística e teoria de redes complexas [5]. A dinâmica não-linear, tendo como caso particular a teoria do caos, busca explicar como uma unidade simples é capaz de prover uma saída complexa, tendo como um dos exemplos mais comuns deste tipo de dinâmica a equação logística, comentada no capítulo 2. Esta técnica é atualmente utilizada em múltiplas áreas do conhecimento, como geofísica, engenharia de controle e neurofisiologia [76]. Por sua vez, a física estatística, junto da relatividade e da mecânica quântica, foi responsável por uma revolução no estudo da Física no início do século XX, introduzindo o conceito de modelos discretos, como modelos Ising, autômatos celulares e modelos baseados em agentes [65, 184].

Dos campos citados anteriormente, a teoria de redes complexas, que será objeto de estudo neste capítulo, é o mais recente. Apesar disso, esta é uma área de suma importância para a compreensão de sistemas complexos, dado que a estrutura subjacente às interações altera o funcionamento global do conjunto. O estudo da natureza da topologia de um sistema complexo é um dos grandes desafios desta área, dado que tais interações se dão frequentemente de forma não-linear e são sujeitas a influências de componentes estocásticos e ruídos externos, podendo a rede de interações ser desconhecida em sua maior parte. A despeito disso, porém, é comum em muitas disciplinas a modelagem de sistemas complexos valendo-se de estruturas em redes [4, 24, 55, 139], mapeando os elementos como vértices e as interações deles como arestas. Nas Ciências Sociais, por exemplo, esta ferramenta é utilizada para o estudo dos relacionamentos sociais entre humanos, criando redes de colaboração científica [14, 136], amizade [152] e contato sexual [95, 115], entre outras. Na Biologia, tais estruturas estão presentes na ecologia, na interação de espécies e evolução em grupos sociais [130, 148], assim como na genômica e proteômica, particularmente na investigação das redes de interações gênicas [162]. Redes são ainda mais frequentes na engenharia, com seu uso indo da estrutura do *backbone* da Internet [70] à representação de um sistema de distribuição elétrica [179], incluindo modelos de comunicação entre múltiplos robôs [143].

A modelagem de sistemas complexos empregando teoria de redes tem ganhado significativa relevância nas últimas décadas, especialmente devido à crescente capacidade de processamento dos computadores modernos. Isto permitiu que fosse percebido que algumas destas redes compartilham

não apenas uma representação, mas também uma série de características não-triviais que não são exibidas nem por redes regulares nem por redes plenamente aleatórias. Uma questão que surge com estas qualidades inesperadas é a razão pela qual elas são tão comuns. Elas contribuem para a robustez e a continuidade do processo coletivo ao qual elas se relacionam, fornecendo alguma vantagem competitiva aos sistemas que as usam? Ou talvez haja alguma característica subjacente comum aos processos que guiam o crescimento de redes dessa classe, resultando nas propriedades observadas? Fato é que essas características garantem um conjunto de efeitos interessantes, como a rápida dispersão de informações ou matéria pelos nós [179], mesmo em redes com conexões esparsas, além de alta resistência a eventos de falha, com baixa degeneração da conectividade da rede quando nós aleatórios são removidos.

4.1 Conceitos sobre grafos

Redes, como as citadas anteriormente, são formalmente estudadas na matemática através da teoria dos grafos, campo cuja concepção inicial é atribuída a Leonhard Euler, em seu trabalho sobre o problema das pontes de Königsberg [69]. Inicialmente, tal estudo concentrava-se em redes de pequeno porte e com alta regularidade, muitas vezes tendo como parte importante a análise visual das redes trabalhadas. No século XX, a teoria dos grafos passou a focar algumas redes maiores. Isso requereu um estudo mais estatístico e baseado em algoritmos. Com vistas a facilitar a leitura das seções seguintes, a seguir serão explicitados alguns conceitos de teoria dos grafos [10].

Definição 1. *Um grafo G é definido por um par de conjuntos $G = \{V, E\}$, onde V é um conjunto de vértices e E é um conjunto de arestas, ou seja, pares que relacionam dois vértices do grafo podendo as arestas possuir, ou não, um custo associado, o qual pode representar informações como a intensidade de uma conexão.*

Definição 2. *Um grafo é dito não-dirigido se seu conjunto de arestas E é formado por pares não-ordenados $\{i, j\}$, onde $i, j \in V$. Caso contrário, se o conjunto E é constituído de pares ordenados (i, j), o grafo é dito dirigido, ou* digrafo*. Uma aresta não-dirigida $\{i, j\}$ permite tanto o fluxo do vértice i com direção ao vértice j como no sentido contrário. Uma aresta dirigida (i, j), porém, tem seu fluxo saindo exclusivamente do vértice i com direção ao vértice j.*

Definição 3. *Um grafo é dito* simples *caso ele não possua laços – arestas que têm ambas as extremidades no mesmo nó – ou multi-arestas – múltiplas arestas com a mesma origem e o mesmo destino.*

Definição 4. *Um grafo pode ser particionado em* componentes conexos*, que são sub-conjuntos $U \subset V$. Um par de vértices $u, v \in V$ está no mesmo componente conexo U se, e somente se, existe*

um caminho $(v_1, ..., v_n)$*, onde* $v_1 = u$*,* $v_n = v$ *e* $(v_i, v_{i+1}) \in E$ *para todo* $1 \leq i < n$*. Em um grafo dirigido, define-se como* componente fortemente conexo *o componente conexo obtido levando em conta a direção das conexões. Quando um componente conexo de um grafo dirigido é obtido ignorando a direção das arestas ele é chamado* componente fracamente conexo*. Caso um grafo seja composto por apenas um componente conexo, o grafo é dito* conexo*.*

No decorrer deste trabalho, os termos rede e grafo, vértice e nó, aresta e conexão serão utilizados como sinônimos. Exceto em situações em que for explicitado o contrário, todos os grafos abordados neste capítulo serão grafos simples e não-dirigidos.

4.2 Características

Conforme afirmado anteriormente, devido à falta de uma análise profunda até poucos anos atrás, as redes do mundo real eram vistas como sendo muito complexas e, assim, tidas como frutos de um processo predominantemente aleatório [12]. Por esta razão, as investigações usualmente se baseavam fortemente nos modelos ER de grafos aleatórios, propostos por Erdös e Rényi [66], onde cada grafo com N nós e M arestas possui igual chance de ser produzido[1]. Com a disponibilidade de maior poder de processamento e de melhores ferramentas analíticas, porém, verificou-se que as redes do mundo real não tinham suas características contempladas plenamente por aquelas geradas pelos modelos ER ou pelos modelos de grafos regulares, como reticulados. Foi possível observar também que essas redes complexas usualmente compartilham algumas características topológicas, como pequenas distâncias médias entre os nós, alta aglomeração [179], distribuição de graus em cauda longa [13], correlações de graus e a presença de padrões recorrentes de conexões, propriedades essas que serão discutidas a seguir.

4.2.1 Graus, distribuições e correlações

Característica mais simples e mais estudada de um vértice isolado, o *grau* k_i é o número total de conexões de um vértice i. Quanto maior o grau de um nó, mais importante ele é na rede, sendo os nós com maior número de conexões chamados *hubs*. Caso o grafo seja dirigido, o grau de cada um dos nós possuirá dois componentes distintos, o grau de entrada k_i^{in}, indicando o número de arestas que chegam ao nó, e o grau de saída k_i^{out}, sinalizando a quantidade de arestas que partem do nó. Deste modo, o grau do nó é definido como $k_i = k_i^{in} + k_i^{out}$. A média de k_i em todo o grafo é chamada de

[1] Este modelo com quantidade pré-determinada de arestas é chamado de modelo G_{NM}, também sendo proposto no mesmo trabalho um modelo alternativo, denominado G_{Np}, onde há uma quantidade determinada N de nós e cada par de nós é conectado com probabilidade p. A principal diferença entre os dois modelos está no fato de que o modelo G_{Np} permite desvios na quantidade final de arestas no grafo.

grau médio, sendo denotada por $< k >$. Desta forma tem-se que, se um grafo não possui conexões com o exterior, $< k_i^{in} >=< k_i^{out} >$.

Uma forma simples de caracterizar uma rede é através de sua distribuição de nós $P(k)$, que é a probabilidade de que um nó selecionado aleatoriamente possua exatamente k arestas, ou seja, $P(k)$ indica a proporção dos nós no grafo que possuem grau k. Analogamente, em grafos dirigidos pode ser também definida a distribuição conjunta de graus $P(k_i^{in}, k_i^{out})$, assim como as distribuições de graus de entrada $P(k_i^{in})$ e de saída $P(k_i^{out})$.

Em uma rede regular – redes onde todos os nós possuem a mesma quantidade de arestas – a distribuição de graus segue uma função delta, sendo totalmente concentrada em apenas um ponto, enquanto que em uma rede plenamente aleatória tal distribuição é binomial ou de Poisson, para redes muito grandes. A cauda de uma distribuição de Poisson é exponencial, de forma que $P(k)$ assume valores muito baixos quando $k >>< k >$. Redes complexas, porém, não possuem distribuições similares nem às de redes aleatórias nem às de redes regulares, comumente seguindo uma distribuição em lei de potência, da forma $P(k) \propto k^{-\gamma}$, onde γ é uma constante positiva. Esta distribuição é uma distribuição dita de cauda longa, a qual cai mais gradualmente que uma exponencial conforme k cresce, de forma que é mais provável a existência de nós com graus mais altos. Uma vez que uma lei de potência não possui qualquer escala característica, redes que possuem uma distribuição de graus segundo esta distribuição são chamadas *redes livres-de-escala*.

É possível observar distribuições em lei de potência em várias situações, como em mapas que indicam a ocorrência de reações entre substratos no metabolismo celular [92, 177], em redes de colaboração entre cientistas [134, 135, 137], na estrutura de comunicação da Internet e na rede formada pelos *hyperlinks* entre páginas da *web* [70] e em redes de ligações telefônicas [2]. Apesar da ocorrência regular de redes livres-de-escala verifica-se em vários casos uma distribuição de graus mista, seguindo uma lei de potência com um limitante exponencial para valores muito grandes de k, caso de redes de interações de proteínas [91]. Outras redes, por sua vez, exibem uma distribuição puramente exponencial, como o sistema elétrico do estado americano da Califórnia e a rede neural do verme *Caenorhabditis elegans* [6], havendo também trabalhos apontando situações em que a distribuição de graus é gaussiana [6]. Um caso interessante é observado em redes de citações entre artigos científicos, onde os graus de entrada seguem uma lei de potência [153], mas os graus de saída distribuem-se em uma forma exponencial [175]. Outro exemplo a ser ressaltado é o das redes de colaboração entre atores, as quais podem apresentar duas distribuições diferentes: quando séries de televisão são contabilizadas, a distribuição de graus possui um limitante exponencial, enquanto que quando estas não constam na rede, a distribuição torna-se puramente livre-de-escala [6, 13]. A razão para isso é que séries são registradas como um único trabalho que, dependendo de sua duração, pode contar com milhares de colaboradores. É importante notar também que, apesar de abundarem trabalhos sobre ca-

racterísticas de redes reais, nem sempre é possível afirmar qual a forma de sua distribuição de graus. Esta situação é vista, por exemplo, em redes tróficas locais onde, devido à pouca quantidade de dados disponíveis, é possível estabelecer tanto uma regressão em lei de potência [130] quanto uma regressão exponencial [33, 34]. Na subseção 4.3, serão exibidos alguns exemplos de dinâmicas de crescimento de redes que podem levar a diferentes distribuições de graus.

Uma distribuição de graus $P(k)$ fornece uma informação completa em redes onde os graus de nós conectados não possuem correlação, ou seja, a probabilidade de um nó possuir um determinado grau k' independe do grau de seus vizinhos. Em situações reais, porém, há um grande número de redes com graus correlacionados, caso das redes silábicas em línguas, como o português [164], onde nós com maior grau costumam ligar-se com nós de menor grau, e de redes sociais, nas quais indivíduos são mais propensos a se conectar a nós com graus semelhantes [138]. Essa correlação pode ser mensurada através do cálculo do grau médio $k_{nn}(k)$ dos nós conectados a vértices com grau k, sendo a inclinação da curva k_{nn}, denotada por ν, usualmente utilizada como um indicador de como se dá esta dependência de graus.

4.2.2 Caminhos entre vértices e efeito *small-world*

Em um grafo, define-se caminho geodésico entre u e v como sendo o caminho com o menor número de arestas conectando os dois vértices, podendo, ou não, haver mais de um destes caminhos entre o mesmo par de vértices. A distância geodésica l_{uv}, por sua vez, indica a quantidade de arestas em um caminho geodésico do nó u ao nó v. Note que em um grafo dirigido não necessariamente l_{uv} é igual a l_{vu}. Quando u e v não pertencem ao mesmo componente conexo, convencionalmente faz-se $l_{uv} = \infty$. As distâncias mínimas entre nós possuem grande impacto sobre a forma com que fenômenos, como comunicação e transporte, evoluem em uma rede. Desta maneira, é possível estabelecer a distância geodésica média em uma rede, denotada por L, da seguinte forma:

$$L = \frac{1}{N(N-1)} \sum_{i,j \in V, i \neq j} l_{ij} \tag{4.1}$$

Esta medida é frequentemente chamada de *diâmetro* da rede, uma denominação ambígua, uma vez que alguns autores utilizam o termo diâmetro para indicar a maior distância entre dois vértices conectados na rede. É importante notar que esta média é calculada sobre todos os pares de vértices da rede, o que faz com que L divirja quando a rede não é conexa. Uma solução para este problema é utilizar para o cálculo da distância média apenas os vértices que se encontrem em um mesmo componente conexo [55, 179]. Uma alternativa menos utilizada é o uso de uma média harmônica das distâncias entre vértices [139], denominada de eficiência da rede [122]:

$$L^{-1} = \frac{1}{N(N-1)} \sum_{i,j \in V, i \neq j} l_{ij}^{-1} \tag{4.2}$$

Tipos diferentes de redes costumam ter distâncias médias características a eles. Em um reticulado de dimensão d contendo N vértices, por exemplo, tem-se que $L \propto N^{1/d}$, enquanto que em uma rede totalmente conectada $L = 1$. Para redes plenamente aleatórias, é possível realizar uma aproximação da distância média. Sendo k a quantidade de vizinhos mais próximos de um determinado nó, então cerca de k^l vértices encontram-se a uma distância de no máximo l arestas a partir do vértice em questão. Desta maneira, tem-se que $N \sim k^L$ e, portanto, $L \sim \log(N)/\log(k)$. Redes com graus distribuídos em lei de potência, por sua vez, possuem distâncias médias ainda menores, sendo no máximo proporcional a $\log(N)/\log(\log(N))$ [25], havendo indícios de que tal crescimento pode ser ainda mais lento [42]. A partir destas relações, é possível notar uma importante propriedade comum a muitas redes: a variação logarítmica da distância geodésica média com o tamanho da rede, relação denominada efeito de mundo pequeno (ou, do inglês, efeito *small-world*) [179]. Esta denominação surge a partir do trabalho de Stanley Milgram que, há cerca de quarenta anos, realizou um experimento no qual um conjunto de pessoas deveria enviar cartas a seus conhecidos e estes a seus contatos, recursivamente, até que estas cartas chegassem a um conjunto de indivíduos alvos [128]. Muitas das cartas se perderam, porém aquelas que chegaram a seu destino passaram, em média, pela mão de seis pessoas, uma distância pequena quando comparada à quantidade de pessoas existentes no planeta e à pequena quantidade de conexões sociais entre estas pessoas.

Em redes do mundo real, como redes biológicas, sociais e tecnológicas [134, 135, 178, 179], é comum observar a ocorrência do efeito *small-world*. Essa pequena distância média possui implicações importantes para os processos que ocorrem nessas redes, uma vez que as perturbações podem se propagar muito rapidamente. Isso pode ser visto na velocidade com que um rumor, ou uma epidemia, se espalha, assim como na necessidade de poucos intermediários para um pacote viajar pela Internet.

Apesar de todas exibirem o efeito *small-world*, as distâncias médias em redes do mundo real costumam ser maiores do que aquelas vistas em redes puramente aleatórias, segundo o modelo ER, ou em redes aleatórias com graus distribuídos em lei de potência. Tal situação indica a existência de alguns componentes não-aleatórios na topologia de redes complexas [4].

Ligada ao caminho geodésico, uma relevante característica relacionada a um nó isolado é a chamada centralidade de intermediação, ou carga (também é utilizado o termo em inglês *beetweeness*) [111]:

$$b_i = \sum_{j,k \in V, j \neq k} \frac{n_{jk}(i)}{n_{jk}}, \tag{4.3}$$

onde n_{jk} é a quantidade de caminhos geodésicos conectando o nó j ao nó k e $n_{jk}(i)$ é a quantidade

destes caminhos que passa pelo nó i. Esta métrica indica a proporção de caminhos geodésicos que passam pelo nó em questão, sendo utilizada, juntamente ao grau e à distância média do nó ao restante dos nós, para quantificar sua importância para a conectividade da rede.

4.2.3 Coeficiente de aglomeração e comunidades

Em redes sociais, é comum verificar-se a formação de círculos de amigos ou conhecidos, nos quais todos os membros têm relações diretas com os outros membros, havendo alta probabilidade de que, se um indivíduo A conhece B e C, então B e C se conheçam. Esta tendência inerente ao agrupamento é chamada de aglomeração ou transitividade e pode ser quantificada de diferentes formas. Uma versão amplamente utilizada, especialmente em análise de dados de redes reais, devido a sua facilidade para ser calculada computacionalmente, é a proposta de Watts e Strogatz [179], que será utilizada durante este trabalho. Nesta abordagem, cada nó possui seu coeficiente de aglomeração próprio, o qual indica a probabilidade de que dois de seus vizinhos, selecionados aleatoriamente, sejam vizinhos entre si. Desta forma, o coeficiente de aglomeração para o nó i é dado por:

$$C_i = \frac{2E_i}{k_i(k_i - 1)}, \tag{4.4}$$

sendo E_i a quantidade de conexões diretas existentes entre os vizinhos do nó i. Assim, para calcular o coeficiente de aglomeração da rede como um todo, é avaliada a média aritmética de C_i para todos os nós:

$$C = \frac{1}{|V|} \sum_{i \in V} C_i. \tag{4.5}$$

Esta medida, no entanto, acaba por dar maior peso para nós que possuem menor grau, não refletindo de forma precisa para toda a rede a probabilidade de que dois vizinhos de um determinado nó tenham uma conexão direta entre si. Logo, é importante mencionar uma segunda maneira de calcular a aglomeração da rede, bastante difundida na sociologia, lá denominada de "fração de triplas transitivas" [24, 139]. Segundo esta proposta, C é calculado da seguinte maneira:

$$C = \frac{6|\{\{i,j,k\}|(i,j),(j,k),(k,i) \in E\}|}{|(i,j,k)|(i,j),(j,k) \in E\}|}, \tag{4.6}$$

onde o numerador quantifica o número de cliques (subgrafos totalmente conectados) de tamanho três (chamadas de triângulos) na rede e o denominador indica a quantidade de caminhos de comprimento dois. O fator multiplicativo 6 é devido ao fato de um triângulo contribuir com seis caminhos de comprimento dois. O coeficiente C, neste caso, indica diretamente a probabilidade de que dois vértices selecionados aleatoriamente na rede, e que possuam ao menos um vizinho em comum, também sejam

vizinhos imediatos entre si.

Em um grafo totalmente conectado, tem-se que $C = 1$, enquanto para grafos estruturados em árvore, $C = 0$. Em uma rede aleatória, segundo o modelo de Erdös e Rényi [66], o coeficiente de aglomeração da rede é igual à probabilidade de existir uma determinada aresta e, desta forma, $C = O(N^{-1})$ para valores grandes de N. Em redes reais, no entanto, tem-se que $C = O(1)$ para $N \to \infty$, de forma que seu coeficiente de aglomeração é muito maior do que o valor previsto para uma rede aleatória com o mesmo número de arestas e de vértices, assemelhando-se, nesse aspecto, mais a um reticulado.

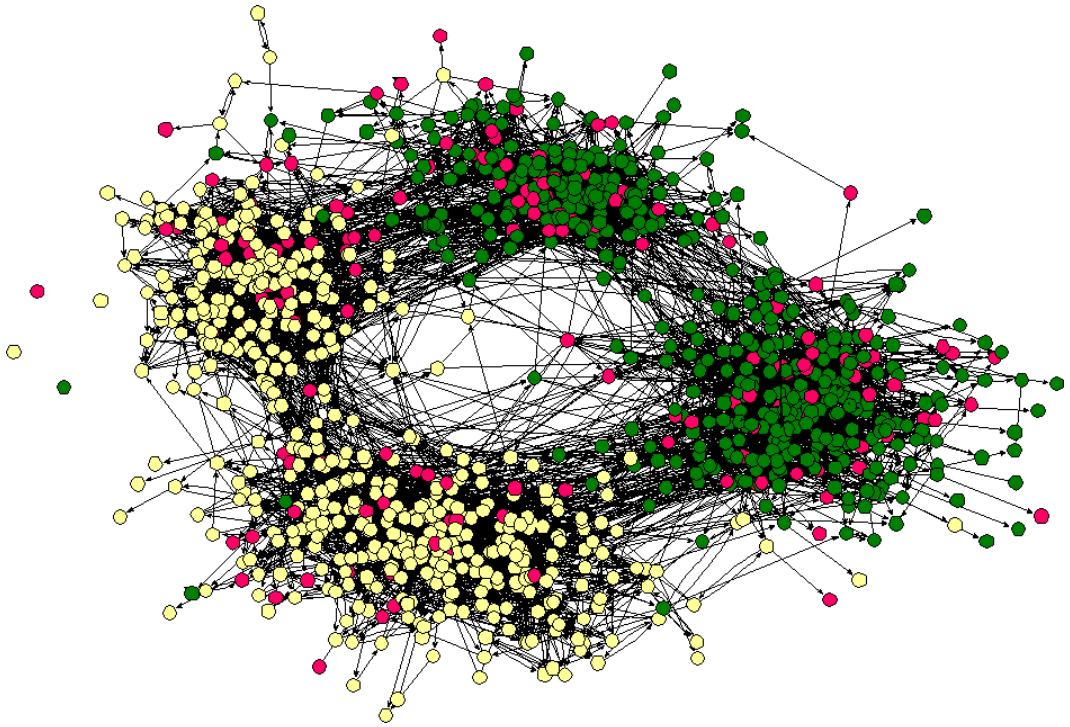


Fig. 4.1: Rede de amizade em uma escola americana: divisão em quatro comunidades. A diagonal que passa pela porção superior esquerda da imagem claramente separando os indivíduos por raça (os vértices são coloridos de acordo com a etnia do indivíduo). A outra diagonal separa alunos do ensino fundamental de alunos do ensino médio.

Uma outra característica usualmente presente em redes sociais é a formação de comunidades, ou seja, conjuntos de vértices dentro dos quais há uma grande quantidade de conexões, porém com relativamente poucas arestas para fora do grupo. De fato, é comum observar que as pessoas possuem maior tendência a se conectar a determinados grupos – sociais, etários ou regionais, por exemplo – com os quais se identificam, o que acaba por dar origem a tais comunidades. Essa situação pode ser vista claramente na figura 4.1, a qual exibe uma rede de amizades entre adolescentes em uma

escola americana [131][2]. De maneira similar, espera-se a formação de comunidades em outras redes complexas, como a divisão de artigos em redes de citações de acordo com áreas de conhecimento, o agrupamento de páginas da *web*, na rede de *hyperlinks*, de acordo com seus assuntos e a organização em módulos de redes neurais ou metabólicas de acordo com suas funcionalidades, suposições essas que parecem ser comprovadas por resultados experimentais [75, 81, 86, 180].

4.2.4 Outras características

Tendo em vista que o estudo de redes complexas é um campo muito recente, há ainda discussões sobre quais as principais características que mereçam ser estudadas em tais redes. Além das apresentadas acima, há diversas outras propriedades presentes em boa parte das redes existentes no mundo, sendo uma destas a formação de motivos (ou, no inglês, *motifs*) [162]. Um motivo é um padrão recorrente de interconexões que ocorre tanto em grafos dirigidos como em não-dirigidos, o qual aparece em uma determinada rede em quantidade superior ao esperado para um grafo aleatório com a igual quantidade de vértices e arestas e com a mesma distribuição de graus [45]. Dada essa alta frequência, imagina-se que tais sub-grafos possuam funções específicas nas redes onde estão presentes [9], o que pode ser visto em redes de transcrição genética, onde laços *feed-forward* (ver figura 4.2), por exemplo, agiriam como filtros, geradores de pulso e aceleradores de resposta [188]. Os motivos estariam, então, interterconectados, formando módulos responsáveis por prover funcionalidades específicas [44]. Sendo assim, motivos atuariam como elementos de circuito simples, o que facilitaria o estudo de dinâmica em redes complexas.

Algumas outras propriedades podem ser vistas como resultados das características já apresentadas. Principalmente devido à forma da distribuição de graus em uma rede livre-de-escala, a qual garante que poucos nós tenham graus mais altos e que a maior parte dos elementos da rede situem-se em sua periferia, essas redes possuem alta robustez estrutural contra a ocorrência de eventos de falha, com uma baixa degeneração de sua conectividade (tamanho do maior componente conexo e caminho geodésico médio) caso sejam retirados nós aleatórios da rede, sendo necessário remover uma proporção significativa dos nós para conseguir comprometer significativamente o funcionamento de uma rede suficientemente grande [41]. Por outro lado, uma distribuição em lei de potência também torna as redes vulneráveis a ataques intencionais aos nós com graus mais altos, sendo necessário retirar apenas 2% dos nós para uma deterioração severa da conectividade em algumas redes [32]. É importante notar que essas análises são meramente estruturais, não levando em conta efeitos dinâmicos na rede. Quando considerado o papel da dinâmica, por um lado resultados experimentais demonstram que redes homogêneas são mais resistentes a falhas em cascata que redes heterogêneas, como redes

[2] A figura em questão, porém, não se encontra no trabalho referido, podendo ser encontrada em `http://www.visualcomplexity.com/vc/project.cfm?id=19`.

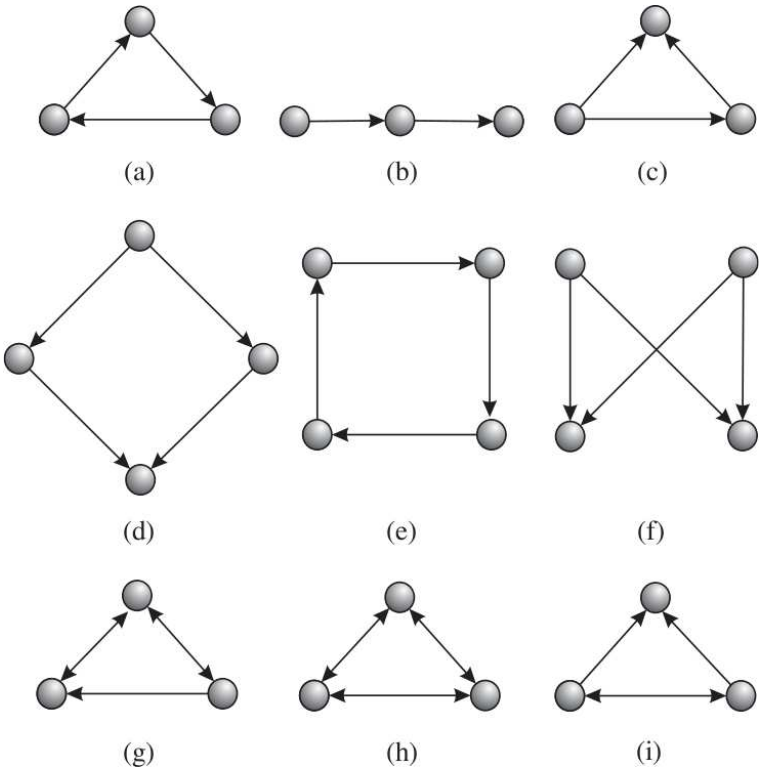


Fig. 4.2: Alguns motivos típicos e suas denominações: (a) laço de realimentação com três vértices, (b) cadeia tripla, (c) laço *feed-forward*, (d) bi-paralelo, (e) laço de realimentação com quatro vértices, (f) *bi-fan*, (g) realimentação com duas conexões mútuas, (h) tríade totalmente conectada, (i) conexão mútua com um nó superior. Imagem retirada de [45].

livres-de-escala [132], mas por outro lado há dados que indicam que redes livres-de-escala apresentam níveis mais baixos de congestionamento do que estruturas com outras distribuições, como grafos aleatórios [172].

Apesar de não ter sido abordada no trabalho de Milgram [128], uma propriedade que pode ser observada na rede social do experimento de envio de cartas (e em muitas outras redes do mundo real) é a existência de estruturas que possibilitam encontrar caminhos curtos baseando-se apenas em informações locais aos nós e sem qualquer método de busca mais rebuscado[3]. Esta propriedade poderia ser utilizada, por exemplo, para a otimização na busca de informação na *web*, na navegação em ruas de uma cidade e no roteamento de pacotes na Internet.

[3]No experimento de Milgram [128], os envolvidos possuíam apenas informações como primeiros nomes, cidades em que residem e emprego dos destinatários, sem qualquer conhecimento mais profundo sobre a topologia da rede social, com cada pessoa repassando a carta a outro indivíduo que ela imaginava estar mais próximo do alvo. Mesmo assim, os participantes foram capazes de encontrar caminhos eficientes conectando as pessoas de origem e de destino.

4.3 Modelos de crescimento

Várias propostas foram apresentadas até o momento para explicar o crescimento de redes complexas. Porém, as duas mais presentes na literatura são os modelos de redes *small-world*, proposto por Watts e Strogatz [179], e o modelo livre-de-escala, proposto por Barabási e Albert [13]. Nesta seção será realizada uma breve descrição destes mecanismos.

4.3.1 Modelo *small-world*

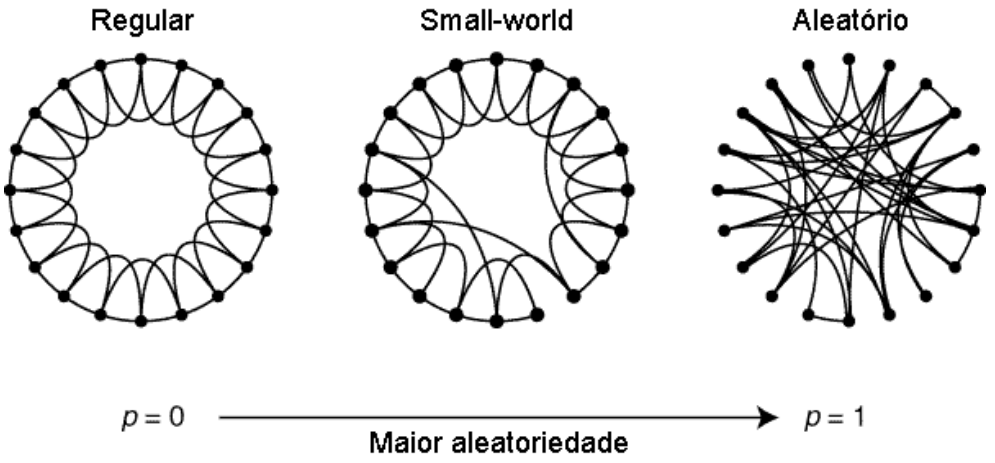


Fig. 4.3: Procedimento de reconexão aleatória no modelo de Watts-Strogatz, o qual realiza uma transição entre um anel regular e uma rede aleatória conforme a variação da probabilidade p de uma aresta sofrer reconexão. Imagem retirada de [179].

Conforme observado na seção 4.2, redes no mundo real costumam exibir duas características de certa forma antagônicas e não observadas em modelos mais simples de redes, como redes aleatórias ou reticulados, a saber: a ocorrência do efeito *small-world* e um alto coeficiente de aglomeração. Tendo isto em mente, Watts e Strogatz [179] propuseram o modelo de redes *small-world*, para a transição de redes entre grafos regulares e grafos aleatórios, dado pelo seguinte algoritmo:

- Início: a rede é iniciada como um anel com N vértices, cada qual conectado a seus K vizinhos mais próximos.
- Aleatoriedade: cada aresta do grafo original é reconectada com probabilidade p, sem a possibilidade da criação de laços ou multi-arestas. A reconexão de uma aresta é feita ao desligá-la de um de seus extremos, contectando-a a um vértice selecionado aleatoriamente da rede.

Pode-se observar que, quando $p = 0$, a rede final mantém-se como um anel ordenado, enquanto que quando, $p = 1$, a rede torna-se plenamente aleatória. O processo em questão cria $pNK/2$ arestas

de longo alcance, unindo nós anteriormente distantes. Desta maneira, quando p é pequeno, a rede gerada é capaz de manter o alto coeficiente de aglomeração da rede original[4] com a criação de alguns atalhos reduzindo rapidamente a distância geodésica média entre os nós, originalmente dada por $L = N/(2K)$, a uma relação logarítmica.

4.3.2 Modelo livre-de-escala

Modelos anteriores, como o modelo de grafo aleatório de Erdös e Rényi [66] e o modelo de redes *small-world* de Wattz e Strogatz [179], propunham formas de geração de grafos em que a quantidade de vértices no sistema era definida *a priori*, sendo permitida a criação apenas de novas ligações no decorrer do tempo. Apesar de suprir algumas das características comuns às redes complexas, esses modelos mostraram-se insuficientes para explicar a origem de tais sistemas. Em ambos os casos, por exemplo, a distribuição de graus segue uma distribuição de Poisson [15], com um pico em $< k >$ e caindo exponencialmente para graus muito grandes, ao contrário da lei de potência observada em várias redes do mundo real.

Através da análise das redes reais, é possível observar que, em geral, elas não são fechadas a novos nós. Em redes de interação social é comum o aparecimento de novos membros – através do nascimento ou da mudança de pessoas, por exemplo –, assim como de tempos em tempos surgem novas espécies em redes biológicas, alterando sua estrutura de relações. Também é possível verificar que, em algumas redes, para cada nova relação criada há uma probabilidade maior de que o nó que a receba seja um nó altamente conectado. Tal característica pode ser vista na rede de citações científicas, onde artigos já bastante citados apresentam maior probabilidade de receber novas citações frente a artigos desconhecidos.

Deste modo, como uma alternativa aos modelos anteriores, Barabási e Albert [13] introduziram o modelo livre-de-escala (também chamado pela sigla SF, do inglês *scale-free*), que mostra-se mais consistente com redes reais. O modelo SF pode ser definido em duas etapas:

- Crescimento: a rede deve ser iniciada com uma quantidade m_0 de nós, sendo adicionado um novo nó a cada passo. Cada nó adicionado possui, a princípio, uma quantidade m ($m <= m_0$) de arestas.

- Conexão preferencial (ou *preferential attachment*): cada aresta adicionada deve ligar o novo nó a um nó i já existente na rede, escolhido com probabilidade proporcional à sua quantidade de

[4]Em um grafo em anel, como o utilizado no modelo de Watts e Strogatz, o coeficiente de aglomeração é dado por:

$$C = \frac{3(K-2)}{4(K-1)}.$$

ligações, conforme a equação

$$p(k) = \frac{k_i}{\sum_j k_j}, \tag{4.7}$$

onde k_i é o grau do nó i.

A combinação destas duas etapas é necessária para a geração de redes com distribuição de graus em lei de potência, distribuição esta que torna-se independente de tempo e tamanho da rede após a execução de alguns passos da etapa de crescimento. Observa-se, também, que este modelo produz redes com distâncias médias menores e coeficientes de aglomeração significantemente superiores aos vistos em grafos aleatórios [4]. Porém, ao contrário do que ocorre tanto em redes reais como em redes *small-world*, onde tal aglomeração mostra-se independente do tamanho da rede, em redes SF C diminui conforme aumenta a quantidade de nós N, com $C \sim N^{-0,75}$.

Variantes propostas

Apesar de ser um modelo amplamente utilizado, o modelo livre-de-escala puro não é capaz de prover algumas das características de redes complexas. Como exemplo, pode-se tomar o expoente γ da distribuição de graus: na natureza, observam-se normalmente $1 \leq \gamma \leq 3$. Porém, redes geradas pelo modelo de Barabási e Albert [13] possuem expoente fixo e igual a 3. Assim, abundam na literatura propostas de alterações do modelo livre-de-escala. Alguns desses visam tornar os resultados previstos mais compatíveis com a realidade, ou tornar a forma de evolução do modelo mais próxima das teorias existentes, ou simplesmente apenas estudar o resultado da flexibilização do modelo. A seguir, serão apresentadas algumas destas propostas.

- **Adição e remoção de arestas**: no modelo original, a cada passo, uma quantidade m de novas arestas, todas ligadas ao nó adicionado no passo atual, é adicionada definitivamente à rede. Apesar de satisfatório em alguns casos, como redes de citação entre artigos, tal modelo não é compatível com muitos outros sistemas reais. Em redes de relações sociais, é comum o aparecimento e desaparecimento de relações durante o tempo, mesmo entre membros antigos da rede. Tendo tal situação em vista, Dorogovtsev e Mendes [53] propuseram a inclusão de um novo mecanismo ao modelo de Barabási e Albert. Além da adição de vértices e arestas presentes no modelo original, a cada passo da evolução cada par de nós poderia ter uma ligação adicionada ou removida da rede, de forma proporcional ao produto dos graus de ambos os nós. Callaway *et al.* [31], por sua vez, propuseram outra alteração à forma de adição de arestas através da flexibilização da origem e destino da nova aresta, permitindo que essa tenha sua origem e destino em qualquer nó da rede. Porém, tal alteração, como proposta, não apresenta

preferential attachment, resultando em uma rede exponencial, e não em lei de potência. No entanto, Dorogovtsev *et al.* [56] propuseram uma generalização desse modelo, suportando o *preferential attachment*, de tal forma que a característica de lei da potência esteja novamente presente.

- ***Fitness***: no modelo livre-de-escala, o grau de um vértice tende a ser proporcional à sua idade. Tal relação, porém, não está presente em redes como a *web*. Adamic e Huberman [3] sugerem que tal comportamento origine-se em uma qualidade interna do nó, denominada *fitness*. A forma com que esse valor influenciaria na atratividade pode relacionar-se, basicamente, de duas formas com o *preferential attachment*, evidenciadas a seguir:

  - Multiplicativa: Bianconi e Barabási [20] propuseram uma alteração ao modelo livre-de-escala, segundo a qual a probabilidade de ligação de um nó torna-se proporcional ao produto $\eta_i k_i$, sendo η_i o *fitness* do nó. Dependendo da distribuição de valores de η, a rede gerada pode chegar a um regime livre-de-escala ou um dos nós pode apresentar um certo monopólio, dominando uma fração finita das ligações.
  - Aditiva: Ergün e Rodgers [67] estudaram uma variação do modelo anterior, onde a influência do *fitness* não é mais multiplicativa, porém aditiva. Assim como no modelo anterior, eles encontraram que, para determinadas distribuições de *fitness*, tal modelo é capaz de gerar redes em lei de potência.

  O uso de *fitness* aditivo é particularmente relevante em modelos que permitam que novos vértices não sejam apontados assim que criados (como o caso de adição de arestas não relacionadas, citada acima, e o de grafos direcionados, a ser explicado a seguir), garantindo que todos os nós possuam certo grau de atratividade inicial.

- **Envelhecimento de vértices**: em certas redes, a atratividade de um nó não depende apenas de seu grau, ou de uma característica fixa intrínseca de sua qualidade, mas depende também de sua idade. Exemplos claros disso podem ser vistos em redes de citação científicas, onde torna-se rara a citação de artigos muito antigos, e em redes de atuação conjunta de atores, onde, com a idade, seus membros tendem a se aposentar. O primeiro caso é contemplado por um mecanismo de envelhecimento gradual proposto por Dorogovtsev e Mendes [52], segundo o qual a atratividade original de um nó seria multiplicada por um fator $t^{-\alpha}$, onde t é a idade do nó e α é uma constante. É mostrado, também, que a rede adquire características livre-de-escala apenas quando $\alpha < 1$. Outras propostas de funções de envelhecimento gradual também foram realizadas, podendo ser citada como exemplo o estudo do uso de uma função exponencial, $e^{\beta t}$, com β sendo uma constante, realizado por Zhu *et al.* [189]. O caso posterior, de envelhecimento

por inatividade, por sua vez, é atendido pela proposta de Amaral *et al.* [6], onde a cada passo o vértice i pode tornar-se inativo (não podendo mais receber ligações) com probabilidade p_i. A rede resultante possui uma distribuição em lei de potência com um limitante exponencial, de forma que variando as taxas de envelhecimento é possível modificar tal limitante, até que para valores muito altos de p_i a característica livre-de-escala desaparece. Efeito semelhante pode ser obtido ao estabelecer limitações de custos e capacidades dos nós em receber conexões.

- **Aceleração do crescimento**: muitas redes não apresentam taxas de crescimento constantes no tempo, como é o caso da rede formada pelos *hyperlinks* na *web*. Para cobrir este caso é proposto em [54] uma generalização, estendendo o modelo de Barabási e Albert para métodos não-lineares de crescimento, onde a cada passo são adicionadas $m \propto t^{\alpha}$ arestas à rede. A rede gerada continua possuindo distribuição de graus em lei de potência, porém de forma mais flexível, com o expoente $\gamma = 1 + 1/(1 + \alpha)$.

- **Grafo direcionado**: o modelo livre-de-escala original foi estruturado apenas para grafos com arestas não-direcionadas. Uma extensão direta do modelo, considerando o *preferential attachment* relacionado apenas ao grau de entrada dos nós e adicionando somente arestas com origem no novo nó, levaria ao problema de como um nó receberia sua primeira referência. Algumas alternativas foram propostas como solução: Dorogovtsev *et al.* [56] propõem a utilização de uma atratividade adicional, um valor fixo a ser somado à atratividade de cada nó, garantindo que todos os nós tenham a chance de receber sua primeira referência. Outra possibilidade [55] é adicionar, juntamente com cada novo nó, n arestas apontando para ele e m novas arestas apontando para nós antigos da rede. Por fim, uma terceira proposta seria considerar o *preferential attachment* sobre a soma dos graus de entrada e saída de cada nó. Porém, essa proposta falha por falta de realismo [139].

Como pode ser observado, na maioria dos casos apresentados, as características de redes complexas são preservadas. As alterações influenciam, basicamente, a forma da lei de potência associada à rede, alterando seu fator de multiplicação, seu expoente ou seu limitante exponencial.

Outros modelos

Um fato importante a se notar é que o comportamento de *preferential attachment* pode estar presente mesmo sem ser implementado explicitamente. Um modelo simples [57] que se encaixa nessa situação pode ser descrito por três regras básicas:

- Início: a rede é iniciada com três nós, cada um desses conectado aos outros dois nós.

- Crescimento: a cada passo um novo nó é adicionado.
- Arestas: é selecionada randomicamente uma aresta a cada passo, sendo ligado o novo nó às duas extremidades da aresta escolhida.

É possível verificar que a conexão preferencial surge naturalmente com esse modelo: uma vez que a seleção de destinos de novas arestas é feita com base nas próprias arestas já existentes, a probabilidade de um vértice ser selecionado para receber uma ligação é proporcional à quantidade de arestas que esse possui.

Modelos de conexão de vizinhos próximos

Algumas modificações surgiram com vistas a fornecer um maior coeficiente de aglomeração ao modelo livre-de-escala. Holme e Kim [87] propuseram o modelo HK, que consiste em conectar o nó adicionado ou a um nó selecionado através da regra de *preferential attachment* ou a um nó que seja vizinho de um vizinho seu, sendo, assim, capaz de produzir redes com distribuições de graus em lei de potência e que possuam um coeficiente de aglomeração de até $0,5$. O modelo DEB de Davidsen *et al.* [46], por sua vez, parte de uma rede inicial, na qual, a cada passo, um nó deve "apresentar" dois vizinhos seus, criando uma conexão entre eles, e, com probabilidade p, um nó é removido da rede e substituído por um novo nó, o qual recebe um vizinho escolhido aleatoriamente. Tal dinâmica é capaz de gerar redes com efeito *small-world*, alta aglomeração e, dependendo do parâmetro p, distribuição de graus em lei de potência ou exponencial.

Assim como o exemplo anterior, o modelo DEB não apresenta o *preferential attachment* implementado de forma explícita. A produção de redes livres-de-escala ocorre, pois quanto mais arestas um nó possui, maior sua quantidade de vizinhos de segundo grau. Sendo assim, a probabilidade de um nó receber uma nova ligação é proporcional a seu grau.

Modelos de duplicação e divergência

Para algumas redes, o modelo livre-de-escala fornece resultados válidos, porém o método de crescimento não condiz com a realidade. Um desses casos é o das redes proteômicas. A origem de novos genes é vista, por exemplo, como resultado de um processo de *crossover* desigual[5] e de mudança na interferência entre genes. Desta forma, o novo gene seria criado não por adição, como no modelo anterior, mas por duplicação mantendo parte das ligações do gene original.

O modelo proposto por Solé *et al.* [165] contempla essa forma de crescimento via divisão de nós, compondo-se das seguintes etapas:

[5]No *crossover* desigual, os dois cromossomos a serem cruzados não se tocam no ponto esperado, fazendo com que um dos cromossomos possua mais genes que o esperado, enquanto que o outro cromossomo torna-se menor.

- Duplicação: a cada passo um nó da rede é escolhido aleatoriamente e duplicado, juntamente de suas conexões.
- Divergência: as arestas do novo nó são removidas com uma probabilidade δ.
- Mutação: αN novas arestas são criadas do novo nó para nós que não sejam vizinhos do nó original.

O *preferential attachment* neste modelo é decorrente do fato de na duplicação de um nó quanto mais arestas um nó que não está sendo copiado possui, maior a chance de que ele receba uma nova conexão. Apesar do modelo em questão produzir redes com distribuição de graus em lei de potência, as correlações de graus e o coeficiente de aglomeração não exibem valores realistas [106].

O modelo "*good gets richer*"

Este modelo difere um pouco mais dos modelos apresentados anteriormente por não envolver necessariamente redes abertas que cresçam com o tempo, e tampouco valer-se de conceitos de *preferential attachment*. Nessa proposta [30], a rede possui uma quantidade de vértices definida *a priori*, embora após alguns passos seja possível acrescentar novos vértices à rede. Cada um desses vértices i possui um *fitness* intrínseco η_i, de forma que, ao ser criada uma nova aresta, os nós relacionados sejam selecionados com probabilidade proporcional a esse *fitness*, segundo a seguinte equação:

$$p(k) = \frac{\eta_i}{\sum_j \eta_j} \tag{4.8}$$

Nem todas as distribuições de *fitness* levarão a rede a uma configuração livre-de-escala. Porém, há alguns formatos de distribuição capazes de fazer a rede chegar a essa condição. O caso mais trivial é o de uma distribuição em lei de potência[6]. Caldarelli *et al.* [30] demonstraram que as características de redes geradas dessa forma são as mesmas das redes geradas com o modelo livre-de-escala.

Otimização topológica

Apesar de grande parte dos estudos sobre a origem de redes complexas focar em procedimentos de crescimento, recentemente foi demonstrado ser possível a obtenção de topologias livres-de-escala como frutos de processos de otimização [72, 174]. Tal resultado abre uma nova possibilidade para uma presença tão difundida de topologias complexas, podendo esta situação ser fruto também de características benéficas que estas redes trazem aos processos que nelas se apoiam. Deste modo,

[6]Distribuições desse tipo são muito comuns também quando relacionadas a critérios de ranqueamento, conforme a lei de Zipf [123].

torna-se cada vez mais importante o estudo das propriedades dinâmicas destas redes, analisando se de fato sua aplicação traz resultados positivos ou se seus mecanismos de crescimentos seriam os únicos responsáveis por suas características.

4.4 Conclusão do capítulo

Análises recentes indicam que muitas das redes observadas no mundo compartilham certas características chaves, como rapidez na propagação de dados, formação de comunidades e robustez contra eventos de falha em nós localizados. Existe na literatura uma série de propostas de modelos de crescimento que sejam capazes de prover as propriedades citadas, verificando-se que, para o surgimento de uma topologia complexa, um fator importante é a presença de atratividade por parte dos nós para receber novas conexões proporcional à quantidade de referências que o nó em questão já possui, o que é denominado *preferential attachment*. Este fator mostra-se compatível com observações empíricas, como na *web*, onde uma página bem conhecida possui maior probabilidade de receber novos *hyperlinks*. Dentre as propostas de crescimento vistas neste capítulo, aquelas baseadas na conexão de vizinhos próximos são de particular interesse por seguir mais firmemente a idéia de auto-organização, uma vez que utilizam apenas informações locais para produzir redes complexas, não necessitando de informações como a distribuição de graus do sistema em um determinado instante. Porém, a despeito da quantidade de alternativas presentes na literatura para evolução topológica, este é um campo relativamente recente, ainda carecendo de estudos mais profundos que possam contemplar de forma mais adequada, por exemplo, o surgimento de comunidades e motivos. Ressalta-se que, devido a sua recorrência, os motivos aparentam possuir uma grande importância na dinâmica de redes complexas, de forma que o estudo de suas funcionalidades torna-se fundamental para uma compreensão mais completa destas.

Por fim, note-se que algumas das propriedades de topologias complexas são extremamente desejáveis na engenharia e que os métodos de crescimento apresentados têm sido ostensivamente avaliados analítica e empiricamente, possuindo um comportamento relativamente bem definido. Assim, ao menos à primeira vista, a aplicação total ou parcial dos modelos de geração de redes apresentados no capítulo atual ao projeto de sistemas tem potenciais efeitos positivos, constituindo uma frente de pesquisa que ainda necessita ser melhor explorada.

Capítulo 5

Aplicação de Redes Complexas na Definição de Vizinhança entre Partículas

Conforme mostrou o capítulo 4, é comum em processos auto-organizados a emergência de topologias complexas de interação. E, uma vez que a inteligência coletiva é um exemplo clássico de auto-organização, pode-se dizer que, na natureza, este tipo de mecanismo é regulado por redes complexas, sendo razoável supor que tais arranjos potencialmente possuem uma influência positiva sobre a eficiência dos procedimentos em questão. Por sua vez, na seção 3.2 observa-se que muitas foram as idéias para melhorar o desempenho da otimização por enxame de partículas. Algumas dessas propostas exploraram modificações na forma como as partículas combinam as informações individuais e aquelas obtidas de seus vizinhos, enquanto que outras propuseram alternativas para a organização topológica da população. Porém, poucos foram os estudos sobre o impacto do uso de redes complexas na definição de vizinhança no PSO, sendo que as raras investigações existentes sobre tal assunto concentram-se na análise de algumas características específicas. Tal situação pode ser vista no trabalho de Kennedy [97], onde é analisado apenas o impacto do efeito *small-world* sobre o desempenho da otimização. Deste modo, esta dissertação propõe a Otimização por Enxame de Partículas com Vizinhança Complexa (ou CNPSO, do inglês *Complex Neighborhood Particle Swarm Optimizer*), uma variante do PSO na qual a vizinhança é determinada através de uma rede complexa dinâmica, a qual será descrita ao longo deste capítulo.

5.1 Os algoritmos CNPSO

Um problema comumente observado no PSO é sua deficiência em trabalhar com problemas multimodais, especialmente quando utilizado com uma topologia global, situação na qual cada partícula possui apenas um atrator global e um individual. Desta forma, todo o enxame é rapidamente atraído

para o melhor resultado já alcançado, facilitando que o mecanismo fique preso em ótimos locais. O CNPSO, por outro lado, busca administrar de forma adequada a influência durante a execução do algoritmo, através da diminuição da vizinhança de cada partícula. A redução de influência poderia ser realizada, por exemplo, através do simples uso de um parâmetro social mais baixo. Porém, ao usar uma topologia complexa, pretende-se que as partículas se auto-organizem em comunidades, de forma que cada comunidade funcione como uma espécie de sub-enxame capaz de refinar bons resultados obtidos por alguns de seus membros, o que aumentaria a habilidade do sistema em trabalhar com múltiplas soluções. Devido ao efeito *small-world*, a topologia escolhida também ajuda a evitar uma desnecessária dispersão das partículas, uma vez que é assegurado um rápido fluxo de informações pela rede. Esta organização tenderia, também, a alocar nós com menos conexões à busca local em torno de regiões promissoras já visitadas, enquanto partículas com graus maiores seriam designadas à exploração do espaço, uma vez que, por receberem mais informações, elas possuem maior probabilidade de terem diferentes atratores em iterações subsequentes.

No CNPSO, para que uma solução consiga dominar todos os indivíduos é necessário que, antes, ela supere suas rivais. Para poder se espalhar efetivamente, também é importante que a solução consiga atingir um *hub*, onde ela terá acesso a uma grande variedade de caminhos na rede. No entanto, a competição para conquistar os *hubs* é acirrada, levando um certo tempo até que uma solução consiga transpor todas as outras em seu caminho. Este processo mais longo, porém, acaba por ser um importante filtro, o qual evita que soluções menos desejáveis dominem toda a rede.

No PSO, muitas vezes as partículas convergem prematuramente, ficando estagnadas em regiões de baixa qualidade, o que prejudica o fluxo de informações e desperdiça poder de processamento. Assim, dado que a troca de informações em redes complexas é altamente dependente de seus *hubs* [68], a primeira versão do CNPSO, aqui chamada de CNPSO-1, introduzida em [78], buscou contornar situações em que estes nós ficavam por diversas iterações em regiões pouco promissoras. Para atingir tal objetivo, um mecanismo de evolução topológica foi adicionado ao sistema, forçando o enxame a uma melhor seleção de quais os nós mais influentes. Tal processo foi cuidadosamente ajustado para manter a topologia com características de redes complexas durante todo o processo de otimização. Deste modo, de tempos em tempos, uma conexão entre partículas é selecionada aleatoriamente e desligada de sua extremidade com pior desempenho sendo, então, conectada a um outro nó da rede, também escolhido aleatoriamente, ambas as amostragens segundo uma distribuição uniforme. Este mecanismo faz com que nós localizados em regiões de baixa qualidade percam gradualmente influência na rede, evitando também que um elemento detenha parte significativa das arestas redirecionadas, situação que tenderia a ocorrer se o nó que recebe novas conexões fosse escolhido diretamente através de comparação de seu desempenho frente ao restante do sistema. Uma característica importante da dinâmica proposta é que o número de conexões no enxame permanece constante, uma vez que, para

cada aresta removida, outra é acrescentada.

Um segunda versão do algoritmo, chamada CNPSO+, buscou flexibilizar a quantidade de arestas na rede e incorporar mecanismos mais gerais para evitar a estagnação das partículas, de forma a não focar apenas nos *hubs*. Em sociedades humanas, observa-se que indivíduos que não conseguem resolver um determinado problema tendem a pedir a ajuda de mais e mais pessoas, enquanto que aqueles cujas estratégias obtêm contínuos sucessos passam a tornar-se mais confiantes em suas experiências e instintos pessoais. Mimetizando este processo, no CNPSO+ uma partícula pode adaptar sua vizinhança de duas formas distintas: caso encontre-se estagnada, não tendo conseguido melhorar os resultados obtidos durante a iteração atual, ela irá buscar novas informações na rede, incluindo em sua vizinhança imediata um nó selecionado aleatoriamente com probabilidade uniforme dentre seus vizinhos de segundo grau. Se não houver nenhum nó a essa distância, ou seja, a partícula em questão possui conexões diretas com todos os nós pertencentes ao seu componente conexo de saída, o nó a ser incluído na vizinhança da partícula é escolhido dentre todos os existentes na rede. Tal funcionamento inspira-se no modelo DEB de crescimento de redes complexas, baseado na conexão de vizinhos próximos (ver subseção 4.3.2). Porém, se a partícula tiver obtido alguma melhoria, ela deverá intensificar seu esforço de busca local, desvencilhando-se de seu pior informante. Deve ser ressaltado que, nesta abordagem, a topologia das partículas é dirigida, o que permite que cada adição ou remoção de arestas não afete sua conexão recíproca, o que ocorreria caso a vizinhança fosse não-dirigida.

Como pode ser visto no algoritmo 2 (CNPSO para minimização), ao início de sua execução, o CNPSO gera uma topologia inicial Z de partículas e inicializa os estados individuais. É importante ressaltar que a forma de geração da rede inicial não é pré-definida, podendo ser considerada um parâmetro. No algoritmo apresentado, $vizinhos(i)$ é o conjunto de vizinhos da partícula i, p_g é a melhor posição já visitada pelos vizinhos da partícula em questão e $estagnada_i$ indica se a partícula não conseguiu melhorar seu resultado na última iteração. A cada iteração, as partículas avaliam seus vizinhos, selecionando aqueles com melhor histórico como suas influências sociais, utilizadas para calcular a direção a ser seguida. As posições das partículas são, então, recalculadas usando as mesmas equações do PSO canônico. Por fim, a cada número $periodo$ de iterações, finalizadas as atualizações dos estados individuais, a rede sofre modificações topológicas, segundo procedimentos explicados anteriormente e cujos pseudo-códigos estão exibidos abaixo.

Pode-se notar que, como o grau médio de cada nó costuma ser um parâmetro dos métodos de crescimento de redes complexas e o CNPSO-1 mantém a quantidade de arestas constante no tempo, a escolha por parte de uma partícula de seu vizinho com melhor histórico de busca pode ser feita em tempo $O(1)$, enquanto que a adaptação topológica é realizada em tempo linear no tamanho do enxame, tendo, assim, baixo impacto na velocidade de execução do algoritmo.

Entrada: ω, c_1, c_2, N
Gerar a vizinhança Z das partículas através de algum método de crescimento de redes complexas;
Inicializar a posição x_i, a velocidade v_i e a melhor posição pessoal p_i das N partículas;
enquanto *o critério de parada não é satisfeito* **faça**
  para cada $j \in V(Z)$ **faça**
    $g \leftarrow \arg\min_{i \in vizinhos(j)} f(p_i)$;
    $v_j \leftarrow \omega v_j + c_1 r_1 (p_j - x_j) + c_2 r_2 (p_g - x_j)$;
    $x_j \leftarrow x_j + v_j$;
    $estagnada_j \leftarrow verdadeiro$;
    se x_j *está dentro do espaço de busca e* $f(x_j) < f(p_j)$ **então**
      $estagnada_j \leftarrow falso$;
      $p_j \leftarrow x_j$;
    fim
  fim
  se $iteracao \bmod periodo = 0$ **então**
    `Adapta topologia(Z);`
  fim
fim

Algoritmo 2: Otimização por Enxame de Partículas baseada em Vizinhança Complexa

Selecione aleatoriamente $(u, v) \in E(Z)$;
se $f(p_u) < f(p_v)$ **então** $n_{melhor} \leftarrow u$;
senão $n_{melhor} \leftarrow v$;
Remova (u, v) de $E(Z)$;
Selecione aleatoriamente $n_{destino} \in V(Z) - vizinhos(n_{melhor}) - \{n_{melhor}\}$;
Adicione $(n_{melhor}, n_{destino})$ em $E(Z)$;

Procedimento `Adapta topologia(Z)`, primeira versão (CNPSO-1)

para cada $j \in V(Z)$ **faça**
  se $estagnada_j$ **então**
    $C \leftarrow \bigcup_{k \in vizinhos(j)} vizinhos(k) - vizinhos(j) - \{j\}$;
    se $|C| = 0$ **então** $C \leftarrow V(Z) - vizinhos(j) - \{j\}$;
    Selecione aleatoriamente $k \in C$;
    Adicione (j, k) em $E(Z)$;
  senão
    $k \leftarrow \arg\max_{i \in vizinhos(j)} f(p_i)$;
    Remova (j, k) de $E(Z)$;
  fim
fim

Procedimento `Adapta topologia(Z)`, segunda versão (CNPSO+)

A quantidade de arestas no CNPSO+, por outro lado, é variável, de acordo com a evolução do processo de busca. Do ponto de vista de tempo de execução, o pior caso ocorre quando o enxame todo encontra-se estagnado por um período significativo, o que faria com que a topologia se tornasse um grafo completo. Nesta situação, a busca pela melhor influência social seria realizada em tempo $O(N)$ e o processo de adaptação tomaria tempo $O(N^3)$, onde N é o número de partículas no enxame. Em todo o caso, em situações em que o grau médio da rede permanece próximo do inicial, tem-se que a seleção do melhor vizinho poderia ser realizada em tempo constante e a adaptação topológica em tempo $O(N^2)$. É importante apontar, porém, que, apesar dos limites teóricos apresentados, observações empíricas indicam que o tempo de execução do CNPSO+ mantém-se próximo do tomado pela primeira versão do algoritmo, muitas vezes executando mais rapidamente que este.

5.2 Experimentos

Nesta subseção, são descritos os experimentos realizados para a avaliação de desempenho do CNPSO. Dado que este trabalho visa analisar os impactos do uso de vizinhanças complexas sobre o funcionamento de mecanismos de inteligência de enxame, buscou-se comparar os resultados obtidos pelo algoritmo proposto com aqueles de outras variantes da otimização por enxame de partículas. Assim sendo, foram avaliados, além do CNPSO e do PSO canônico, os algoritmos AR-PSO, *Bare Bones Particle Swarm*, FIPS, DMS-PSO, H-PSO e TRIBES, todos apresentados na seção 3.2. Tanto o *Bare Bones Particle Swarm* quanto a versão canônica foram testados usando as vizinhanças *gbest*, *lbest* e von Neumann, enquanto que o FIPS foi avaliado utilizando apenas a vizinhança von Neumann. As topologias *gbest* e *lbest* foram escolhidas por serem amplamente utilizadas na literatura, ao passo que a von Neumann foi selecionada por ser apontada como aquela que apresenta os melhores resultados. A não aplicação das vizinhanças global e local ao FIPS ocorre pois a literatura indica resultados inferiores para estas combinações. Apesar de a proposta original do AR-PSO ter abrangido apenas a topologia *gbest*, aqui foi analisada também sua utilização com vizinhança von Neumann, dados os resultados positivos obtidos por esta em outras variantes. Uma vez que os três primeiros algoritmos listados consistem apenas de modificações na forma de interação das partículas, mantendo intacta a topologia do enxame, foi avaliada também a combinação das dinâmicas utilizadas nestas propostas com a vizinhança do CNPSO. Como o TRIBES envolve modificações não só referentes à topologia, mas também à regra de atualização do PSO, foi utilizada neste estudo uma versão simplificada do algoritmo, preservando todos seus mecanismos de modificação da estrutura do enxame, mas usando as equações de cálculo de velocidade e de posição do PSO canônico. Tal restrição visa minimizar outras influências, comparando apenas o efeito exercido pela vizinhança.

Para garantir que todos os testes fossem realizados segundo os mesmos critérios e sob as mesmas

condições, todos os algoritmos escolhidos foram implementados pelo autor deste trabalho. Exceto para o *Bare Bones Particle Swarm*, que não utiliza tais parâmetros, foi adotado um fator de inércia ω decrescendo linearmente de acordo com a iteração atual, indo de $0,9$ até alcançar $0,4$ no final da busca, enquanto que $c_1 = c_2 = 1,496$. No FIPS não há tal divisão entre influência pessoal e social, de modo que, neste caso, $c = 2,992$. Para os parâmetros específicos de cada algoritmo, foram selecionados valores indicados na literatura como capazes de prover bons resultados, a saber: influência igual para todos os informantes de uma partícula no FIPS, sub-enxames com 3 partículas e reconstrução da topologia a cada 5 iterações no DMS-PSO, ocorrência de 4 ramos por nó no H-PSO e limitantes inferior e superior de diversidade de $5 * 10^{-6}$ e $0,25$, respectivamente, para o AR-PSO [157]. A inicialização das posições e velocidades das partículas foi feita de forma aleatória e uniforme, sendo as velocidades restritas a $|v_i^d| \leq \frac{X_{MAX}^d - X_{MIN}^d}{10}$ durante toda a otimização. Apesar de as posições iniciais serem limitadas, conforme indicado na tabela 5.2, durante os testes os locais visitados pelas partículas não foram contidos ao espaço de busca. Resultados obtidos fora deste, no entanto, não eram registrados como influências sociais ou pessoais, forçando, assim, as partículas a retornarem aos limites estabelecidos. O critério de parada utilizado, em todos os testes, foi o número de avaliações da função objetivo, limitadas a 100000.

Após ter os parâmetros de sua topologia ajustados, cada algoritmo foi executado 30 vezes para cada função de teste. A partir dos valores obtidos, foram calculadas suas médias e os índices de desempenho. Também foi definido um objetivo para cada função, delimitando o ponto a partir do qual uma busca pode ser considerada como obtendo sucesso. Assim, para cada bateria de testes de uma função também foi calculada a taxa de sucesso e a quantidade de avaliações da função objetivo esperadas para atingir a meta especificada, definida por:

$$AF_{esp} = \frac{< AF >}{t_{sucesso}}, \tag{5.1}$$

onde $< AF >$ é a média de avaliações da função objetivo até atingir a meta designada e $t_{sucesso}$ é a proporção de execuções que obtiveram sucesso na função em questão. Esta métrica é observada também em outros trabalhos [43, 90], sendo denominada, em alguns casos, taxa de desempenho. Quando um algoritmo obtém taxa de sucesso igual a 0% em alguma função de teste, são indefinidos tanto o valor de AF_{esp} para a função quanto sua média sobre todos os testes, sendo tais situações indicadas apenas por um traço ("–") nas tabelas de resultados. Todos os experimentos foram executados utilizando o ambiente Python 2.6.

Tab. 5.1: Funções de teste utilizadas para avaliação dos algoritmos.

Função	
Ackley	$f_1(x) = \sum_{i=1}^{N-1} (20 + e - 20e^{-0,2\sqrt{0,5({x_{i+1}}^2 + {x_i}^2)}} - e^{0,5(\cos(2\pi x_{i+1}) + cos(2\pi x_i))})$
Griewank	$f_2(x) = \frac{\sum_{i=1}^{N} {x_i}^2}{4000} - \prod_{i=1}^{N} \cos(\frac{x_i}{\sqrt{i}}) + 1$
Rastrigin	$f_3(x) = \sum_{i=1}^{N} ({x_i}^2 - 10\cos(2\pi x_i) + 10)$
Rosenbrock	$f_4(x) = \sum_{i=1}^{N-1} (100 * ({x_{i+1}} - {x_i}^2)^2 + (x_i - 1)^2)$
Salomon	$f_5(x) = 1 - \cos(2\pi\sqrt{\sum_{i=1}^{N} x_i^2}) + \sum_{i=1}^{N} \frac{x_i^2}{10}$
Schwefel	$f_6(x) = 418,9829N - \sum_{i=1}^{N} x_i \sin(\sqrt{\lvert x_i \rvert})$
Sphere	$f_7(x) = \sum_{i=1}^{N} {x_i}^2$

5.2.1 Funções de teste

Para a avaliação de desempenho dos algoritmos em questão, foram utilizadas sete funções de teste comumente adotadas na literatura para explorar suas capacidades, especialmente em problemas multimodais. Em todos os casos, os problemas foram avaliados em $N = 30$ dimensões, a exemplo de [97], sendo procurado o valor mínimo da função objetivo. As funções utilizadas podem ser vistas na tabela 5.1. A função Sphere é unimodal, consistindo apenas de uma bacia de atração que se estende por todo o espaço. A função Rosenbrock é uma função multimodal, apresentando dois mínimos locais para $N > 2$ [160]. Sua dificuldade reside na existência de um amplo platô no entorno do ótimo, levemente inclinado em direção a este. Por sua vez, as funções Ackley e Griewank possuem um grande número de mínimos locais, porém são consideradas funções multimodais simples. A função Rastrigin possui mínimos locais igualmente espaçados, enquanto que a Salomon é formada por anéis em torno do ótimo global, alternando regiões de mínimo e de máximo, de forma que ambas são consideradas difíceis para a maioria dos métodos de otimização. A função Schwefel é considerada particularmente complicada, dado que seu ótimo global localiza-se próximo aos limites do espaço de busca, sendo cercado por mínimos locais de baixa qualidade. Os gráficos das funções de teste, em duas dimensões, podem ser visualizados no apêndice A.

Dado que alguns algoritmos podem se beneficiar de situações em que o ótimo global situa-se na origem do eixo de coordenadas (geralmente adotado como o centro de dispersão das partículas) [113], o espaço de inicialização foi deslocado para algumas funções, o que também é útil para uma melhor avaliação da capacidade do enxame se mover em busca da solução. Os limites do espaço de busca e de inicialização para cada função de teste encontram-se listados na tabela 5.2, assim como a localização do mínimo global e o objetivo a ser alcançado. Verifica-se que, em todas as funções utilizadas, o seu valor ótimo é tal que $f_i(x^*) = 0$. Os valores de objetivos selecionados para as funções Ackley,

Tab. 5.2: Espaços de busca e de inicialização, posição do mínimo global e valores a serem alcançados para cada uma das funções de teste.

Função	Espaço de busca	Espaço de inicialização	Mínimo global	Objetivo
Ackley	$[-32,768; 32,768]^N$	$[-32,768; 16]^N$	$(0; ...; 0)$	0,1
Griewank	$[-500; 500]^N$	$[200; 500]^N$	$(0; ...; 0)$	0,1
Rastrigin	$[-5,12; 5,12]^N$	$[-5,12; 2]^N$	$(0; ...; 0)$	100
Rosenbrock	$[-30; 30]^N$	$[-30; 30]^N$	$(1; ...; 1)$	100
Salomon	$[-100; 100]^N$	$[10; 100]^N$	$(0; ...; 0)$	0,4
Schwefel	$[-500; 500]^N$	$[-500; 500]^N$	$(420,97; ...; 420,97)$	5000
Sphere	$[-100; 100]^N$	$[50; 100]^N$	$(0; ...; 0)$	0,01

Griewank, Rosenbrock e Sphere são comumente encontrados na literatura, como pode ser visto no trabalho de Janson e Middendorf [90]. Para as funções restantes, foi definido um valor que fornecesse um limitante adequado frente aos resultados obtidos por outros PSOs.

5.2.2 Ajustes da topologia

Antes de proceder a comparação de desempenho dos algoritmos, foi necessário estabelecer quais arranjos topológicos seriam capazes de produzir bons resultados. Tal etapa é de particular importância para o CNPSO, uma vez que ela permite averiguar qual sua sensibilidade a ajustes paramétricos.

Assim, um primeiro passo para o CNPSO é analisar qual o impacto sobre seu desempenho do uso de diferentes formas de geração de redes complexas no estabelecimento da topologia inicial do enxame. Para tal estudo, foi comparada a aplicação do modelo livre-de-escala (SF), proposto por Barabási e Albert [13], com a de outros dois modelos baseados na conexão de vizinhos próximos, capazes de gerar redes com coeficientes de aglomeração mais realistas, a saber, os modelos HK, proposto por Holme e Kim [87], e DEB, proposto por Davidsen *et al.* [46]. Tal comparação foi feita produzindo enxames com 250 partículas e grau médio de 14 conexões, valores capazes de produzir bons resultados, de acordo com análises previamente realizadas pelo autor. Nesta parte dos experimentos, foi suprimido o mecanismo de adaptação topológica, para que os efeitos da escolha da topologia inicial fossem maximizados. Escolhida a forma de geração da rede inicial, buscou-se, então, avaliar qual a influência do número de partículas e do número de conexões sobre o desempenho do algoritmo. Para isso, foram testadas todas as combinações de tamanhos de enxame $|V| \in \{25, 50, 100, 250, 500, 1000\}$ e graus médios $< k > \in \{2, 4, 6, 8, 10, 12, 14, 16\}$, adaptando a topologia a cada 80 iterações (valores razoáveis, de acordo com testes prévios), tanto para o CNPSO-1 quanto para o CNPSO+. Para melhor averiguar os efeitos do uso das topologias dinâmicas apresentadas, nesta etapa foi analisada também

uma versão rígida do CNPSO, ou seja, com uma vizinhança fixa. Por fim, após esta seleção, foi analisada a sensibilidade do CNPSO à alteração do parâmetro $periodo$, comparando resultados obtidos quando realizadas 0, 3, 7, 20, 50, 150 e 400 adaptações topológicas no decorrer de cada busca. Os mesmos procedimentos de ajuste se repetiram para as versões híbridas do CNPSO.

Os demais algoritmos também receberam alguns ajustes. Cada um deles foi analisado com número de partículas $|V| \in \{25, 50, 100, 250, 500, 1000\}$[1], sendo utilizados na comparação final os tamanhos de enxame que tiveram os melhores desempenhos para cada algoritmo testado. É importante ressaltar que, para evitar que os resultados fossem prejudicados por alguma bateria de testes que tenha obtido resultados excepcionalmente bons, em cada etapa dos experimentos foram executados todos os testes indicados, mesmo quando dados para a execução de um determinado algoritmo com um certo conjunto de parâmetros já tivessem sido obtidos em fases anteriores.

5.2.3 Índice de desempenho

Na literatura, são encontrados alguns índices de desempenho que buscam combinar adequadamente todos os resultados obtidos por um algoritmo (ou conjunto de parâmetros) testado para múltiplas funções. Um exemplo é o proposto por Mendes *et al.* [127], no qual tal valor é obtido pela média dos resultados padronizados conseguidos pelo algoritmo. Nesta padronização, todos os algoritmos que estão sendo comparados têm seus valores escalados da mesma maneira, de forma que, para cada função, a média dos resultados seja 0 e seu desvio-padrão, 1. O problema desta abordagem é que o índice de desempenho pode variar radicalmente para um mesmo conjunto de resultados, dependendo de com quais outros valores ele está sendo comparado. Assim, um conjunto de resultados A pode tanto ser considerado como melhor quanto ser visto como pior do que um conjunto de resultados B, de acordo com outros resultados que estão sendo analisados juntamente.

Neste trabalho, foi definido um índice de desempenho no qual cada função é escalada de forma diferente, conforme os resultados para ela esperados. Cada algoritmo testado recebe, então, uma nota d_f para cada uma das funções f analisadas, dada por:

$$d_f = \frac{1}{1 + \frac{f_{objetivo}}{f_{media}}}, \tag{5.2}$$

onde $f_{objetivo}$ é a meta definida (vide tabela 5.2) e f_{media} é a média dos resultados obtidos para f no teste em questão. O índice de desempenho final do algoritmo é obtido, assim, através da média de d_f sobre todas as funções utilizadas. Os valores de d_f são restritos ao intervalo $[0, 1]$, de modo que, se algum algoritmo fornecer resultados muito bons ou muito ruins para uma função específica, eles

[1] As topologias von Neumann testadas, em todos os casos, possuiam duas dimensões definidas pelo par de divisores de $|V|$ mais próximos entre si.

não terão uma influência tão grande no índice que possa mascarar o desempenho em outras funções. Note que, para casos de minimização, quanto mais próximo de 0 o índice d_f, melhores os resultados obtidos.

Como buscou-se nos experimentos realizados avaliar a habilidade dos algoritmos em fornecer bons resultados em diferentes funções, evitando, assim, a necessidade de ajustes específicos a cada problema abordado, o índice aqui apresentado foi utilizado para definir qual arranjo de parâmetros é capaz de prover os melhores resultados para cada algoritmo.

5.3 Ajustes de parâmetros

Nesta seção, encontram-se os resultados relativos ao ajustes de parâmetros relativos à topologia, tanto para as versões do CNPSO quanto para as variantes presentes na literatura. Ressalta-se que aqui são exibidos, em sua maioria, índices e valores médios que buscam apontar de forma mais geral o funcionamento dos algoritmos testados. As tabelas aqui exibidas foram coloridas de modo a facilitar suas análises: a cor vermelha indica os melhores resultados, enquando valores mais negativos são indicados pelas cores verde, em situações em que são variados dois parâmetros, e branca, quando apenas um parâmetro é modificado.

5.3.1 CNPSO e supressão de ajustes topológicos

Geração da rede inicial

Conforme apontado anteriormente, as versões do CNPSO apresentadas neste capítulo não definem, *a priori*, um modelo de geração da topologia inicial do enxame, sendo este um parâmetro dos algoritmos. Os resultados dos testes realizados[2] são indicados a seguir.

Tab. 5.3: Testes de modelos de geração da rede inicial do CNPSO – Resultados médios e indices de desempenho (d_f).

	Funções							
Modelo	Salomon	Rastrigin	Schwefel	Rosenbrock	Sphere	Griewank	Ackley	Desempenho
SF	2,1987E-01	3,3663E+01	4,3265E+03	3,5909E+01	1,6067E-09	4,0239E-03	1,7212E-01	0,2866
DEB	2,0998E-01	3,5454E+01	4,4482E+03	3,5564E+01	1,1591E-09	7,6337E-03	9,4099E-05	0,2016
HK	2,2660E-01	3,5188E+01	4,3016E+03	4,1948E+01	1,7485E-09	6,3964E-03	6,5214E-01	0,3296

[2]No modelo DEB, a probabilidade de remoção de um nó, substituindo-o por um novo, foi definida igual a 1%, e no modelo HK cada novo nó tem 50% de chances de se conectar a um vizinho de segundo grau.

Tab. 5.4: Testes de modelos de geração da rede inicial do CNPSO – Desvios-padrão.

	Funções						
Modelo	Salomon	Rastrigin	Schwefel	Rosenbrock	Sphere	Griewank	Ackley
SF	4,0000E-02	7,9724E+00	7,5785E+02	2,7330E+01	9,1716E-10	5,4871E-03	9,2621E-01
DEB	2,9969E-02	8,2186E+00	4,7360E+02	2,3507E+01	5,0509E-10	8,3646E-03	2,1953E-05
HK	4,4185E-02	8,0153E+00	7,2766E+02	3,1957E+01	1,2247E-09	1,1129E-02	3,5113E+00

Tab. 5.5: Testes de modelos de geração da rede inicial do CNPSO – Taxas de sucesso.

	Funções							
Modelo	Salomon	Rastrigin	Schwefel	Rosenbrock	Sphere	Griewank	Ackley	Média
SF	100%	100%	83%	97%	100%	100%	97%	97%
DEB	100%	100%	87%	97%	100%	100%	100%	98%
HK	100%	100%	87%	93%	100%	100%	97%	97%

Tab. 5.6: Testes de modelos de geração da rede inicial do CNPSO – Avaliações de função esperadas até atingir o objetivo (AF_{esp}).

Modelo	Salomon	Rastrigin	Schwefel	Rosenbrock	Sphere	Griewank	Ackley	Média
SF	63783,3	33066,7	39996,0	57859,7	67316,7	62425,0	77951,8	57485,6
DEB	64941,7	32058,3	42315,1	59135,0	67116,7	62200,0	74625,0	57484,5
HK	63850,0	33483,3	40473,4	60545,3	67916,7	62591,7	77711,1	58081,6

Comparando os resultados por função das tabelas 5.3, 5.4, 5.5 e 5.6, não é possível observar grandes diferenças entre os modelos de geração utilizados. À exceção da função Ackley, todos os modelos atingiram resultados médios com a mesma ordem de grandeza e com velocidades esperadas de convergência próximas, tendo uma diferença máxima de 5,8%, para a função Schwefel. No entanto, nota-se que os enxames com topologias geradas pelo modelo DEB obtiveram os melhores resultados em 4 das 7 funções de teste utilizadas, situação refletida no seu índice de desempenho. Analisando os desvios-padrão e as taxas de sucesso verifica-se, também, que este modelo foi capaz de prover resultados mais regulares, além de possuir convergência levemente mais rápida.

Tamanho e grau médio do enxame

Escolhido o modelo DEB para a geração da topologia inicial das partículas, procedeu-se a análise de sensibilidade das versões do CNPSO à variação do tamanho do enxame e de sua quantidade de conexões. Para facilitar a comparação, em cada uma das métricas analisadas, para cada combinação de parâmetros está marcada em negrito a versão do CNPSO que obteve os melhores resultados.

Nos resultados do CNPSO rígido (tabelas 5.7, 5.8 e 5.9) observa-se, no geral, uma tendência a me-

Tab. 5.7: Testes de tamanhos do enxame e graus médios para a versão do CNPSO sem atualizações topológicas – Índices de desempenho (d_f). Ressalta-se que menores valores do índice de desempenho indicam melhores resultados.

	Número de partículas					
Grau médio	25	50	100	250	500	1000
2	0,3985	0,3793	0,3640	0,3713	0,7193	0,8102
4	0,3770	0,3400	**0,2015**	0,2145	0,5720	0,7434
6	**0,3716**	0,3520	**0,2045**	0,1955	0,4529	0,7085
8	0,3693	0,3579	**0,2047**	0,2889	0,3739	0,6836
10	**0,3756**	**0,3566**	**0,3002**	0,2841	0,3241	0,6705
12	**0,3795**	0,3635	0,2125	0,2013	0,3070	0,6550
14	0,3868	0,3759	0,3479	0,2009	0,2920	0,6486
16	0,3856	**0,3607**	0,3478	0,3313	0,2791	0,6328

Tab. 5.8: Testes de tamanhos do enxame e graus médios para a versão do CNPSO sem atualizações topológicas – Taxas de sucesso.

	Número de partículas					
Grau médio	25	50	100	250	500	1000
2	79%	83%	87%	67%	31%	21%
4	86%	**95%**	97%	99%	48%	26%
6	**86%**	**93%**	97%	99%	67%	30%
8	87%	**92%**	**98%**	**99%**	84%	35%
10	83%	**93%**	**98%**	**99%**	85%	40%
12	82%	90%	**97%**	98%	84%	41%
14	82%	89%	**96%**	96%	86%	42%
16	**84%**	**92%**	92%	**97%**	90%	42%

lhores valores quando utilizados enxames maiores. Um exemplo ocorre com a função Schwefel, para a qual são conseguidos resultados significativamente melhores usando enxames com 1000 partículas, em especial devido a sua alta multimodalidade, o que demanda maior capacidade de exploração. Esta situação, porém, não é observada para as funções com menos ótimos locais: para a função Rosenbrock, em que se verifica uma baixa sensibilidade a esses parâmetros, e a função Sphere, que tem seus melhores resultados com enxames menores, especificamente com 50 partículas. Ainda, observa-se que, para cada tamanho, há uma densidade de conexões específica capaz de prover os melhores resultados, crescendo juntamente com a quantidade de partículas.

Confrontando índices de desempenho e taxas de sucesso, verifica-se, para situações com menos partículas, uma maior sensibilidade à variação do tamanho do enxame do que à modificação de sua densidade de conexões. Tal dependência se modifica em enxames com 500 ou mais partículas, onde

Tab. 5.9: Testes de tamanhos do enxame e graus médios para a versão do CNPSO sem atualizações topológicas – Médias das avaliações de funções esperadas até atingir o objetivo (AF_{esp}).

	Número de partículas					
Grau médio	25	50	100	250	500	1000
2	51156,9	54360,6	70030,8	–	–	–
4	36159,9	38746,5	46939,6	66391,2	–	–
6	**35137,3**	37073,1	44436,1	62138,0	–	–
8	**33044,4**	36708,5	42673,3	59924,6	–	–
10	37308,2	35625,1	42453,7	57954,5	–	–
12	36442,8	38329,3	41752,6	57908,5	–	–
14	37789,4	37795,6	41288,6	57959,9	131943,1	–
16	**33526,0**	**35020,9**	42301,5	57395,5	94960,0	–

tornam-se evidentes os benefícios do aumento da quantidade de conexões para além dos limites analisados. Dentre as combinações de parâmetros comparadas, observa-se que os melhores resultados são obtidos para enxames de tamanho 100 ou 250, sendo o melhor índice de desempenho obtido pela configuração com 250 partículas e grau médio igual a 6 (tabela 5.7).

Tab. 5.10: Testes de tamanhos do enxame e graus médios para a primeira versão do algoritmo - CNPSO-1 – Índices de desempenho (d_f).

	Número de partículas					
Grau médio	25	50	100	250	500	1000
2	**0,3786**	**0,3504**	**0,2148**	0,2837	0,6901	0,8009
4	**0,3635**	**0,3282**	0,2111	0,2065	0,5551	0,7415
6	0,3727	**0,3515**	0,2979	0,1964	0,4361	0,7072
8	**0,3638**	**0,3386**	0,3035	**0,1945**	0,3638	0,6828
10	0,3828	0,3623	0,3271	**0,1912**	0,3271	0,6657
12	0,3815	**0,3588**	0,3228	0,1997	0,3092	0,6574
14	**0,3781**	**0,3559**	**0,3213**	**0,1994**	0,2872	0,6456
16	0,3852	0,3660	**0,3261**	0,2046	0,2837	0,6409

Para o CNPSO-1 (tabelas 5.10, 5.11 e 5.12) são observados resultados bem semelhantes àqueles obtidos com a topologia rígida. Também aqui os índices de desempenho e taxas de sucesso atingem seus melhores valores para enxames com entre 100 e 250 partículas e ao aumentar a quantidade de conexões, para enxames maiores. O melhor desempenho foi alcançado quando utilizadas 250 partículas, com grau médio igual a 10, uma configuração bem próxima daquela considerada a melhor para o CNPSO rígido. Também a exemplo do teste anterior, é possível observar uma maior rapidez do enxame em atingir o objetivo estabelecido quando ele possui menos partículas e mais conexões,

Tab. 5.11: Testes de tamanhos do enxame e graus médios para a primeira versão do algoritmo - CNPSO-1 – Taxas de sucesso.

	Número de partículas					
Grau médio	25	50	100	250	500	1000
2	**88%**	**91%**	95%	90%	37%	22%
4	**88%**	94%	**98%**	**100%**	50%	28%
6	85%	**93%**	**98%**	**100%**	72%	30%
8	**87%**	92%	96%	**99%**	85%	39%
10	**83%**	89%	95%	99%	86%	42%
12	**84%**	**93%**	94%	**99%**	86%	40%
14	**83%**	**91%**	95%	96%	87%	42%
16	81%	89%	**93%**	96%	90%	40%

Tab. 5.12: Testes de tamanhos do enxame e graus médios para a primeira versão do algoritmo - CNPSO-1 – Médias das avaliações de funções esperadas até atingir o objetivo (AF_{esp}).

	Número de partículas					
Grau médio	25	50	100	250	500	1000
2	40560,9	42911,7	53084,0	95233,2	–	–
4	35841,3	37390,6	46085,2	65219,6	–	–
6	36496,2	37435,8	44369,7	61120,2	–	–
8	33598,7	36839,2	44279,8	59597,5	–	–
10	**34923,0**	38331,8	43676,0	57840,7	–	–
12	**35551,4**	35897,1	42680,9	56821,3	485161,1	–
14	**35612,5**	35504,6	42311,2	57786,0	167395,3	–
16	36557,4	36600,7	42569,2	57042,6	95254,5	–

decrescendo a quantidade média de avaliações esperadas da função objetivo, aparentemente até um limite em torno de 30000.

Porém, ao contrário do verificado para a versão rígida, nos testes do CNPSO-1 não ocorreram quaisquer variações bruscas nos resultados para pequenas modificações paramétricas, o que indica maior estabilidade deste algoritmo, uma vez que tais flutuações são usualmente decorrentes de algumas execuções com resultados de qualidade muito mais baixa que o esperado. Uma segunda diferença que vale ser relatada é a grande melhoria dos resultados para a função Sphere, advinda do mecanismo de seleção de nós mais influentes, que garante mais conexões para partículas com melhores históricos, proporcionando uma melhor busca local.

Observando os resultados obtidos pelo CNPSO+ (tabelas 5.13, 5.14 e 5.15), uma diferença drástica com relação às outras versões apresentadas é o fato de ser praticamente eliminada a relação entre o desempenho do algoritmo e a densidade de conexões da topologia inicial, restando apenas a depen-

Tab. 5.13: Testes de tamanhos do enxame e graus médios para a segunda versão do algoritmo - CNPSO+ – Índices de desempenho (d_f).

	Número de partículas					
Grau médio	25	50	100	250	500	1000
2	0,3799	0,3511	0,3115	**0,1924**	**0,2677**	**0,6500**
4	0,3724	0,3688	0,2095	**0,2018**	**0,2619**	**0,6459**
6	0,3854	0,3657	0,3156	**0,1952**	**0,2597**	**0,6415**
8	0,3730	0,3653	0,3236	0,1987	**0,2526**	**0,6333**
10	0,3875	0,3653	0,3445	0,1956	**0,2490**	**0,6287**
12	0,3886	0,3614	**0,2089**	**0,1962**	**0,2401**	**0,6280**
14	0,3905	0,3746	0,3528	0,2005	**0,2374**	**0,6188**
16	**0,3848**	0,3762	0,3266	**0,1989**	**0,2405**	**0,6231**

Tab. 5.14: Testes de tamanhos do enxame e graus médios para a segunda versão do algoritmo - CNPSO+ – Taxas de sucesso.

	Número de partículas					
Grau médio	25	50	100	250	500	1000
2	83%	91%	**95%**	**98%**	**92%**	**41%**
4	83%	91%	95%	96%	**95%**	**40%**
6	79%	91%	94%	96%	**96%**	**40%**
8	84%	**92%**	96%	96%	**96%**	**41%**
10	79%	90%	94%	97%	**98%**	**46%**
12	78%	91%	94%	97%	**99%**	**43%**
14	76%	91%	92%	**97%**	**97%**	**46%**
16	82%	89%	92%	95%	**97%**	**43%**

Tab. 5.15: Testes de tamanhos do enxame e graus médios para a segunda versão do algoritmo - CNPSO+ – Médias das avaliações de funções esperadas até atingir o objetivo (AF_{esp}).

	Número de partículas					
Grau médio	25	50	100	250	500	1000
2	**34401,2**	**35348,8**	**41926,3**	**56817,5**	**84987,3**	–
4	**33948,4**	**34952,3**	**41899,8**	**56488,1**	**78313,4**	–
6	37128,6	**34776,7**	**41486,3**	**57338,1**	**76894,3**	–
8	33145,0	**34850,0**	**40203,0**	**56077,7**	**75484,4**	–
10	38759,9	**34632,1**	**40922,9**	**54977,9**	**72892,5**	–
12	38807,7	**34112,4**	**40689,2**	**54156,7**	**71874,7**	–
14	39918,2	**33758,4**	**41136,7**	**54525,8**	**72996,9**	–
16	33633,8	35582,2	**41134,5**	**54333,0**	**71596,1**	–

dência sobre a quantidade de partículas. Ressalta-se que a redução na sensibilidade à quantidade inicial de conexões na rede ocorre de forma que, na maior parte dos casos, os resultados médios obtidos pelo CNPSO+ são similares aos melhores alcançados pelas propostas anteriores, quando comparados dados para um mesmo tamanho de enxame. Esta baixa dependência é advinda da flexibilização da quantidade de arestas, uma vez que, ao contrário do CNPSO-1, no CNPSO+ o enxame pode, no decorrer das iterações, não apenas modificar quais os nós mais influentes, mas também ajustar sua densidade, de acordo com a ocorrência, ou não, de estagnação das partículas. Em todo o caso, analisando as funções de forma isolada, ainda é possível notar uma leve dependência sobre o grau médio para enxames maiores. Para 1000 partículas, por exemplo, observa-se tal situação para as funções Salomon, Sphere, Rosenbrock, Griewank e Ackley, não sendo isto plenamente captado pelo índice de desempenho, dado que as variações são relativamente pequenas e os resultados obtidos são de baixa qualidade. Tal relação ocorre pois, para este tamanho de enxame, há poucas iterações disponíveis – apenas 100 –, de modo que não há espaço para deixar apenas para o final da execução uma busca local mais eficiente, induzida pela adição de conexões à rede, assim como a possibilidade de variação de densidade mantém-se relativamente baixa, podendo ser alterada em, no máximo, 5%.

Outros pontos que saltam aos olhos com relação aos CNPSOs anteriores são os índices de desempenho e taxas de sucesso muito melhores obtidos por enxames com 500 e 1000 partículas e o aumento da velocidade de convergência em quase todas as configurações paramétricas testadas, com a quase independência entre essa velocidade e a quantidade de conexões na rede inicial. Esta maior rapidez apenas não ocorre para enxames com 25 partículas, para os quais as quantidades de avaliações de funções esperadas são similares às das versões anteriores. Como contrapartida, porém, verifica-se uma leve redução na taxa de sucesso para enxames com menos de 500 partículas, assim como uma piora nos resultados obtidos para a função Sphere, especialmente quando o algoritmo é comparado ao CNPSO-1. Com a exceção das funções mais unimodais Sphere e Rosenbrock, os melhores resultados do CNPSO+ foram sempre obtidos para enxames com 250 ou 500 partículas. A configuração capaz de prover o melhor índice de desempenho foi aquela com 250 nós e grau médio igual a 2, uma rede inicial significativamente mais esparsa que as encontradas para os outros CNPSOs.

Ao se comparar os índices de desempenho médios obtidos por cada algoritmo (tabelas 5.7, 5.10 e 5.13), o CNPSO+ mostra-se melhor que seus pares, ficando o CNPSO-1 em segundo lugar. Comparando os resultados obtidos pelas três versões para cada configuração paramétrica, novamente esta situação se repete, com o CNPSO+ tendo 69% de seus resultados melhores do que aqueles conseguidos pela versão sem adaptações topológicas, enquanto que 67% dos resultados do CNPSO-1 superaram o de sua versão rígida. Tais fatos indicam uma influência positiva das adaptações topológicas propostas, ao menos para a redução, sem perda de desempenho, da sensibilidade do mecanismo de otimização à variação de parâmetros.

Confome indicado no capítulo 3, na literatura encontram-se diversas propostas apontando benefícios de um simples aumento na densidade do enxame conforme a busca se aproxima de seu final, como maneira de intensificar o refinamento das boas soluções encontradas. Tal adição simples de conexões poderia ser a responsável pelos resultados positivos do CNPSO+, de forma que tanto a verificação de estagnação das partículas quanto a possibilidade de remoção de arestes poderiam ser desnecessárias ou até negativas. Resultados contrários a essa hipótese, no entanto, foram encontrados em testes prévios realizados, mostrando uma piora de desempenho quando, a cada adaptação, a partícula apenas podia ganhar conexões.

Quantidade de adaptações topológicas

Após a escolha do número de partículas e de conexões da topologia inicial para as três versões do CNPSO, foram averiguados os efeitos da variação da quantidade de adaptações topológicas sofridas pelo enxame no decorrer da busca. O número máximo de 400 adaptações foi utilizado por ser igual ao limite de iterações para enxames com 250 partículas.

Tab. 5.16: Testes de quantidade de adaptações topológicas para as versões do CNPSO – Índices de desempenho (d_f).

	Adaptações						
Algoritmo	0	3	7	20	50	150	400
CNPSO-1	0,1915	0,1915	0,1930	0,1985	0,2803	0,1905	0,1941
CNPSO+	0,3822	0,2245	0,2030	0,1932	0,1880	0,1932	0,2992

Tab. 5.17: Testes de quantidade de adaptações topológicas para as versões do CNPSO – Taxas de sucesso.

	Adaptações						
Algoritmo	0	3	7	20	50	150	400
CNPSO-1	98%	99%	100%	98%	99%	98%	99%
CNPSO+	70%	99%	98%	97%	97%	96%	95%

Comparando os índices de desempenho (tabela 5.16), verifica-se uma divergência entre as relações das duas versões do CNPSO com a quantidade de adaptações topológicas: enquanto que a primeira versão mostra-se pouco sensível ao aumento ou à redução da quantidade de modificações na rede de influências, não sendo possível observar qualquer tendência em nenhuma das funções testadas, o CNPSO+ apresenta resultados que variam fortemente com este parâmetro. Para este, ao

Tab. 5.18: Testes de quantidade de adaptações topológicas para as versões do CNPSO – Médias das avaliações de funções esperadas até atingir o objetivo (AF_{esp}).

	Adaptações						
Algoritmo	0	3	7	20	50	150	400
CNPSO-1	58144,3	58553,1	57758,4	58563,0	57436,2	58807,7	58762,3
CNPSO+	–	69610,1	68286,3	64331,0	59795,6	55360,7	52787,8

aumentar-se a quantidade de adaptações, é observada uma melhoria contínua até um ponto de inflexão, em 50 adaptações, após o qual há uma piora do desempenho geral. As taxas de sucesso para os dois algoritmos, por outro lado, permanecem com valores próximos em quase todo o intervalo analisado, havendo uma única exceção quando utilizado o CNPSO+ sem adaptações. Isso ocorre pois, nesta situação, a topologia inicial utilizada, muito esparsa para esta versão do algoritmo, é mantida durante toda a busca. Apesar da proximidade destas taxas, é possível notar na tabela 5.17 um declínio suave do sucesso conforme são realizadas mais modificações no CNPSO+. Com relação à velocidade de convergência (tabela 5.18), assim como às outras métricas analisadas, o CNPSO-1 não mostrou mudanças significativas ao alterar a rede mais vezes, enquanto que o CNPSO+ indicou que, quanto mais modificações fossem realizadas, maior sua rapidez para atingir o objetivo estabelecido.

Para o CNPSO+, o contraste apresentado entre a taxa de sucesso e o índice de desempenho indica que todos os arranjos foram capazes de atingir o objetivo em proporções similares. Porém as configurações com 20 e 50 adaptações foram aquelas que conseguiram melhor refinar os resultados obtidos. Tais configurações proveem uma solução de compromisso adequada, apresentando resultados de boa qualidade para todas as funções testadas, mesmo com o algoritmo em questão possuindo uma quantidade adequada de adaptações diferente para cada função, divergência esta que pode ser vista nos comportamentos para a função Sphere, onde quanto mais a topologia é modificada melhores os resultados obtidos, e para a função Schwefel, que tem desempenho um pouco superior quando há menos adaptações topológicas.

Por fim, aponta-se que, após analisar os índices de desempenho da tabela 5.16, foram fixados em 150 e 50, respectivamente, os números de adaptações topológicas do CNPSO-1 e do CNPSO+.

5.3.2 Propostas vistas na literatura

Findos os ajustes paramétricos para as versões do CNPSO, foram realizados os testes para estabelecer quais os tamanhos de enxame a serem utilizados para as versões do PSO presentes na literatura, com as quais o CNPSO será comparado.

Para os testes realizados utilizando vizinhança global (tabelas 5.19, 5.20 e 5.21), observou-se uma

Tab. 5.19: Testes de tamanhos do enxame para outros PSOs, vistos na literatura – Índices de desempenho (d_f).

		Número de partículas					
Algoritmo	Vizinhança	25	50	100	250	500	1000
PSO	Global	0,4049	0,3988	0,3988	0,3892	0,3947	0,4321
	Local	0,2174	0,2109	0,2122	0,5124	0,7520	0,8342
	von Neumann	0,3249	0,1986	0,1838	0,2084	0,5922	0,7500
FIPS	von Neumann	0,2176	0,2266	0,2472	0,4066	0,7472	0,7825
Bare Bones PSO	Global	0,3962	0,3663	0,3459	0,3883	0,7628	0,8505
	Local	0,3837	0,2708	0,5005	0,8435	0,8766	0,8938
	von Neumann	0,3789	0,3575	0,2612	0,8083	0,8583	0,8861
H-PSO		0,3667	0,3270	0,1996	0,2116	0,5176	0,7678
DMS-PSO		0,3456	0,1816	0,1719	0,1870	0,4348	0,7129
AR-PSO	Global	0,5193	0,4698	0,4323	0,4270	0,4380	0,4889
	von Neumann	0,3534	0,2932	0,2740	0,2721	0,5864	0,7482

Tab. 5.20: Testes de tamanhos do enxame para outros PSOs, vistos na literatura – Taxas de sucesso.

		Número de partículas					
Algoritmo	Vizinhança	25	50	100	250	500	1000
PSO	Global	70%	78%	75%	78%	75%	73%
	Local	94%	97%	95%	58%	28%	20%
	von Neumann	92%	97%	97%	100%	43%	29%
FIPS	von Neumann	86%	86%	80%	62%	15%	14%
Bare Bones PSO	Global	83%	93%	97%	80%	17%	6%
	Local	82%	73%	53%	13%	0%	0%
	von Neumann	88%	93%	78%	14%	10%	0%
H-PSO		83%	87%	88%	85%	52%	0%
DMS-PSO		93%	98%	95%	92%	63%	28%
AR-PSO	Global	61%	69%	74%	75%	73%	66%
	von Neumann	83%	87%	89%	95%	43%	28%

baixa sensibilidade do índice de desempenho e da taxa de sucesso à variação do enxame, em especial quando esta topologia foi combinada com o PSO canônico e o AR-PSO. Apesar de tal topologia apresentar valores melhores que seus pares ao utilizar enxames maiores, em nenhuma das configurações foi exibido um resultado satisfatório, sendo particularmente negativos os valores obtidos para a função Ackley. A topologia local, por outro lado, foi capaz de obter resultados com boa qualidade, sempre com populações menores, perdendo efetividade ao aumentar a topologia. Tal relação indica que, nesta situação, os benefícios obtidos para funções multimodais com o aumento da população, e sua subsequente maior diversidade, são contrabalançados pelo crescimento mais rápido na distância

Tab. 5.21: Testes de tamanhos do enxame para outros PSOs, vistos na literatura – Médias das avaliações de funções esperadas até atingir o objetivo (AF_{esp}).

		Número de partículas					
Algoritmo	Vizinhança	25	50	100	250	500	1000
PSO	Global	49569,9	49293,0	69904,7	72746,8	150743,7	–
	Local	34440,1	42829,3	60873,9	–	–	–
	von Neumann	35969,6	37161,1	47032,9	66494,0	–	–
FIPS	von Neumann	–	–	–	–	–	–
Bare Bones PSO	Global	27185,5	29494,8	41390,4	136512,7	–	–
	Local	57717,0	512805,6	–	–	–	–
	von Neumann	50117,1	46641,5	–	–	–	–
H-PSO		40814,1	52629,4	85909,2	–	–	–
DMS-PSO		34103,8	41276,3	61081,3	96533,1	–	–
AR-PSO	Global	–	–	127454,1	–	–	–
	von Neumann	63155,2	49906,7	66110,7	72992,8	–	–

das partículas nesta vizinhança. Conforme o esperado, a topologia von Neumann foi capaz de superar as duas vizinhanças anteriores, quando se analisando o índice de desempenho de suas melhores configurações. Comparando a sensibilidade ao tamanho do enxame, verifica-se que esta topologia obtém resultados melhores para 50 a 250 partículas, deteriorando-se bruscamente para enxames com mais de 500 elementos, uma situação similar à do DMS-PSO, mas contrária à vista no CNPSO+. Estima-se que tal degeneração deva-se ao fato de o grau de cada nó ser fixo em 4 conexões, o que causa lentidão no fluxo de informações em populações maiores. Uma dependência paramétrica bem similar à vista no CNPSO-1 e em sua versão rígida (para graus médios menores) é observada para o H-PSO, fato possivelmente advindo de densidades de conexões parecidas, da presença do efeito *small-world* e de a conectividade nas três topologias depender de poucos nós (*hubs*).

Comparando-se o PSO canônico com o AR-PSO, nota-se que, para uma mesma vizinhança, os resultados conseguidos pela versão com repulsão são substancialmente piores. Conforme esperado, há melhoria de desempenho para a função Schwefel, sendo este o único benefício registrado. No entanto, ocorre uma notável piora para as funções Sphere, Ackley e, em menor grau, Griewank. Particularmente, a função Ackley mostra-se um ponto fraco para o AR-PSO, dadas as baixas taxas de sucesso obtidas pela proposta, algo visto de forma mais intensa quando usada a topologia global. Outra variante que mostrou-se incapaz de obter bons resultados foi o FIPS, não conseguindo, inclusive, sequer uma execução de sucesso para a função Schwefel em todos os testes realizados.

Através da análise dos índices de desempenho (tabela 5.19), escolheram-se as seguintes configurações paramétricas: FIPS com 25 partículas; PSO e *Bare Bones Particle Swarm* locais com 50 partículas; PSO von Neumann, *Bare Bones* global e von Neumann, H-PSO e DMS-PSO com 100

partículas; PSO global e AR-PSO global e von Neumann com 250 partículas. Como pode ser observado, em mais de 70% dos casos testados, os enxames com melhores desempenhos são aqueles com 100 ou mais elementos, o que contraria a proposta de se restringir os enxames de PSO aos tamanhos tidos como clássicos, entre 20 e 50 partículas, especialmente ao trabalhar com problemas de maior dimensionalidade.

Comparando os resultados aqui vistos com os do CNPSO+, observa-se que o algoritmo proposto neste trabalho produz índices de desempenho quase sempre melhores para populações com mais elementos, sendo capaz de manter taxas de sucesso superiores a 90% para enxames com 500 partículas, uma situação em que 7 das 11 outras variantes analisadas tiveram sucesso menos de metade das execuções, havendo, inclusive, algoritmos incapazes de obter sucesso em qualquer das funções de teste.

5.3.3 Versões híbridas

Dado seu comportamento exibir aspectos mais interessantes, foi escolhida a segunda proposta de adaptação topológica do CNPSO para ser utilizada na definição de vizinhança no FIPS, no *Bare Bones Particle Swarm* e no AR-PSO. Segue-se a análise de sensibilidade destas versões híbridas à definição de tamanho de enxame, grau médio e quantidade de adaptações.

Tamanho e grau médio do enxame

Tendo em mãos os parâmetros adequados obtidos para o CNPSO+, foram utilizadas 50 adaptações topológicas nos testes de tamanho e densidade de enxame para os modelos híbridos.

Tab. 5.22: Testes de tamanhos do enxame e graus médios para a versão híbrida do FIPS com o CNPSO+ – Índices de desempenho (d_f).

	Número de partículas					
Grau médio	25	50	100	250	500	1000
2	0,4116	0,4392	0,4134	0,4286	0,4199	0,6064
4	0,4186	0,4043	0,4207	0,4564	0,4604	0,6333
6	0,4342	0,4388	0,4625	0,4794	0,4763	0,6082
8	0,4694	0,4674	0,4940	0,4991	0,4930	0,5876
10	0,5504	0,4793	0,4984	0,5057	0,4996	0,5812
12	0,6980	0,4928	0,5004	0,5054	0,5160	0,6019
14	0,7537	0,6293	0,5325	0,5998	0,5602	0,6419
16	0,7556	0,7378	0,6889	0,7384	0,7477	0,6838

Analisando a relação entre resultados e parâmetros topológicos (tabelas 5.22 e 5.23), é possível observar que o híbrido entre o CNPSO+ e o FIPS apresenta pouca dependência sobre o tamanho da

Tab. 5.23: Testes de tamanhos do enxame e graus médios para a versão híbrida do FIPS com o CNPSO+ – Taxas de sucesso.

	Número de partículas					
Grau médio	25	50	100	250	500	1000
2	75%	79%	79%	58%	58%	32%
4	69%	71%	62%	50%	45%	31%
6	59%	60%	55%	48%	46%	29%
8	57%	57%	50%	47%	44%	29%
10	54%	56%	50%	46%	44%	31%
12	43%	55%	49%	47%	45%	31%
14	28%	50%	44%	44%	44%	29%
16	28%	31%	37%	27%	17%	27%

população, sendo apenas registrada uma piora brusca para enxames com 1000 partículas. Na maioria dos casos, especialmente para a função Ackley, verifica-se que a aplicação de vizinhança complexa dinâmica traz piores resultados ao FIPS, só não sendo verificada tal deterioração para enxames com 500 ou mais partículas, o que ocorre pois, nesta situação, os resultados da dinâmica original são fracos e a adaptação topológica adicionada permite que as partículas ganhem conexões, acelerando a convergência. Contrário ao visto no CNPSO+ mas de forma similar ao FIPS, é possível notar uma grande sensibilidade do mecanismo à definição da densidade inicial do enxame, sendo verificados piores resultados para enxames mais densos. Tal situação ocorre pois, nessa dinâmica, quando todas as influências têm o mesmo peso a combinação de muitos atratores acaba por levá-la a regiões intermediárias, com resultados mais fracos. Assim, o enxame perde desempenho na fase mais inicial, de exploração, não conseguindo escolher uma região adequada para a qual convergir.

De forma geral, observam-se resultados e taxas de sucesso ruins para esta combinação, apontando piora com relação aos dois algoritmos isolados. Conforme visto na tabela 5.22, os melhores resultados foram obtidos para população de tamanho 50 e grau médio 4.

Analisando a variação dos resultados para a aplicação do CNPSO+ ao *Bare Bones Particle Swarm* (tabelas 5.24 e 5.25), nota-se que, em enxames menores, a quantidade média de conexões tem pouca influência. Com 100 partículas, observa-se a formação de dois grupos de índices de desempenho, induzida por uma certa instabilidade dessa configuração ao otimizar a função Ackley. Para populações com 250 ou mais elementos, no entanto, é possível observar resultados continuamente melhores ao aumentar a quantidade de conexões. O aparecimento desta relação para enxames menores do que ocorre com o CNPSO+ é decorrência de uma convergência significativamente mais lenta, que é fruto também da realização de apenas 50 adaptações topológicas nesta bateria de testes, frente às 80 utilizadas nos ajustes de parâmetros anteriores. É possível observar, assim como no CNPSO+, um vínculo

Tab. 5.24: Testes de tamanhos do enxame e graus médios para a versão híbrida do *Bare Bones Particle Swarm* com o CNPSO+ – Índices de desempenho (d_f).

	Número de partículas					
Grau médio	25	50	100	250	500	1000
2	0,3722	0,3453	0,2303	0,7374	0,8446	0,8834
4	0,3639	0,3638	0,2290	0,7063	0,8406	0,8829
6	0,3714	0,3586	0,3132	0,6784	0,8386	0,8819
8	0,3671	0,3521	0,2265	0,6522	0,8320	0,8799
10	0,3977	0,3660	0,3362	0,6447	0,8318	0,8752
12	0,3860	0,3570	0,3094	0,6131	0,8329	0,8766
14	0,3759	0,3671	0,3606	0,6036	0,8270	0,8741
16	0,3850	0,3608	0,3132	0,5666	0,8254	0,8704

Tab. 5.25: Testes de tamanhos do enxame e graus médios para a versão híbrida do *Bare Bones Particle Swarm* com o CNPSO+ – Taxas de sucesso.

	Número de partículas					
Grau médio	25	50	100	250	500	1000
2	90%	98%	99%	16%	13%	1%
4	93%	97%	98%	17%	12%	0%
6	89%	99%	98%	18%	12%	0%
8	89%	97%	99%	23%	14%	0%
10	86%	95%	97%	28%	14%	1%
12	86%	94%	98%	32%	13%	1%
14	88%	97%	97%	34%	14%	2%
16	83%	95%	98%	41%	14%	3%

bem marcado entre desempenho da otimização e o tamanho do enxame. Este híbrido, porém, registra uma piora intensa, em todas as funções testadas, ao aumentar a população para além do limite de 100 partículas, característica que o assemelha mais ao *Bare Bones* PSO.

Destaca-se, ainda, a obtenção de resultados muito positivos para a função Schwefel, em um dos raros casos em que um algoritmo conseguiu 100% de sucesso neste teste. A única situação em que se observa desempenho semelhante é para o *Bare Bones Particle Swarm* com vizinhança global. Com base nos resultados vistos na tabela 5.24, foi definida como topologia inicial uma rede com 100 nós e, na média, 8 conexões por vértice.

Observa-se que, assim como no CNPSO+, o desempenho de sua combinação com o AR-PSO (tabelas 5.26 e 5.27) apresenta uma dependência maior sobre o tamanho do enxame do que sobre o grau médio. Também de forma similar ao PSO com vizinhança complexa, verifica-se que, em enxames maiores, há uma melhora dos resultados ao aumentar o grau das partículas. Para populações peque-

Tab. 5.26: Testes de tamanhos do enxame e graus médios para a versão híbrida do AR-PSO com o CNPSO+ – Índices de desempenho (d_f).

	Número de partículas					
Grau médio	25	50	100	250	500	1000
2	0,4599	0,3725	0,2976	0,2624	0,3064	0,6482
4	0,4454	0,3394	0,3322	0,2662	0,2892	0,6412
6	0,4583	0,3756	0,3308	0,2656	0,2804	0,6398
8	0,4475	0,3976	0,3219	0,2988	0,2681	0,6343
10	0,4711	0,4028	0,3207	0,2818	0,2643	0,6278
12	0,4671	0,3973	0,3270	0,2753	0,2641	0,6237
14	0,4754	0,3981	0,3635	0,3032	0,2658	0,6221
16	0,4988	0,4058	0,3615	0,2719	0,2694	0,6191

Tab. 5.27: Testes de tamanhos do enxame e graus médios para a versão híbrida do AR-PSO com o CNPSO+ – Taxas de sucesso.

	Número de partículas					
Grau médio	25	50	100	250	500	1000
2	69%	81%	88%	92%	83%	40%
4	73%	81%	90%	94%	84%	42%
6	71%	84%	89%	92%	85%	41%
8	75%	79%	90%	96%	93%	40%
10	65%	78%	87%	92%	95%	43%
12	73%	84%	87%	93%	95%	44%
14	67%	79%	86%	94%	96%	48%
16	70%	76%	85%	92%	94%	48%

nas, no entanto, os algoritmos diferem de comportamento, com a versão com repulsão apresentando perda de desempenho em redes mais densas. Em enxames menores, assim como ocorre no AR-PSO (visto na tabela 5.19), é notada menor efetividade na busca local, refletida nos resultados para as funções Rosenbrock, Sphere e Ackley. Tal efeito provém do fato de enxames menores perderem diversidade mais rapidamente ao realizarem a explotação de uma região, ativando, assim, a repulsão das partículas. Enxames maiores, por outro lado, permitem que conjuntos de partículas convirjam de forma independente, mantendo, assim, a diversidade global mais alta. De forma geral, poucos são os efeitos do uso de topologia complexa dinâmica vistos neste caso, apenas podendo ser apontada uma melhoria na taxa de sucesso para a função Ackley. Os parâmetros da topologia inicial escolhidos foram 250 partículas e grau médio igual a 2.

Quantidade de adaptações topológicas

Tab. 5.28: Testes de quantidade de adaptações topológicas para as versões híbridas do CNPSO+ – Índices de desempenho (d_f).

	Adaptações						
Algoritmo	0	3	7	20	50	150	400
AR-PSO	0,4085	0,2482	0,2619	0,2707	0,2623	0,3187	0,3392
Bare Bones PSO	0,2514	0,2464	0,3119	0,3211	0,3180	0,2172	0,2260
FIPS	0,2364	0,2477	0,2845	0,3961	0,3991	0,4219	0,4334

Tab. 5.29: Testes de quantidade de adaptações topológicas para as versões híbridas do CNPSO+ – Taxas de sucesso.

	Adaptações						
Algoritmo	0	3	7	20	50	150	400
AR-PSO	69%	99%	96%	94%	93%	89%	89%
Bare Bones PSO	87%	89%	90%	96%	99%	99%	97%
FIPS	83%	84%	79%	75%	73%	63%	56%

Nas tabelas 5.28 e 5.29, observa-se, para o FIPS, resultados mais positivos quanto menor a quantidade de modificações topológicas realizadas, o que indica que a idéia de buscar novas influências quando uma partícula encontra-se estagnada não é benéfica a esta dinâmica. Isto ocorre pois o atrator da partícula localiza-se em um ponto obtido pela combinação de todas as influências que esta recebe e não em uma localização positiva previamente visitada. Deste modo, ao receber conexões que apresentem informações acerca de uma diferente porção do espaço, a partícula ficará perdida, atraída pelo ponto médio das localizações promissoras, o que reduz, assim, sua capacidade de busca local. Tal deterioração é mais intensa quando há múltiplos ótimos locais, o que explica que os resultados sejam tão negativos em todos os testes, exceto para as funções Rosenbrock e Sphere, menos multimodais. Deste modo, optou-se pela supressão de quaisquer atualizações topológicas na versão híbrida entre o FIPS e o CNPSO+.

De forma contrária à que ocorre com o FIPS, para o híbrido com o *Bare Bones Particle Swarm* é visto um claro benefício ao se aumentar a quantidade de adaptações topológicas, somente não sendo obtidos resultados melhores para a função Griewank. Tal desempenho ocorre devido ao aumento da regularidade da otimização, o que é refletido nas taxas de sucesso médias e por função. Através da análise dos índices de desempenho, foi fixada em 150 a quantidade adaptações para a combinação entre o CNPSO+ e o *Bare Bones Particle Swarm*.

A relação dos resultados do AR-PSO híbrido com o número de adaptações realizadas na rede de influências tende a ser semelhante à do CNPSO+. A exceção ocorre com as funções Ackley e Sphere, para as quais é benéfica a realização de menos modificações topológicas. Isto ocorre pois, como os melhores resultados para este algoritmo foram obtidos para redes esparsas, aumentar a quantidade de adaptações significa tornar a rede mais densa, facilitando o fluxo de informações e acelerando a convergência. Conforme dito anteriormente, uma convergência mais rápida do enxame faz com que este entre em fase de repulsão mais regularmente, deteriorando, assim, sua habilidade na execução de busca local. Deste modo, foi verificada que a realização de apenas 3 modificações durante toda a execução do algoritmo é capaz de prover o melhor resultado geral.

5.4 Comparação de resultados após ajustes de parâmetros

Tendo ajustado os parâmetros da topologia utilizada, foram realizados os testes finais para a comparação de algoritmos e vizinhanças. Para aumentar a isenção das análises, nesta parte do trabalho, além do índice de desempenho aqui apresentado, também foi utilizado o índice proposto por Mendes *et al.* [127], explicado na seção 5.2.3. Os resultados obtidos por cada algoritmo, discriminados por função, são exibidos nas tabelas 5.30, 5.31, 5.32 e 5.33 e no apêndice C, onde são exibidos os gráficos de barras com valores médios e desvios-padrão obtidos por cada algoritmo para cada função, enquanto que o ranqueamento dos algoritmos, segundo seus desempenhos médios, estão indicados nas tabelas 5.34, 5.35, 5.36 e 5.37. Uma segunda bateria de testes foi realizada para a elaboração de gráficos de convergência que permitam acompanhar, na média, qual o melhor resultado já obtido por um algoritmo até a avaliação de função atual. Estes gráficos podem ser vistos no apêndice B.

Ao analisar apenas abordagens em que a população é mantida com um número fixo de partículas, é possível reparar que a principal diferença na forma do gráfico de convergência foi relativa à velocidade de busca, com topologias mais esparsas apresentando uma convergência mais lenta e obtendo, assim, melhores resultados devido à menor tendência de ocorrer uma estagnação prematura. Tanto para a dinâmica canônica quanto para aquela com repulsão, observam-se dois perfis típicos de evolução dos resultados da busca, separados por função: na maior parte dos casos, em uma primeira fase são realizadas melhorias intensas, ocorrendo uma transição súbita para uma etapa de progressos mais lentos até que seja atingida, por fim, uma situação de estabilização, com a ocorrência de poucas melhorias pontuais. Para as funções Ackley e Sphere, cujos resultados finais são mais dependentes de busca local, é visto o segundo perfil ao utilizar vizinhanças com menos conexões, marcado por uma melhoria mais lenta a princípio, seguida por uma sequência de resultados com maior qualidade, que se estende até o limite de avaliações de funções. Topologias globais, porém, mostraram-se mais propensas à estagnação nestas funções, possivelmente devido a uma perda prematura de diversidade, que

faz com as partículas fiquem presas nas adjacências de regiões ótimas, sem informações suficientes para detectar outras regiões promissoras.

Para o *Bare Bones Particle Swarm*, observam-se algumas mudanças de comportamento: a evolução dos resultados torna-se mais lenta, há uma transição mais suave entre a fase de melhorias mais intensas e a estagnação e, ao otimizar as funções Schwefel e Rosenbrock, nota-se que há um aumento na etapa de exploração, particularmente benéfica para a primeira função citada, sendo seus melhores resultados conseguidos nesta dinâmica. O FIPS, por sua vez, apresenta uma convergência prematura, de modo que o enxame encontra-se estagnado durante grande parte da busca. Conseguindo alguns resultados melhores que aqueles obtidos pelo PSO canônico apenas para a função Schwefel, o AR-PSO não consegue os efeitos positivos que promete, especialmente se usada sua topologia original (global). Os valores ótimos obtidos são, em geral, similares àqueles da versão simples, mas, conforme dito anteriormente, há uma perda de capacidade de busca local, refletida em uma parada brusca na melhoria de resultados para as funções Sphere e Ackley, mesmo ao utilizar topologias esparsas.

No PSO canônico, as vizinhanças estáticas capazes de fornecer os melhores resultados foram as topologias complexa e von Neumann, este último, porém, exibe valores pouco satisfatórios ao ser utilizado com o *Bare Bones Particle Swarm*. Observam-se, também, algumas situações comuns a todas as dinâmicas: ao utilizar a topologia global, é recorrente observar o enxame ficar preso em um mínimo local na função Ackley, assim como os resultados e perfis de convergência mostram-se muito similares para vizinhanças locais e von Neumann, com a última normalmente tendo um índice de desempenho um pouco melhor.

Mesmo sendo um método bem simples de modificação topológica, o DMS-PSO obtém, frequentemente, resultados bons, tendo seus índices de desempenho superados apenas pelo TRIBES. A evolução dos resultados é similar à que ocorre com vizinhanças von Neumann, ocorrendo, porém, um prolongamento da etapa de exploração das funções Rastrigin e Schwefel, que são marcadas, em quase todos os algoritmos testados, por melhorias mais modestas durante a otimização do que outras funções, onde as soluções finais são ordens de grandeza menores do que aquelas obtidas no início da busca. Devido a sua dinâmica, o DMS-PSO traz uma difusão mais lenta das informações na rede, atraso esse proporcional ao período de reconstrução da topologia. Um benefício oriundo deste fluxo é o fato de que, antes de ser amplamente difundida por toda a população, uma informação tende a ser refinada pelos sub-enxames, tornando-a mais competitiva, o que permite manter a diversidade em níveis mais altos. Esta característica, porém, faz com que, atingido o critério de parada, ainda não tenham sido obtidos bons resultados para a função Schwefel. Por fim, nota-se, de forma bem clara, a modificação de comportamento que ocorre quando, após passado 90% do limite de iterações, o enxame passa a utilizar uma vizinhança global. Nesta situação, ocorre uma breve aceleração da convergência, seguida por uma tendência à estagnação, devida à intensa explotação realizada.

Assim como para o DMS-PSO, o progresso da otimização do H-PSO é similar ao visto para o PSO com vizinhança von Neumann, sendo um pouco mais lento e fornecendo resultados de mais baixa qualidade em alguns casos. Tal deterioração é devida à estrutura em árvore utilizada por esse algoritmo, que faz com que quaisquer trocas de informações entre dois ramos seja mediada pela promoção de um destes nós a uma posição superior. Particularmente, há uma predominância muito grande da influência do nó raiz sobre o funcionamento do enxame como um todo, de modo que são verificados resultados melhores que aqueles retornados ao utilizar uma vizinhança von Neumann ou o CNPSO+ apenas em situações que necessitam de uma convergência mais rápida (funções Ackley e Sphere).

Dentre os algoritmos com vizinhança complexa, nota-se que a versão sem adaptações topológicas e o CNPSO-1 possuem desempenhos parecidos, diferenciando-se do CNPSO+, o que indica que os resultados alcançados pelo PSO dependem mais da densidade de conexões do que das posições exploradas pelos *hubs*. O CNPSO-1, no entanto, conseguiu a maior taxa de sucesso dentre todos os mecanismos testados, sendo o único a alcançar o objetivo estipulado em todas as execuções realizadas. Tal situação aponta que, apesar de ser uma condição normalmente superável, um mau posicionamento dos nós mais conectados pode fazer com que, eventualmente, o enxame fique preso em um mínimo local. Quanto à velocidade de convergência, verifica-se que a segunda proposta de atualização topológica torna a busca mais rápida, exceto para a função Schwefel, mas sem levar a uma estagnação prematura, o que comprometeria o resultado geral.

Conforme esperado, os CNPSOs mostraram uma alta capacidade de otimização para a função Schwefel, a mais multimodal do conjunto de testes. Dos oito algoritmos com melhores resultados para esta função, cinco se valeram de redes complexas. A única exceção a estes resultados positivos ocorre para a versão híbrida com o FIPS, já discutida anteriormente. O melhor resultado geral para esta função foi obtido pela combinação do CNPSO+ e do *Bare Bones Particle Swarm*, situação na qual ocorre um comportamento diferenciado desta vizinhança. Apesar de, quando aplicado a outros mecanismos, o CNPSO+ ter um perfil que se assemelha ao das outras vizinhanças esparsas, como a von Neumann, no caso em questão, são apresentados resultados e convergência mais parecidos com os obtidos ao utilizar vizinhança global. Isto ocorre por esta dinâmica ter sido inicialmente projetada para reproduzir o comportamento do PSO canônico, porém, sem o uso de parâmetros, gerando as posições das partículas por uma distribuição em forma de sino (no caso uma gaussiana) sem o estabelecimento de qualquer limitante para a máxima diferença entre dois pontos visitados por uma partícula em iterações subsequentes – o que seria equivalente à restrição de velocidades no PSO original. Assim, na fase inicial, quando as partículas estão mais espalhadas e, portanto, as velocidades tendem a ser maiores com menos possibilidades para o refino de resultados, há maior probabilidade dos indivíduos serem considerados em estagnação, ganhando a rede, rapidamente, maior densidade

de conexões.

Em cada uma das dinâmicas analisadas, nota-se que a aplicação do CNPSO+ traz resultados positivos frente às topologias clássicas, obtendo os melhores desempenhos – caso da versão canônica e do AR-PSO – ou bem próximo dos melhores – para o *Bare Bones Particle Swarm* e para o FIPS, nos quais a vizinhança complexa fica em segundo lugar quando analisado pelo índice de desempenhos aqui proposto e em primeiro quando utilizando o índice de Mendes *et al.* [127].

Ao comparar os resultados do TRIBES, melhores em quase todas as funções de teste, saltam aos olhos os benefícios obtidos ao tornar o tamanho da população variável. Esta flexibilização permite selecionar o momento em que deve haver uma busca mais ampla, contando com mais partículas, ou que deve ser focada a exploração de uma região, dispensando indivíduos que não seriam capazes de oferecer informações novas. Outro aspecto positivo é a inserção de novos elementos com posição aleatória e em uma tribo à parte: isto permite manter maior diversidade, uma vez que tais informações podem ser melhor apuradas, evitando que estes novos dados, potencialmente benéficos, sejam asfixiados por um conjunto de vizinhos com posições não necessariamente melhores, mas que já tenham sido refinadas. A evolução da busca no TRIBES revela alguns aspectos interessantes, como a ocorrência de melhoras súbitas, frutos da descoberta de regiões promissoras pelas partículas adicionadas ao sistema. Esta diferença mais acentuada no perfil de convergência indica que a exploração de posições aleatórias no decorrer do processo e o controle do tamanho do enxame são modificações mais impactantes do que apenas alterar a forma de conexão entre as partículas.

5.5 Conclusão do capítulo

De forma geral, os resultados obtidos pelos modelos baseados em vizinhança complexa são muito próximos daqueles vistos para o DMS-PSO e para topologias von Neumann. Notam-se maiores diferenças apenas na otimização das funções Sphere e Ackley, onde são verificados perfis de busca similares, tendo, porém, os CNPSOs uma menor velocidade de convergência, advinda de sua preferência por enxames maiores. Deste modo, verifica-se que a introdução de algum mecanismo para acelerar a otimização, como a eliminação de algumas partículas na etapa final, é potencialmente benéfica. Por outro lado, este melhor desempenho em grandes populações, quando comparado às outras topologias, aponta uma importante consequência da combinação do efeito *small-world* à flexibilidade de escolha da densidade do enxame, que é um crescimento menor, conforme aumentado o enxame, da quantidade de avaliações de função necessárias para atingir o objetivo estabelecido.

Tendo em vista a semelhança dos resultados obtidos pelo CNPSO com aqueles retornados ao usar vizinhança von Neumann, à primeira vista, não haveria vantagem em aplicar redes complexas na otimização por enxame de partículas, dado o aumento do esforço de implementação e do custo

computacional, frente a uma vizinhança estática já bem estabelecida na literatura. No entanto, é necessário ressaltar que, em todas as dinâmicas onde foi aplicada, a topologia complexa foi capaz de conseguir resultados competitivos, inclusive na versão híbrida com o *Bare Bones Particle Swarm*, situação na qual a topologia von Neumann mostra resultados mais fracos. Aponta-se, também, que tal desempenho positivo foi obtido ao utilizar valores de parâmetros comuns na literatura, os quais foram ajustados para as topologias mais clássicas, empiricamente ou em análises teóricas de versões simplificadas do PSO. Assim, ao aplicar outros modelos de vizinhança, faz-se necessária uma análise de quais parâmetros seriam os mais recomendados. Para o próprio CNPSO-1, por exemplo, uma configuração capaz de prover melhores resultados no conjunto de teste utilizado se valhe de fatores de inércia, sociais e congnitivos diferentes dos aqui utilizados, como pode ser visto em [78]. No entanto, a própria seleção dos parâmetros de uma meta-heurística é, por si só, um problema de otimização combinatória de alta complexidade computacional, de modo que mecanismos livres de parâmetros (como o TRIBES) e meta-otimizações capazes de encontrar automaticamente boas combinações tem sido objetivo de vários artigos [38, 99]. Assim, dado o foco deste trabalho na comparação de vizinhanças e o grande número de algoritmos analisados, optou-se por valores mais consagrados para alguns parâmetros, com a adequação apenas daqueles relativos à topologia, deixando um ajuste fino mais preciso para trabalhos posteriores.

Potenciais benefícios da aplicação de redes complexas à otimização por enxame de partículas não se restringem apenas a suas qualidades intrínsecas, mas incluem o ferramental que elas disponibilizam: há uma ampla gama de mecanismos para a geração de redes heterogêneas, com um comportamento relativamente bem mapeado. Ao contrário da topologia von Neumann, por exemplo, em redes complexas é possível adicionar e remover nós e arestas, ajustando a rede de influência para apresentar diversas características, como o envelhecimento de partículas ou *fitness* diferenciados. Muitos destes mecanismos são baseados em regras locais, o que, em conjunto com os bons resultados obtidos em enxames maiores, é uma característica desejável para a paralelização da busca.

Por fim, a partir dos resultados obtidos pelo TRIBES, verifica-se que uma direção a ser tomada em próximas etapas é o uso de enxames de tamanho variável, com a inclusão de mecanismos de refino de novas soluções, algo que se mostrou benéfico também ao DMS-PSO.

Tab. 5.30: Comparação de desempenho dos algoritmos – Resultados médios das funções objetivo. Ressalta-se que, quanto mais intenso o vermelho no fundo de uma célula, melhor o resultado nela contido.

		Funções						
Algoritmo	Vizinhança	Salomon	Rastrigin	Schwefel	Rosenbrock	Sphere	Griewank	Ackley
CNPSO Rígido		2,3971E-01	3,5874E+01	4,3280E+03	2,9686E+01	2,4559E-07	1,7290E-03	1,7520E-03
CNPSO-1		2,1987E-01	3,6881E+01	4,2990E+03	3,7544E+01	8,7234E-09	4,0184E-03	2,6565E-04
CNPSO+		2,0654E-01	3,6183E+01	4,6530E+03	2,6365E+01	2,6899E-09	5,7511E-04	1,0309E-04
PSO	Global	3,7321E-01	5,9167E+01	5,1741E+03	3,8073E+01	4,6710E-09	1,2230E-02	2,3810E+01
	Local	3,1321E-01	4,9537E+01	4,7355E+03	2,3112E+01	1,5940E-21	3,3255E-04	8,5580E-12
	von Neumann	2,0321E-01	3,4031E+01	4,5169E+03	2,9163E+01	3,4345E-17	2,2257E-03	6,1249E-09
FIPS	von Neumann	2,2048E-01	4,5022E+01	1,0033E+04	2,9139E+01	2,0622E-10	1,2180E-03	4,7321E-04
	CNPSO+	3,8057E-01	5,1938E+01	7,3725E+03	2,7284E+01	4,4241E-06	1,3061E-03	4,0748E-02
Bare Bones PSO	Global	3,2324E-01	3,6847E+01	3,6851E+03	4,6212E+01	5,4906E-13	8,5982E-03	1,6385E+00
	Local	5,8812E-01	6,4910E+01	5,5303E+03	5,7235E+01	3,2384E-10	1,8885E-03	5,9531E-07
	von Neumann	6,5134E-01	4,8145E+01	5,0216E+03	8,5651E+01	5,8945E-06	2,6502E-03	1,7210E-01
	CNPSO+	3,3656E-01	3,6617E+01	3,5435E+03	3,9360E+01	8,8026E-14	4,8431E-03	1,7200E-01
H-PSO		2,2987E-01	3,4598E+01	5,7006E+03	3,1373E+01	1,7459E-20	4,3735E-03	1,4336E-10
DMS-PSO		1,8987E-01	2,3257E+01	4,8279E+03	2,6880E+01	2,1985E-15	2,4653E-04	8,1425E-08
AR-PSO	Global	3,3321E-01	6,0064E+01	5,3494E+03	5,0470E+01	1,8711E-03	1,4245E-02	2,1360E+01
	von Neumann	2,8654E-01	3,4502E+01	4,2683E+03	2,5655E+01	7,0887E-04	4,5834E-04	8,8104E-02
	CNPSO+	2,1769E-01	5,0868E+01	4,5909E+03	3,2867E+01	4,3214E-05	1,1021E-03	4,2335E-02
TRIBES		6,3879E-02	2,3041E-08	4,8079E+03	2,0047E+01	9,3006E-50	0,0000E+00	0,0000E+00
Valor do ótimo global		0,0000E+00	0,0000E+00	0,0000E+00	0,0000E+00	0,0000E+00	0,0000E+00	0,0000E+00

Tab. 5.31: Comparação final de desempenho dos algoritmos – Desvios-padrão.

		Funções						
Algoritmo	Vizinhança	Salomon	Rastrigin	Schwefel	Rosenbrock	Sphere	Griewank	Ackley
CNPSO Rígido		4,7211E-02	6,2735E+00	4,8802E+02	9,9010E-08	9,9010E-08	5,0544E-03	5,9071E-04
CNPSO-1		3,9999E-02	6,9925E+00	3,9485E+02	6,7326E-09	6,7326E-09	8,0363E-03	8,2262E-05
CNPSO+		2,4944E-02	6,9953E+00	4,8102E+02	1,6409E-09	1,6409E-09	2,1752E-03	3,5160E-05
PSO	Global	6,2893E-02	1,1997E+01	7,3051E+02	1,4939E-08	1,4939E-08	1,1846E-02	2,5821E+01
	Local	4,2687E-02	8,1049E+00	5,1480E+02	1,4985E-21	1,4985E-21	1,7908E-03	3,9759E-12
	von Neumann	3,1447E-02	5,2645E+00	4,2896E+02	2,6712E-17	2,6712E-17	3,7450E-03	2,4851E-09
FIPS	von Neumann	3,9807E-02	6,3577E+00	4,7590E+02	4,6364E-10	4,6364E-10	4,2329E-03	1,4585E-03
	CNPSO+	8,2675E-02	7,5545E+00	7,6423E+02	2,2693E-05	2,2693E-05	3,8152E-03	2,0764E-01
Bare Bones PSO	Global	4,2291E-02	9,6584E+00	1,0404E+03	1,1218E-12	1,1218E-12	1,4012E-02	7,0257E+00
	Local	5,5071E-02	1,3661E+01	9,4534E+02	1,1517E-09	1,1517E-09	4,4820E-03	5,9663E-07
	von Neumann	9,2552E-02	1,0383E+01	6,3287E+02	5,8106E-06	5,8106E-06	6,3450E-03	9,2624E-01
	CNPSO+	5,4631E-02	6,9057E+00	5,8943E+02	1,1349E-13	1,1349E-13	6,7675E-03	9,2622E-01
H-PSO		4,5826E-02	8,1968E+00	8,0717E+02	1,6419E-20	1,6419E-20	6,7867E-03	8,0367E-11
DMS-PSO		3,0000E-02	5,1612E+00	6,8577E+02	2,2554E-15	2,2554E-15	1,3276E-03	4,4838E-08
AR-PSO	Global	6,9921E-02	1,2892E+01	7,1144E+02	3,5283E+01	1,2838E-03	1,1616E-02	2,7047E+01
	von Neumann	3,3990E-02	6,3337E+00	4,0104E+02	2,1953E+00	2,6844E-04	1,3116E-03	1,3882E-02
	CNPSO+	1,7951E-02	1,0884E+01	6,4278E+02	1,4466E+00	8,9104E-10	3,6224E-03	2,8881E-05
TRIBES		9,7855E-02	1,2392E-07	5,9748E+02	4,9568E-49	4,9568E-49	0,0000E+00	3,6616E-15

Tab. 5.32: Comparação final de desempenho dos algoritmos – Taxas de sucesso.

		Funções						
Algoritmo	Vizinhança	Salomon	Rastrigin	Schwefel	Rosenbrock	Sphere	Griewank	Ackley
CNPSO Rígido		100%	100%	90%	97%	100%	100%	100%
CNPSO-1		100%	100%	100%	100%	100%	100%	100%
CNPSO+		100%	100%	80%	100%	100%	100%	100%
PSO	Global	90%	100%	43%	97%	100%	100%	20%
	Local	100%	100%	67%	100%	100%	100%	100%
	von Neumann	100%	100%	90%	100%	100%	100%	100%
FIPS	von Neumann	100%	100%	0%	100%	100%	100%	100%
	CNPSO+	77%	100%	0%	100%	100%	100%	97%
Bare Bones PSO	Global	97%	100%	93%	97%	100%	100%	90%
	Local	0%	100%	37%	93%	100%	100%	100%
	von Neumann	0%	100%	57%	80%	100%	100%	97%
	CNPSO+	100%	100%	97%	100%	100%	100%	97%
H-PSO		100%	100%	20%	100%	100%	100%	100%
DMS-PSO		100%	100%	50%	100%	100%	100%	100%
AR-PSO	Global	93%	100%	33%	93%	100%	100%	0%
	von Neumann	100%	100%	100%	100%	100%	100%	87%
	CNPSO+	100%	100%	80%	100%	100%	100%	90%
TRIBES		100%	100%	30%	100%	100%	100%	100%

Tab. 5.33: Comparação final de desempenho dos algoritmos – Avaliações de função esperadas até atingir o objetivo (AF_{esp}).

		Funções						
Algoritmo	Vizinhança	Salomon	Rastrigin	Schwefel	Rosenbrock	Sphere	Griewank	Ackley
CNPSO Rígido		74258,3	34016,7	40761,3	61533,9	73483,3	67933,3	82291,7
CNPSO-1		65866,7	32966,7	37458,3	57916,7	69000,0	64116,7	76908,3
CNPSO+		62550,0	37691,7	53085,9	54366,7	69925,0	64791,7	75858,3
PSO	Global	59197,5	25808,3	58934,9	46935,2	51700,0	48658,3	302291,7
	Local	60390,0	22168,3	39450,0	42945,0	47060,0	43821,7	52525,0
	von Neumann	57273,3	25370,0	35127,6	45376,7	54126,7	50553,3	59776,7
FIPS	von Neumann	13076,7	14641,7	–	10670,0	16062,5	14825,0	19294,2
	CNPSO+	33904,5	17110,0	–	13183,3	22145,0	20280,0	27240,8
Bare Bones PSO	Global	60381,7	16283,3	49798,5	44810,9	42460,0	36223,3	37987,7
	Local	–	22660,0	209938,0	76565,1	47866,7	41725,0	50553,3
	von Neumann	–	23606,7	146387,5	101390,6	69463,3	58483,3	64362,7
	CNPSO+	66166,7	18200,0	55298,5	51640,0	46693,3	41333,3	43983,4
H-PSO		53253,3	42500,0	168666,7	45253,3	49540,0	46510,0	55770,0
DMS-PSO		58119,6	46886,4	134151,6	45068,1	52816,5	50463,6	59664,0
AR-PSO	Global	57149,2	26733,3	87225,0	49142,2	52441,7	47825,0	–
	von Neumann	81416,7	36858,3	42450,0	63066,7	80333,3	74233,3	103369,1
	CNPSO+	79491,7	42083,3	58671,9	68091,7	82366,7	76458,3	100709,9
TRIBES		15519,8	2826,1	60741,1	9165,7	19399,3	13858,0	23267,5

Tab. 5.34: Classificação dos algoritmos segundo o índice de desempenho proposto neste trabalho.

Algoritmo	Vizinhança	Desempenho
TRIBES		0,1136
DMS-PSO		0,1737
CNPSO+		0,1862
PSO	von Neumann	0,1876
CNPSO Rígido		0,1951
CNPSO-1		0,2001
H-PSO		0,2051
PSO	Local	0,2068
FIPS	von Neumann	0,2251
AR-PSO	CNPSO+	0,2469
AR-PSO	von Neumann	0,2682
Bare Bones PSO	Local	0,2709
FIPS	CNPSO+	0,2775
Bare Bones PSO	CNPSO+	0,3001
Bare Bones PSO	Global	0,3540
Bare Bones PSO	von Neumann	0,3665
PSO	Global	0,3919
AR-PSO	Global	0,4228

Tab. 5.35: Classificação dos algoritmos segundo seus índices de desempenho (Mendes *et al.* [127]).

Algoritmo	Vizinhança	Desempenho
TRIBES		-1,0536
DMS-PSO		-0,6200
CNPSO+		-0,4869
PSO	von Neumann	-0,4404
CNPSO Rígido		-0,4154
PSO	Local	-0,2758
CNPSO-1		-0,2750
AR-PSO	CNPSO+	-0,2439
AR-PSO	von Neumann	-0,2409
H-PSO		-0,1934
Bare Bones PSO	CNPSO+	-0,1806
Bare Bones PSO	Global	0,0480
FIPS	CNPSO+	0,1546
FIPS	von Neumann	0,1948
Bare Bones PSO	Local	0,6150
Bare Bones PSO	von Neumann	0,7655
PSO	Global	0,9650
AR-PSO	Global	1,6830

Tab. 5.36: Classificação dos algoritmos segundo suas taxas de sucesso médias.

Algoritmo	Vizinhança	Sucesso
CNPSO-1		100%
Bare Bones PSO	CNPSO+	99%
PSO	von Neumann	99%
AR-PSO	von Neumann	98%
CNPSO Rígido		98%
CNPSO+		97%
Bare Bones PSO	Global	97%
AR-PSO	CNPSO+	96%
PSO	Local	95%
DMS-PSO		93%
TRIBES		90%
H-PSO		89%
FIPS	von Neumann	86%
FIPS	CNPSO+	82%
PSO	Global	79%
Bare Bones PSO	von Neumann	76%
Bare Bones PSO	Local	76%
AR-PSO	Global	74%

Tab. 5.37: Classificação dos algoritmos segundo a quantidade de avaliações de função esperadas até atingir o objetivo (AF_{esp}). Aponta-se que os espaços sinalizados com "–" indicam que o algoritmo testado não foi capaz de obter ao menos uma execução bem sucedida em alguma das funções de teste.

Algoritmo	Vizinhança	Avaliações
TRIBES		20682,5
Bare Bones PSO	Global	41135,1
PSO	Local	44051,4
Bare Bones PSO	CNPSO+	46187,9
PSO	von Neumann	46800,6
CNPSO-1		57747,6
CNPSO+		59752,8
CNPSO Rígido		62039,8
DMS-PSO		63881,4
H-PSO		65927,6
AR-PSO	von Neumann	68818,2
AR-PSO	CNPSO+	72553,3
PSO	Global	84789,4
FIPS	von Neumann	–
FIPS	CNPSO+	–
Bare Bones PSO	Local	–
Bare Bones PSO	von Neumann	–
AR-PSO	Global	–

Capítulo 6

Conclusões e Perspectivas Futuras

Neste trabalho, foram analisados os impactos do uso de topologias complexas para a definição da rede de influências na otimização por enxame de partículas. Com este objetivo, foram aqui propostas variantes da técnica em questão, nas quais a vizinhança é estabelecida de forma dinâmica, porém com diferentes graus de flexibilidade.

O PSO é uma técnica extremamente distribuída, na qual a idéia de auto-organização é um princípio norteador. Ao contrário do que ocorre em outras meta-heurísticas, aqui não se verifica a presença de quaisquer mecanismos centralizados, como a seleção natural nos algoritmos genéticos, de modo que a coordenação do enxame e sua capacidade de chegar a bons resultados são frutos apenas de regras simples internas às partículas, aplicadas a informações obtidas pela intensa comunicação de uma partícula com suas vizinhas. Logo, a topologia que define essas interações no enxame é determinante de seu comportamento, de modo que se encontram na literatura diversas alternativas de vizinhança, conforme apresentado neste trabalho.

Apesar de ser um tópico relativamente recente, muitos estudos já foram realizados sobre vários aspectos de redes complexas. Assim sendo, este trabalho focou na revisão bibliográfica sobre as principais características destas topologias e, especialmente, nos métodos propostos para sua evolução, esforço que serviu de base para o projeto e análise do CNPSO.

Em uma primeira etapa da pesquisa, foi proposta uma modificação ao PSO, com o uso de uma vizinhança complexa adaptável, na qual tanto a quantidade de partículas quanto a de conexões entre elas era mantida constante durante a busca, focando a evolução da rede apenas na distribuição de graus entre os nós. Comparado a versões do algoritmo em que a topologia foi mantida fixa, foram verificadas poucas mudanças, de forma que a única melhoria notável foi uma taxa de sucesso levemente maior. Uma semelhança observada entre essas topologias, estática e dinâmica, foi a alta sensibilidade à quantidade de arestas na rede de influências do enxame. Assim, em uma segunda etapa, procurou-se tornar a vizinhança ainda mais flexível, permitindo o aumento e a redução da den-

sidade de conexões, conforme fosse verificada a estagnação de algumas partículas durante o processo de otimização. Nessa versão foi verificado um sensível ganho de desempenho e uma redução na dependência sobre a vizinhança inicial.

Comparando estas variantes com outras propostas vistas na literatura, é possível notar duas características principais dos CNPSOs, a saber, melhores resultados em problemas mais multimodais e preferência por enxames maiores, os quais tornam mais proeminentes as características da vizinhança complexa adotada. Observa-se, também, que mesmo a versão rígida é capaz de prover bons resultados, similares àqueles obtidos por topologias altamente difundidas por sua eficiência, como a von Neumann. Um outro ponto a ser mencionado é o fato de o CNPSO ser capaz de obter resultados competitivos, mesmo ao ser aplicado a outras variantes de otimização por enxame de partículas.

Apesar de apresentarem resultados semelhantes, verifica-se que o uso de uma vizinhança von Neumann é significativamente mais fácil que o de uma vizinhança complexa, como o CNPSO+, em especial devido à quantidade de parâmetros a serem ajustados. Por outro lado, a implementação de mecanismos de ajuste automático da topologia do enxame tender a ser especialmente benéfica ao CNPSO, uma vez que esta abundância de parâmetros indica uma maior adaptabilidade dos mecanismos aqui propostos a problemas específicos.

A análise dos perfis de convergência do CNPSO indicou ser ele capaz de identificar boas regiões também em funções mais simples, apresentando como deficiência, porém, uma maior lentidão no refino de soluções. Deste modo, verifica-se que a técnica aqui proposta se beneficiaria com a adição de métodos para acelerar o processo de busca local, como a eliminação de partículas que não tragam informações novas. Através da comparação de resultados obtidos por outras dinâmicas, verifica-se, também, uma potencial melhoria ao fortalecer a organização em sub-enxames, como visto no DMS-PSO e no TRIBES. No entanto, é preciso notar que os procedimentos atuais para a geração de redes complexas ainda carecem de meios para contemplar, de maneira mais adequada, a formação de comunidades, necessitando esta área de mais pesquisas.

Algumas perspectivas futuras de trabalho na área contemplam a busca de valores mais adequados ao CNPSO para os fatores de inércia e de influência, assim como a utilização de uma população com tamanho definido dinamicamente com o decorrer da busca. Também deve ser explorado o uso de mecanismos de adaptação topológica específicos ao utilizar diferentes dinâmicas do PSO, o que evitaria situações em que a modificação da vizinhança acarrete uma perda de desempenho, como visto no FIPS. Por fim, o estudo da aplicação de vizinhanças complexas também deve ser feito para problemas de otimização dinâmica e multi-objetivo, nos quais é necessário promover uma maior diversidade do enxame.

Bibliografia

[1] A. M. Abdelbar e S. Abdelshahid. Instinct-based pso with local search applied to satisfiability. In *Proceedings of 2004 IEEE International Joint Conference on Neural Networks*, volume 3, pg. 2291–2295, Julho 2004.

[2] J. Abello, P. Pardalos, e M. G. C. Resende. On maximum clique problems in very large graphs. In *External memory algorithms*, pg. 119–130. American Mathematical Society, 1999.

[3] L. A. Adamic, B. A. Huberman, A.-L. Barabási;, R. Albert, H. Jeong, e G. Bianconi;. Power-law distribution of the world wide web. *Science*, 287(5461):2115a–, 2000.

[4] R. Albert e A.-L. Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47, 2002.

[5] L. A. N. Amaral e J. M. Ottino. Complex networks. *Eur. Phys. J. B*, 38(2):147–162, Março 2004.

[6] L. A. N. Amaral, A. Scala, M. Barthélémy, e H. E. Stanley. Classes of small-world networks. *Proceedings of the National Academy of Sciences of the United States of America*, 97(21):11149–11152, Outubro 2000.

[7] W. R. Ashby. Principles of the self-organizing dynamic system. *Journal of General Psychology*, 37:125–128, 1947.

[8] R. Axelrod. The dissemination of culture: A model with local convergence and global polarization. *Journal of Conflict Resolution*, 41(2):203–226, 1997.

[9] D. Balcan, A. Kabakçioğlu, M. Mungan, e A. Erzan. The information coded in the yeast response elements accounts for most of the topological properties of its transcriptional regulation network. *PLoS ONE*, 2(6):e501, 2007.

[10] J. Bang-Jensen e G. Gutin. *Digraphs: Theory, Algorithms and Applications*. Springer, Agosto 2002.

[11] A. Banks, J. Vincent, e C. Anyakoha. A review of particle swarm optimization. part ii: Hybridisation, combinatorial, multicriteria and constrained optimization, and indicative applications. *Natural Computing*, 7(1):109–124, Março 2008.

[12] A.-L. Barabási. *Linked: How Everything Is Connected to Everything Else and What It Means for Business, Science, and Everyday Life*. Plume Books, Abril 2003.

[13] A.-L. Barabási e R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, Outubro 1999.

[14] A.-L. Barabási, H. Jeong, Z. Néda, E. Ravasz, A. Schubert, e T. Vicsek. Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and its Applications*, 311(3):590–614, 2002.

[15] A. Barrat e M. Weigt. On the properties of small-world network models. *Eur. Phys. J. B*, 13(3):547–560, 2000.

[16] R. Beckers, J. L. Deneubourg, e S. Goss. Modulation of trail laying in the ant lasius niger (hymenoptera: Formicidae) and its role in the collective selection of a food source. *Journal of Insect Behavior*, 6(6):751–759, Novembro 1993.

[17] R. Beckers, J. L. Deneubourg, S. Goss, e J. M. Pasteels. Collective decision making through food recruitment. *Insectes Sociaux*, 37(3):258–267, 1990.

[18] M. Beekman e F. L. W. Ratniek. Long-range foraging by the honey-bee, apis mellifera l. *Functional Ecology*, pg. 490–496, 2000.

[19] M. Beekman, G. A. Sword, e S. J. Simpson. Biological foundations of swarm intelligence. In *Swarm Intelligence*, pg. 3–41. Springer, 2008.

[20] G. Bianconi e A.-L. Barabási. Competition and multiscaling in evolving networks. *Europhys. Lett.*, 54(4):436–442, Maio 2001.

[21] T. Blackwell e P. Bentley. Improvised music with swarms. In *Proceedings of the 2002 Congress on Evolutionary Computation, 2002. CEC '02*, volume 2, pg. 1462–1467, 2002.

[22] T. Blackwell e J. Branke. Multi-swarm optimization in dynamic environments. *Lecture Notes in Computer Science*, 3005:489–500, 2004.

[23] C. Blum e D. Merkle. *Swarm intelligence: Introduction and applications*. Springer, 2008.

[24] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, e D.-U. Hwang. Complex networks: Structure and dynamics. *Physics Reports*, 424(4-5):175–308, Fevereiro 2006.

[25] B. Bollobás e O. Riordan. The diameter of a scale-free random graph. *Combinatorica*, 24(1):5–34, 2004.

[26] E. Bonabeau, M. Dorigo, e G. Theraulaz. *Swarm Intelligence: From natural to artificial systems (Santa Fe Institute Studies in the Sciences of Complexity Proceedings)*. Oxford University Press, USA, 1 edição, Setembro 1999.

[27] A. Brabazon, A. Silva, T. F. Sousa, M. O'Neill, R. Matthews, e E. Costa. Investigating strategic inertia using orgswarm. *Informatica*, 29(2):125–141, 2005.

[28] C. M. Brooks e K. Koizumi. The hypothalamus and control of integrative processes. In *Medical physiology*, pg. 813–836. C. V. Mosby, Saint Louis, 1974.

[29] J. L. Burton e N. R. Franks. The foraging ecology of the army ant eciton rapax: An ergonomic enigma? *Ecological Entomology*, 10(2):131–141, 1985.

[30] G. Caldarelli, A. Capocci, P. De Los Rios, e M. A. Munoz. Scale-free networks without growth or preferential attachment: Good get richer. *cond-mat/0207366*, Julho 2002.

[31] D. S. Callaway, J. E. Hopcroft, J. M. Kleinberg, M. E. J. Newman, e S. H. Strogatz. Are randomly grown graphs really random? *Phys. Rev. E*, 64(4):041902, Setembro 2001.

[32] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, e D. J. Watts. Network robustness and fragility: Percolation on random graphs. *Phys. Rev. Lett.*, 85(25):5468–5471, Dezembro 2000.

[33] J. Camacho, R. Guimerà, e L. A. N. Amaral. Analytical solution of a model for complex food webs. *Phys. Rev. E*, 65(3):030901, Fevereiro 2002.

[34] J. Camacho, R. Guimerà, e L. A. N. Amaral. Robust patterns in food web structure. *Phys. Rev. Lett.*, 88(22):228102, Maio 2002.

[35] S. Camazine, J.-L. Deneubourg, N. R. Franks, J. Sneyd, G. Theraulaz, e E. Bonabeau. *Self-organization in biological systems*. Princeton University Press, Agosto 2001.

[36] M. V. Carneiro e I. C. Charret. A criticalidade auto-organizada na pilha de areia. *Revista Brasileira de Ensino de Física*, 27:571–576, Dezembro 2005.

[37] M. Clerc. Binary particle swarm optimisers: toolbox, derivations, and mathematical insights. 2005.

[38] M. Clerc. *Particle swarm optimization*. ISTE Publishing Company, Fevereiro 2006.

[39] M. Clerc e J. Kennedy. The particle swarm - explosion, stability, and convergence in a multidimensional complex space. *IEEE Transactions on Evolutionary Computation*, 6(1):58–73, Fevereiro 2002.

[40] L. S. Coelho e R. Krohling. Predictive controller tuning using modified particle swarm optimization based on cauchy and gaussian distributions. In *Soft Computing: Methodologies and Applications*, pg. 287–298. Springer, 2005.

[41] R. Cohen, K. Erez, D. ben Avraham, e S. Havlin. Resilience of the internet to random breakdowns. *Phys. Rev. Lett.*, 85(21):4626–4628, Novembro 2000.

[42] R. Cohen e S. Havlin. Scale-free networks are ultrasmall. *Phys. Rev. Lett.*, 90(5):058701, Fevereiro 2003.

[43] Y. Cooren, M. Clerc, e P. Siarry. Performance evaluation of tribes, an adaptive particle swarm optimization algorithm. *Swarm Intelligence*, 3(2):149–178, Junho 2009.

[44] L. F. Costa, O. N. Oliveira Jr., G. Travieso, F. A. Rodrigues, P. R. Villas Boas, L. Antiqueira, M. P. Viana, e L. E. C. Rocha. Analyzing and modeling real-world phenomena with complex networks: A survey of applications. *arXiv:0711.3199*, Novembro 2007.

[45] L. F. Costa, F. A. Rodrigues, G. Travieso, e P. R. Villas Boas. Characterization of complex networks: A survey of measurements. *Advances in Physics*, 56(1):167–242, 2007.

[46] J. Davidsen, H. Ebel, e S. Bornholdt. Emergence of a small world from local interactions: Modeling acquaintance networks. *Phys. Rev. Lett.*, 88(12):128701, Março 2002.

[47] J. L. Deneubourg, S. Goss, N. Franks, e J. M. Pasteels. The blind leading the blind: Modeling chemically mediated army ant raid patterns. *Journal of Insect Behavior*, 2(5):719–725, 1989.

[48] D. Dhar. Self-organized critical state of sandpile automaton models. *Phys. Rev. Lett.*, 64(14):1613–1616, Abril 1990.

[49] M. Dorigo. *Optimization, Learning and Natural Algorithms*. Tese de doutorado, Politecnico di Milano, Milão, Itália, 1992.

[50] M. Dorigo e T. Stützle. *Ant colony optimization*. MIT Press, 2004.

[51] M. Dorigo, E. Tuci, R. Groß, V. Trianni, T. H. Labella, S. Nouyan, C. Ampatzis, J. L. Deneubourg, G. Baldassarre, S. Nolfi, et al. The swarm-bots project. *Lecture Notes in Computer Science*, 3342:31–44, 2005.

[52] S. N. Dorogovtsev e J. F. F. Mendes. Evolution of networks with aging of sites. *Phys. Rev. E*, 62(2):1842–1845, Agosto 2000.

[53] S. N. Dorogovtsev e J. F. F. Mendes. Scaling behaviour of developing and decaying networks. *Europhys. Lett.*, 52(1):33–39, Outubro 2000.

[54] S. N. Dorogovtsev e J. F. F. Mendes. Effect of the accelerating growth of communications networks on their structure. *Phys. Rev. E*, 63(2):025101, Janeiro 2001.

[55] S. N. Dorogovtsev e J. F. F. Mendes. Evolution of networks. *Advances in Physics*, 51(4):1079–1187, Setembro 2002.

[56] S. N. Dorogovtsev, J. F. F. Mendes, e A. N. Samukhin. Structure of growing networks with preferential linking. *Phys. Rev. Lett.*, 85(21):4633–4636, Novembro 2000.

[57] S. N. Dorogovtsev, J. F. F. Mendes, e A. N. Samukhin. Size-dependent degree distribution of a scale-free growing network. *Phys. Rev. E*, 63(6):062101, Maio 2001.

[58] H. A. Downing e R. L. Jeanne. Nest construction by the paper wasp, polistes: a test of stigmergy theory. *Animal Behaviour*, 36(6):1729–1739, 1988.

[59] J. P. Dupuy. Nas origens das ciências cognitivas. *Biblioteca básica*, 1996.

[60] R. C. Eberhart e Xiaohui Hu. Human tremor analysis using particle swarm optimization. In *Proceedings of the 1999 Congress on Evolutionary Computation. CEC 99.*, volume 3, pg. – 1930, 1999.

[61] R. C. Eberhart e J. Kennedy. A new optimizer using particle swarm theory. In *Proceedings of the Sixth International Symposium on Micro Machine and Human Science, 1995. MHS '95*, pg. 39–43, Outubro 1995.

[62] R. C. Eberhart e Y. Shi. Comparing inertia weights and constriction factors in particle swarm optimization. In *Proceedings of the 2000 Congress on Evolutionary Computation*, volume 1, pg. 84–88, 2000.

[63] R. C. Eberhart e Y. Shi. "particle swarm optimization and its applications to VLSI design and video technology". In *Proceedings of 2005 IEEE International Workshop on VLSI Design and Video Technology*, pg. xxiii–xxiii, Maio 2005.

[64] R. C. Eberhart, P. Simpson, e R. Dobbins. *Computational intelligence PC tools*. Academic Press Professional, Inc., San Diego, CA, EUA, 1996.

[65] J. M. Epstein e R. L. Axtell. *Growing Artificial Societies: Social Science from the Bottom Up*. MIT Press, 1 edição, Novembro 1996.

[66] P. Erdös e A. Rényi. On random graphs. *Publicationes Mathematicae (Debrecen)*, 6:290–297, 1959.

[67] G. Ergün e G. J. Rodgers. Growing random networks with fitness. *Physica A: Statistical Mechanics and its Applications*, 303(1-2):261–272, 2002.

[68] P. A. Estevez, P. Vera, e K. Saito. Selecting the most influential nodes in social networks. In *International Joint Conference on Neural Networks, 2007. IJCNN 2007*, pg. 2397–2402, Agosto 2007.

[69] L. Euler. Solutio problematis ad geometriam situs pertinentis. *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, 8:128–140, 1736.

[70] M. Faloutsos, P. Faloutsos, e C. Faloutsos. On power-law relationships of the internet topology. In *SIGCOMM '99: Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communication*, pg. 251–262, Nova Iorque, NY, EUA, 1999. ACM.

[71] T. A. Feo e M. G. C. Resende. Greedy randomized adaptive search procedures. *Journal of Global Optimization*, 6(2):109–133, Março 1995.

[72] R. Ferrer i Cancho e R. V. Solé. Optimization in complex networks. In *Statistical Mechanics of Complex Networks*, pg. 114–126. Springer, 2003.

[73] F. Gao e H. Tong. Control a novel discrete chaotic system through particle swarm optimization. In *The Sixth World Congress on Intelligent Control and Automation, 2006. WCICA 2006*, volume 1, pg. 3330–3334, Junho 2006.

[74] Michael R. Garey e David S. Johnson. *Computers and Intractability; A Guide to the Theory of NP-Completeness*. W. H. Freeman & Co., New York, NY, USA, 1990.

[75] M. Girvan e M. E. J. Newman. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America*, 99(12):7821–7826, 2002.

[76] L. Glass e M. C. Mackey. *From clocks to chaos: The rhythms of life*. Princeton Univesity Press, Princeton, NJ, USA, 1988.

[77] F. Glover. Tabu search. In *Encyclopedia of Operations Research and Management Science*, pg. 821–827. Springer, 2001.

[78] A. Godoy e F. J. Von Zuben. A complex neighborhood based particle swarm optimization. In *IEEE Congress on Evolutionary Computation, 2009. CEC '09*, pg. 720–727, Maio 2009.

[79] P.-P. Grassé. La reconstruction du nid et les coordinations interindividuelles chez bellicositermes natalensis et cubitermes sp. la théorie de la stigmergie: Essai d'interprétation du comportement des termites constructeurs. *Insectes Sociaux*, 6(1):41–80, Março 1959.

[80] P.-P. Grassé. Termitologia, tome ii. *Fondation des Sociétés. Construction.*, 1984.

[81] R. Guimerà, L. Danon, A. Díaz-Guilera, F. Giralt, e A. Arenas. Self-similar community structure in a network of human interactions. *Phys. Rev. E*, 68(6):065103, Dezembro 2003.

[82] F. Heppner e U. Grenander. A stochastic nonlinear model for coordinated bird flocks. *American Association for the Advancement of Science*, 1990.

[83] J. M. Hereford. A distributed particle swarm optimization algorithm for swarm robotic applications. In *IEEE Congress on Evolutionary Computation, 2006. CEC 2006*, pg. 1678–1685, Setembro 2006.

[84] N. Higashi e H. Iba. Particle swarm optimization with gaussian mutation. In *Proceedings of the 2003 IEEE Swarm Intelligence Symposium, 2003. SIS '03*, pg. 72–79, Abril 2003.

[85] J. H. Holland. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Michigan, MI, EUA, 1975.

[86] P. Holme, M. Huss, e H. Jeong. Subnetwork hierarchies of biochemical pathways. *Bioinformatics*, 19(4):532–538, 2003.

[87] P. Holme e B. J. Kim. Growing scale-free networks with tunable clustering. *Phys. Rev. E*, 65(2):026107, Janeiro 2002.

[88] B. Hölldobler e E. O. Wilson. The multiple recruitment systems of the african weaver ant oecophylla longinoda (latreille) (hymenoptera: Formicidae). *Behavioral Ecology and Sociobiology*, 3(1):19–60, 1978.

[89] B. Hölldobler e E. O. Wilson. *The ants*. Belknap Press, Março 1990.

[90] S. Janson e M. Middendorf. A hierarchical particle swarm optimizer and its adaptive variant. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 35(6):1272–1282, Dezembro 2005.

[91] H. Jeong, S. P. Mason, A.-L. Barabasi, e Z. N. Oltvai. Lethality and centrality in protein networks. *Nature*, 411(6833):41–42, Maio 2001.

[92] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, e A.-L. Barabasi. The large-scale organization of metabolic networks. *Nature*, 407(6804):651–654, Outubro 2000.

[93] C. Ji, Y. Zhang, S. Gao, P. Yuan, e Z. Li. Particle swarm optimization for mobile ad hoc networks clustering. In *IEEE International Conference on Networking, Sensing and Control*, volume 1, pg. 372–375, Março 2004.

[94] N. Jin e Y. Rahmat-Samii. Advances in particle swarm optimization for antenna designs: Real-number, binary, single-objective and multiobjective implementations. *IEEE Transactions on Antennas and Propagation*, 55(3):556–567, Março 2007.

[95] J. H. Jones e M. S. Handcock. An assessment of preferential attachment as a mechanism for human sexual network formation. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1520):1123–1128, 2003.

[96] J. Kennedy. Thinking is social: Experiments with the adaptive culture model. *The Journal of Conflict Resolution*, 42(1):56–76, 1998.

[97] J. Kennedy. Small worlds and mega-minds: Effects of neighborhood topology on particle swarm performance. In *Proceedings of the 1999 Congress on Evolutionary Computation, 1999. CEC 99*, volume 3, pg. –1938, 1999.

[98] J. Kennedy. Stereotyping: Improving particle swarm performance with cluster analysis. In *Proceedings of the 2000 Congress on Evolutionary Computation*, volume 2, pg. 1507–1512, 2000.

[99] J. Kennedy. Bare bones particle swarms. In *Proceedings of the 2003 IEEE Swarm Intelligence Symposium, 2003. SIS '03*, pg. 80–87, Abril 2003.

[100] J. Kennedy. Swarm intelligence. In *Handbook of nature-inspired and innovative computing*, pg. 187–219. Springer, 2006.

[101] J. Kennedy e R. C. Eberhart. Particle swarm optimization. In *Proceedings of IEEE International Conference on Neural Networks*, volume 4, pg. 1942–1948, Dezembro 1995.

[102] J. Kennedy e R. C. Eberhart. A discrete binary version of the particle swarm algorithm. In *IEEE International Conference on Systems, Man, and Cybernetics, 1997. 'Computational Cybernetics and Simulation'*, volume 5, pg. 4104–4108, Outubro 1997.

[103] J. Kennedy e R. Mendes. Population structure and particle swarm performance. In *Proceedings of the 2002 Congress on Evolutionary Computation, 2002. CEC '02*, volume 2, pg. 1671–1676, 2002.

[104] J. Kennedy e R. Mendes. Neighborhood topologies in fully informed and best-of-neighborhood particle swarms. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 36(4):515–519, Julho 2006.

[105] N. Khemka, C. Jacob, e G. Cole. Making soccer kicks better: a study in particle swarm optimization and evolution strategies. In *The 2005 IEEE Congress on Evolutionary Computation*, volume 1, pg. 735–742, Setembro 2005.

[106] J. Kim, P. L. Krapivsky, B. Kahng, e S. Redner. Infinite-order percolation and giant fluctuations in a protein interaction network. *Phys. Rev. E*, 66(5):055101, Novembro 2002.

[107] S. Kirkpatrick, Jr. Gelatt, C. D., e M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.

[108] P.-C. Ko e P.-C. Lin. A hybrid swarm intelligence based mechanism for earning forecast. *Asian Journal of Information Technology*, 3(3):180–187, 2004.

[109] T. Kohonen. The self-organizing map. *Neurocomputing*, 21(1-3):1–6, Novembro 1998.

[110] H. Kruuk. Predators and anti-predator behavoiour of the black-headed gull (larus ridibundus l.). *Behaviour Supplement*, 11:1–129, 1964.

[111] V. Latora e M. Marchiori. A measure of centrality based on network efficiency. *New Journal of Physics*, 9(6):188, 2007.

[112] S. Li, X. Sun, e Y. Xu. Particle swarm optimization for route planning of unmanned aerial vehicles. In *IEEE International Conference on Information Acquisition*, pg. 1213–1218, Agosto 2006.

[113] J. J. Liang, A. K. Qin, P. M. Suganthan, e S. Baskar. Particle swarm optimization algorithms with novel learning strategies. In *IEEE International Conference on Systems, Man and Cybernetics*, volume 4, pg. 3659–3664, Outubro 2004.

[114] J. J. Liang e P. N. Suganthan. Dynamic multi-swarm particle swarm optimizer with local search. In *The 2005 IEEE Congress on Evolutionary Computation*, volume 1, pg. 522–528, Setembro 2005.

[115] F. Liljeros, C. R. Edling, L. A. N. Amaral, H. E. Stanley, e Y. Åberg. The web of human sexual contacts. *Nature*, 411(6840):907–908, 2001.

[116] D. S. Liu, K. C. Tan, C. K. Goh, e W. K. Ho. On solving multiobjective bin packing problems using particle swarm optimization. In *IEEE Congress on Evolutionary Computation, 2006. CEC 2006*, pg. 2095–2102, Setembro 2006.

[117] J. E. Lloyd. Bioluminescence and communication in insects. *Annual Review of Entomology*, 28(1):131–160, 1983.

[118] H.-Y. Lv, W.-G. Zhou, e C.-G. Zhou. A discrete particle swarm optimization algorithm for phylogenetic tree reconstruction. In *Proceedings of 2004 International Conference on Machine Learning and Cybernetics, 2004*, volume 4, pg. 2650–2654, Agosto 2004.

[119] M. Lüscher. Der lufterneuerung im nest der termite macrotermes natalensis (haviland). *Insectes Sociaux*, 3(2):273–276, 1956.

[120] M. Lüscher. Air-conditioned termite nests. *Scientific American*, 205(1):138–145, 1961.

[121] E. Mallon, S. Pratt, e N. Franks. Individual and collective decision-making during nest site selection by the ant leptothorax albipennis. *Behavioral Ecology and Sociobiology*, 50(4):352–359, 2001.

[122] M. Marchiori e V. Latora. Harmony in the small-world. *Physica A: Statistical Mechanics and its Applications*, 285(3-4):539–546, 2000.

[123] M. Marsili e Y.-C. Zhang. Interacting individuals leading to zipf's law. *Phys. Rev. Lett.*, 80(12):2741–2744, Março 1998.

[124] M. Maruyama. The second cybernetics: Deviation-amplifying mutual causal processes. *American scientist*, 51:164–179, 1963.

[125] R. M. May. Biological populations with nonoverlapping generations: Stable points, stable cycles, and chaos. *Science*, 186(4164):645–647, 1974.

[126] R. M. May. Simple mathematical models with very complicated dynamics. *Nature*, 261(5560):459–467, Junho 1976.

[127] R. Mendes, J. Kennedy, e J. Neves. The fully informed particle swarm: simpler, maybe better. *IEEE Transactions on Evolutionary Computation*, 8(3):204–210, Junho 2004.

[128] S. Milgram. The small world problem. *Psychology Today*, 1:61–67, Maio 1967.

[129] Y.-B. Mo e H.-T. Liu. The design of neural network direct inverse controller based on complex particle swarm optimization algorithm. In *Eighth International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, 2006. SYNASC '06*, pg. 382–388, Setembro 2006.

[130] J. M. Montoya e R. V. Solé. Small world patterns in food webs. *Journal of Theoretical Biology*, 214(3):405–412, Fevereiro 2002.

[131] J. Moody. Race, school integration, and friendship segregation in america. *American Journal of Sociology*, 107(3):679–716, 2001.

[132] A. E. Motter e Y.-C. Lai. Cascade-based attacks on complex networks. *Phys. Rev. E*, 66(6):065102, Dezembro 2002.

[133] V. B. Mountcastle. *Medical physiology*. C. V. Mosby, Saint Louis, 1974.

[134] M. E. J. Newman. Scientific collaboration networks. i. network construction and fundamental results. *Phys. Rev. E*, 64(1):016131, Junho 2001.

[135] M. E. J. Newman. Scientific collaboration networks. II. shortest paths, weighted networks, and centrality. *Phys. Rev. E*, 64(1):016132, Junho 2001.

[136] M. E. J. Newman. The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences of the United States of America*, 98(2):404–409, 2001.

[137] M. E. J. Newman. The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences of the United States of America*, 98(2):404–409, 2001.

[138] M. E. J. Newman. Assortative mixing in networks. *Phys. Rev. Lett.*, 89(20):208701, Outubro 2002.

[139] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.

[140] M. O'Neill e A. Brabazon. Grammatical swarm. *Lecture Notes in Computer Science*, pg. 163–174, 2004.

[141] G. C. Onwubolu e M. Clerc. Optimal path for automated drilling operations by a new heuristic approach using particle swarm optimization. *International Journal of Production Research*, 42(19):473–491, 2004.

[142] J. M. Ottino. Complex networks. *AIChE Journal*, 49(2):292–299, Abril 2004.

[143] R. W. Paine e J. Tani. How hierarchical control self-organizes in artificial adaptive systems. *Adaptive Behavior*, 13(3):211–225, Setembro 2005.

[144] K. E. Parsopoulos e M. N. Vrahatis. Recent approaches to global optimization problems through particle swarm optimization. *Natural Computing*, 1(2):235–306, 2002.

[145] B. L. Partridge, T. Pitcher, J. M. Cullen, e J. Wilson. The three-dimensional structure of fish schools. *Behavioral Ecology and Sociobiology*, 6(4):277–288, Março 1980.

[146] A. Paul, T. A. A. Victoire, e A. E. Jeyakumar. Particle swarm approach for retiming in VLSI. In *IEEE International Symposium on Micro-NanoMechatronics and Human Science*, volume 3, pg. 1532–1535, Dezembro 2003.

[147] T. Peram, K. Veeramachaneni, e C. K. Mohan. Fitness-distance-ratio based particle swarm optimization. In *Proceedings of the 2003 IEEE Swarm Intelligence Symposium, 2003. SIS '03*, pg. 174–181, Abril 2003.

[148] S. L. Pimm. *Food webs*. The University of Chicago Press, Maio 1982.

[149] R. Poli. Analysis of the publications on the applications of particle swarm optimisation. *Journal of Artificial Evolution and Applications*, 2008(1):1–10, Janeiro 2008.

[150] R. Poli, J. Kennedy, e T. Blackwell. Particle swarm optimization. *Swarm Intelligence*, 1(1):33–57, 2007.

[151] R. R. Provine. Contagious yawning and laughter: Significance for sensory feature detection, motor pattern generation, imitation, and the evolution of social behavior. In *Social learning in animals: The roots of culture*, pg. 179–208. Elsevier, 1996.

[152] A. Rapoport e W. J. Horvath. A study of a large sociogram. *Behavioral Science*, 6(4):279–291, 1961.

[153] S. Redner. How popular is your paper? an empirical study of the citation distribution. *The European Physical Journal B*, 4(2), 1998.

[154] M. Reyes-Sierra e C. A. C. Coello. Multi-objective particle swarm optimizers: A survey of the state-of-the-art. *International Journal of Computational Intelligence Research*, 2(3):287–308, 2006.

[155] C. W. Reynolds. Flocks, herds and schools: A distributed behavioral model. In *Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, pg. 25–34. ACM New York, NY, USA, 1987.

[156] T. J. Richer e T. M. Blackwell. The Lévy particle swarm. In *IEEE Congress on Evolutionary Computation, 2006. CEC 2006*, pg. 808–815, Setembro 2006.

[157] J. Riget e J. Vesterstroem. A diversity-guided particle swarm optimizer - the ARPSO. Technical report, Aarhus University - Department of Computer Science, 2002.

[158] T. D. Seeley. The honey bee colony as a superorganism. *American Scientist*, 77:546–553, Novembro 1989.

[159] T. D. Seeley e W. F. Towne. Tactics of dance choice in honey bees: Do foragers compare dances? *Behavioral Ecology and Sociobiology*, 30(1):59–69, 1992.

[160] Y.-W. Shang e Y.-H. Qiu. A note on the extended Rosenbrock function. *Evol. Comput.*, 14(1):119–126, 2006.

[161] P. S. Shelokar, V. K. Jayaraman, e B. D. Kulkarni. An ant colony approach for clustering. *Analytica Chimica Acta*, 509(2):187–195, 2004.

[162] S. S. Shen-Orr, R. Milo, S. Mangan, e U. Alon. Network motifs in the transcriptional regulation network of Escherichia coli. *Nat Genet*, 31(1):64–68, Abril 2002.

[163] Y. Shi e R. C. Eberhart. Empirical study of particle swarm optimization. In *Proceedings of the 1999 Congress on Evolutionary Computation, 1999. CEC 99.*, volume 3, pg. –1950, 1999.

[164] M. M. Soares, G. Corso, e L. S. Lucena. The network of syllables in Portuguese. *Physica A: Statistical Mechanics and its Applications*, 355(2-4):678–684, 2005.

[165] R. V. Solé, R. Pastor-Satorras, E. Smith, e T. B. Kepler. A model of large-scale proteome evolution. *Advances in Complex Systems*, 5(1):43–54, Janeiro 2002.

[166] L. Steels. Towards a theory of emergent functionality. In *Proceedings of the First International Conference on Simulation of Adaptive Behavior on From Animals to Animats*, pg. 451–461. MIT Press Cambridge, MA, USA, 1991.

[167] A. M. Stuart. Alarm, defense, and construction behavior relationships in termites (isoptera). *Science*, 156(3778):1123–1125, 1967.

[168] P. N. Suganthan. Particle swarm optimiser with neighbourhood operator. In *Proceedings of the 1999 Congress on Evolutionary Computation, 1999. CEC 99*, volume 3, pg. –1962, 1999.

[169] M. Sugisaka e X. Fan. An effective search method for neural network based face detection using particle swarm optimization. *IEICE Transactions on Information and Systems*, 88(2):214–222, 2005.

[170] D. J. T. Sumpter e M. Beekman. From nonlinearity to optimality: Pheromone trail foraging by ants. *Animal Behaviour*, 66(2):273–280, 2003.

[171] G. A. Sword, P. D. Lorch, e D. T. Gwynne. Insect behaviour: Migratory bands give crickets protection. *Nature*, 433(7027):703+, Fevereiro 2005.

[172] Z. Toroczkai e K. E. Bassler. Network dynamics: Jamming is limited in scale-free systems. *Nature*, 428(6984):716, Abril 2004.

[173] S. Ujjin e P. J. Bentley. Particle swarm optimization recommender system. In *Proceedings of the 2003 IEEE Swarm Intelligence Symposium, 2003. SIS '03*, pg. 124–131, Abril 2003.

[174] S. Valverde, R. Ferrer i Cancho, e R. V. Solé. Scale-free networks from optimal design. *Europhys. Lett.*, 60(4):512–517, Novembro 2002.

[175] A. Vazquez. Statistics of citation networks. *cond-mat/0105031*, Maio 2001.

[176] K. von Frisch. *The dance language and orientation of bees*. Harvard University Press, Cambridge, MA, EUA, 1967.

[177] A. Wagner e D. A. Fell. The small world inside large metabolic networks. *Proceedings. Biological Sciences / The Royal Society*, 268(1478):1803–1810, Setembro 2001.

[178] D. J. Watts. *Small Worlds: The Dynamics of Networks Between Order and Randomness*. Princeton University Press, 2003.

[179] D. J. Watts e S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, Junho 1998.

[180] D. M. Wilkinson e B. A. Huberman. A method for finding communities of related genes. *Proceedings of the National Academy of Sciences of the United States of America*, 101(Suppl 1):5241–5248, 2004.

[181] E. O. Wilson. The relation between caste ratios and division of labor in the ant genus pheidole (hymenoptera: Formicidae). *Behavioral Ecology and Sociobiology*, 16(1):89–98, 1984.

[182] E. O. Wilson. *Sociobiology: The new synthesis*. Belknap Press, 2000.

[183] M. L. Winston. *The Biology of the honey bee*. Harvard University Press, Cambridge, MA, EUA, 1987.

[184] S. Wolfram. *A new kind of science*. Wolfram Media Inc., Champaign, IL, USA, 1 edição, Maio 2002.

[185] R. Xu, G. C. Anagnostopoulos, e D. C. Wunsch II. Multiclass cancer classification using semisupervised ellipsoid artmap and particle swarm optimization with gene expression data. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 4(1):65–77, Março 2007.

[186] R. Xu e D. C. Wunsch II. Gene regulatory networks inference with recurrent neural network models. In *Proceedings of 2005 IEEE International Joint Conference on Neural Networks, 2005. IJCNN '05*, volume 1, pg. 286–291, Agosto 2005.

[187] H. Yoshida, K. Kawata, Y. Fukuyama, S. Takayama, e Y. Nakanishi. A particle swarm optimization for reactive power and voltage control considering voltage security assessment. *IEEE Transactions on Power Systems*, 15(4):1232–1239, Novembro 2000.

[188] A. Zaslaver, A. E. Mayo, R. Rosenberg, P. Bashkin, H. Sberro, M. Tsalyuk, M. G. Surette, e U. Alon. Just-in-time transcription program in metabolic pathways. *Nat Genet*, 36(5):486–491, Maio 2004.

[189] H. Zhu, X. Wang, e J.-Y. Zhu. Effect of aging on network structure. *Phys. Rev. E*, 68(5):056121, Novembro 2003.

[190] Q. Zhu, L. Qian, Y. Li, e S. Zhu. An improved particle swarm optimization algorithm for vehicle routing problem with time windows. In *IEEE Congress on Evolutionary Computation, 2006. CEC 2006*, pg. 1386–1390, Setembro 2006.

Apêndice A

Gráficos das funções de teste utilizadas

Embora os experimentos realizados e reportados no capítulo 5 tenham empregado dimensão 30 para o argumento das funções de teste, são apresentados a seguir os gráficos considerando argumentos de dimensão 2, permitindo a visualização e uma análise qualitativa das principais características e demandas de otimização das funções de teste.

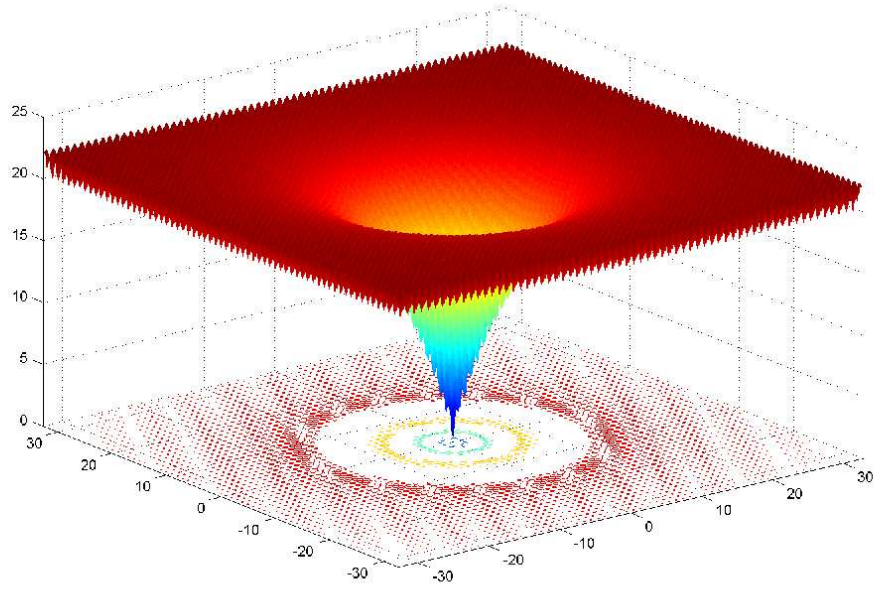


Fig. A.1: Função Ackley.

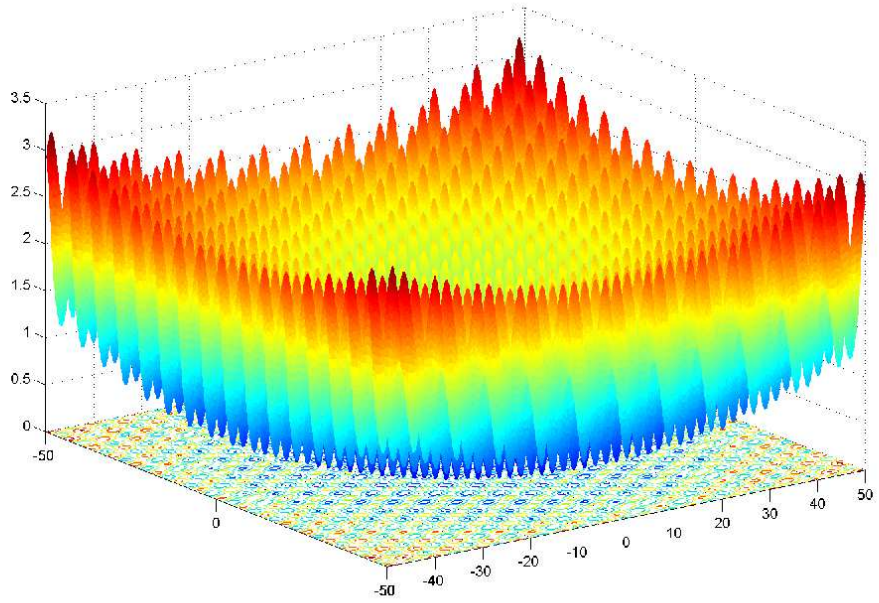


Fig. A.2: Função Griewank.

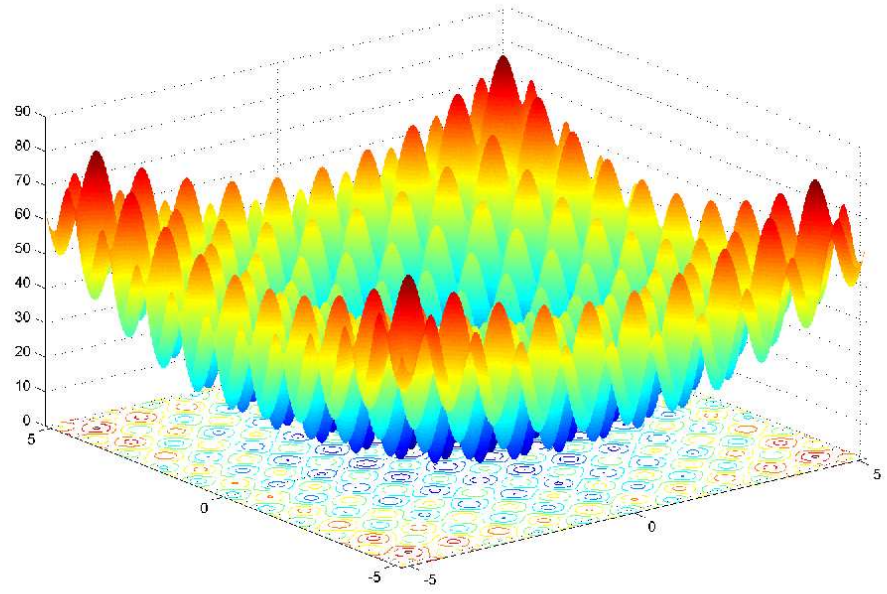


Fig. A.3: Função Rastrigin.

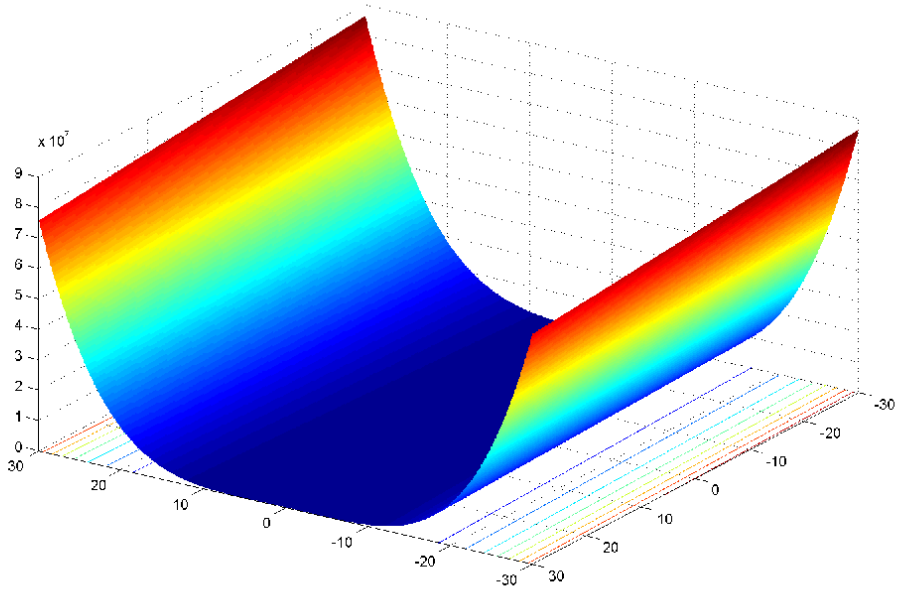


Fig. A.4: Função Rosenbrock.

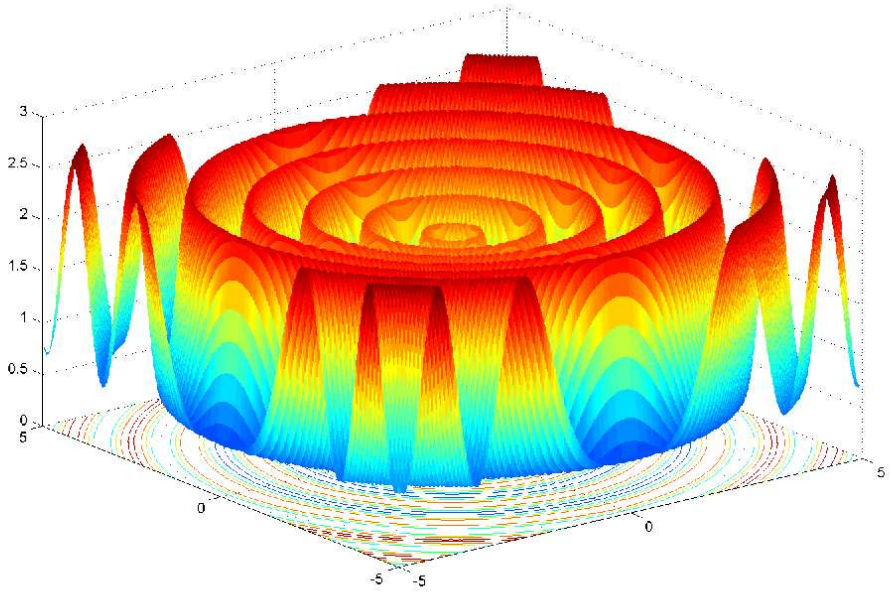


Fig. A.5: Função Salomon.

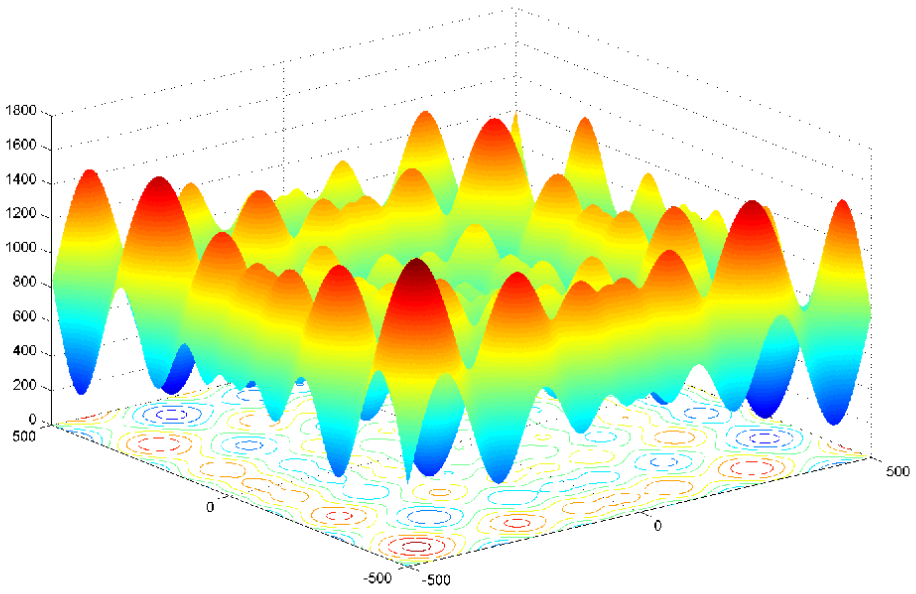


Fig. A.6: Função Schwefel.

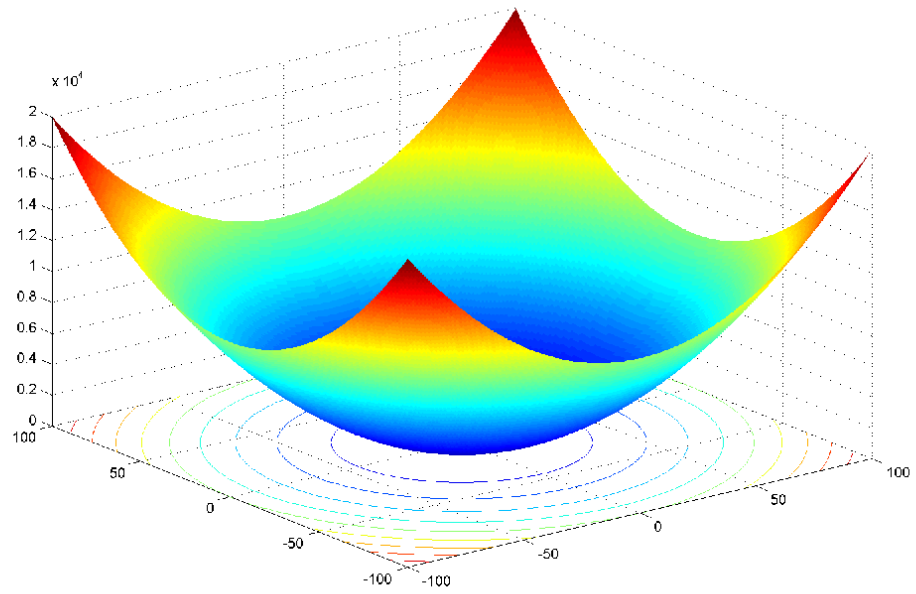


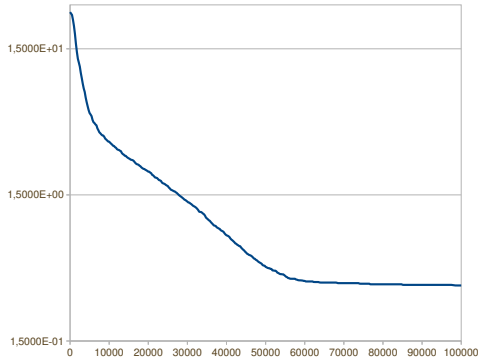
Fig. A.7: Função Sphere.

Apêndice B

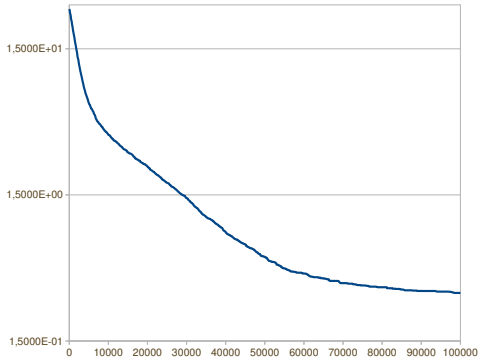
Gráficos de convergência

Para facilitar comparações, todos os gráficos para uma mesma função encontram-se em uma mesma escala. À exceção dos gráficos relativos às funções Rastrigin e Schwefel, o eixo vertical (resultados obtidos) encontra-se em escala logarítmica.

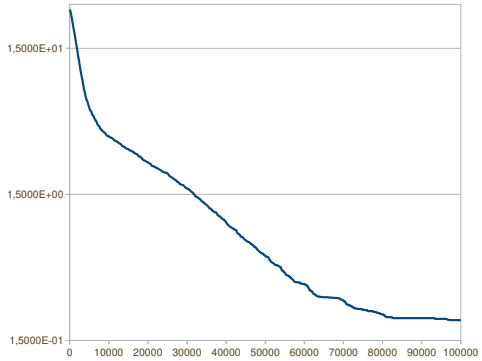
B.0.1 Função Salomon



(a) Vizinhaça global

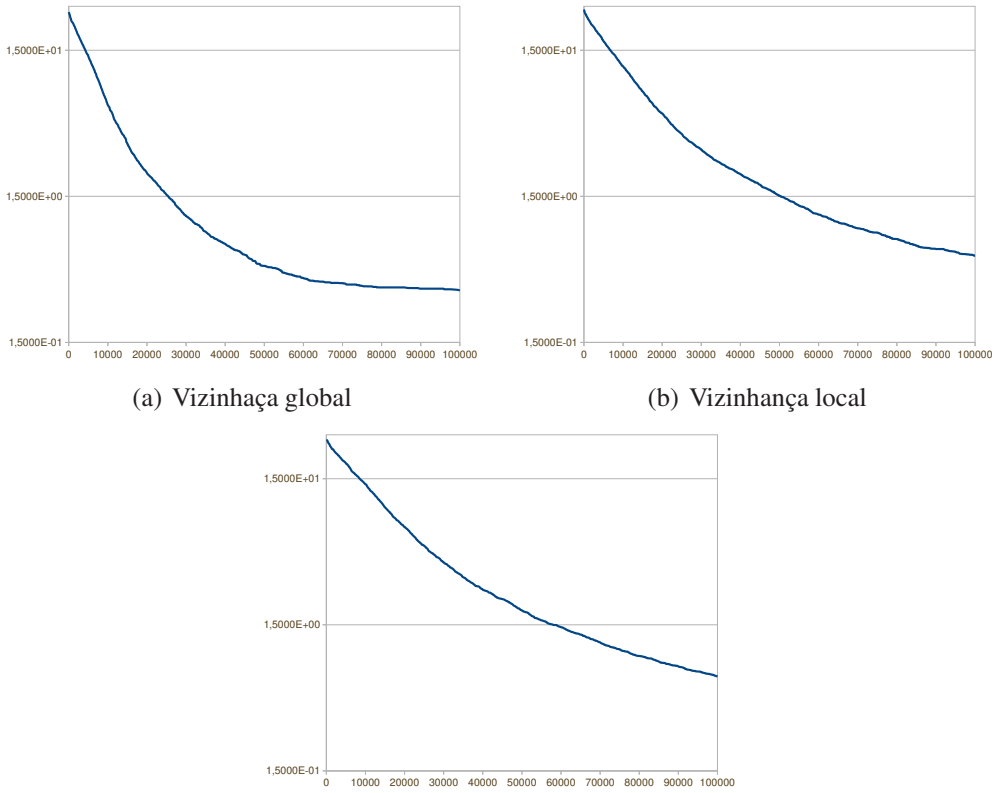


(b) Vizinhança local



(c) Vizinhança von Neumann

Fig. B.1: Gráficos de convergência do PSO para a função Salomon.

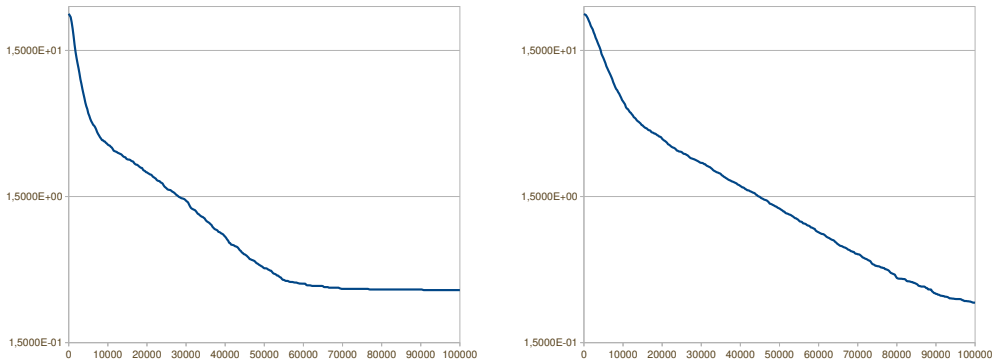


(a) Vizinhaça global

(b) Vizinhança local

(c) Vizinhança von Neumann

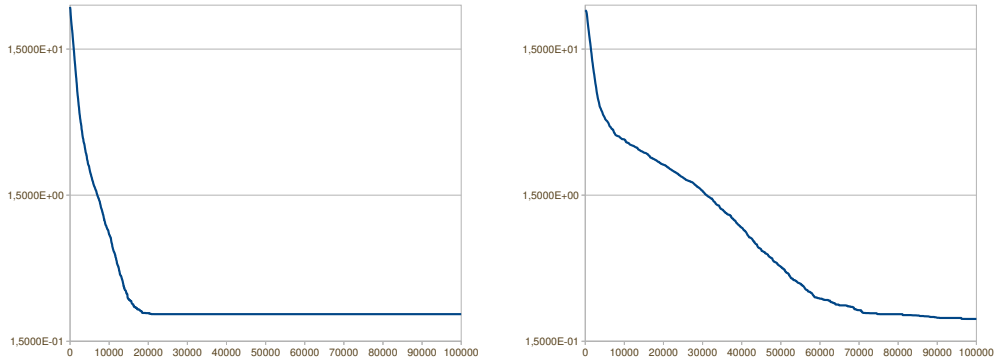
Fig. B.2: Gráficos de convergência do *Bare Bones Particle Swarm* para a função Salomon.



(a) Vizinhaça global

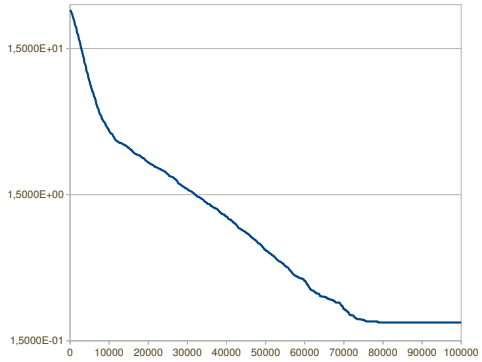
(b) Vizinhança von Neumann

Fig. B.3: Gráficos de convergência do ARPSO para a função Salomon.

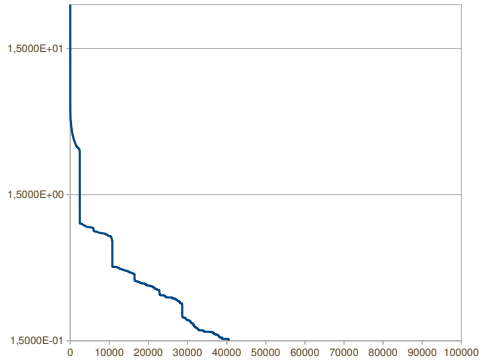


(a) FIPS com vizinhança von Neumann

(b) H-PSO

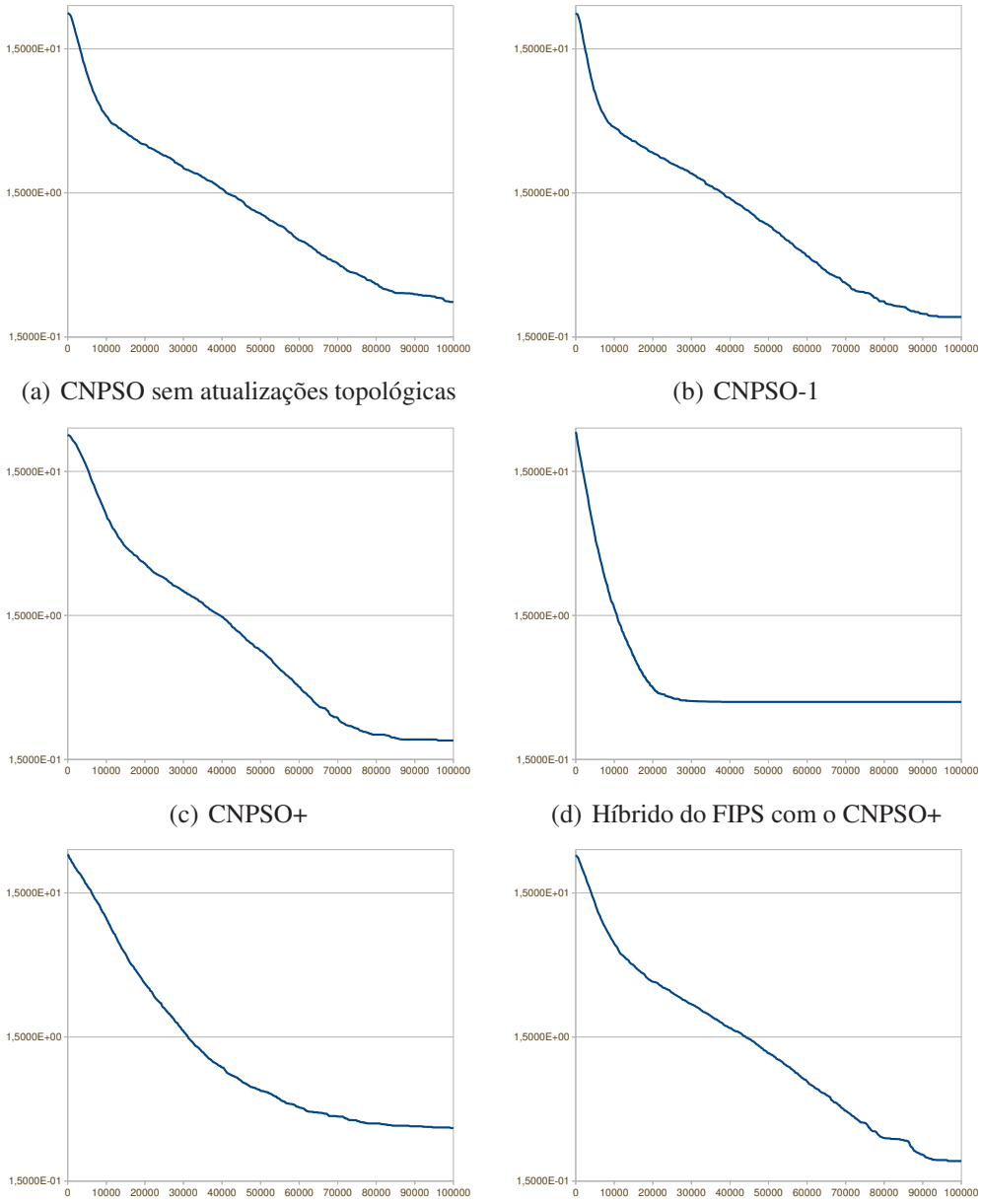


(c) DMS-PSO



(d) TRIBES

Fig. B.4: Gráficos de convergência dos demais PSOs vistos na literatura para a função Salomon.



(a) CNPSO sem atualizações topológicas

(b) CNPSO-1

(c) CNPSO+

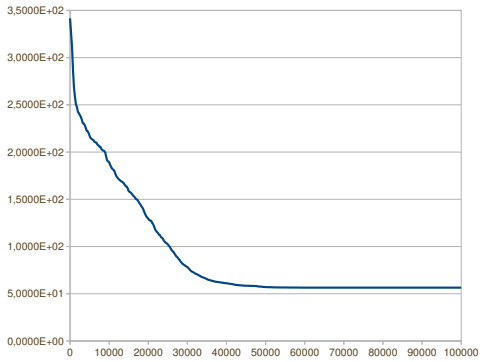
(d) Híbrido do FIPS com o CNPSO+

(e) Híbrido do *Bare Bones Particle Swarm* com o CNPSO+

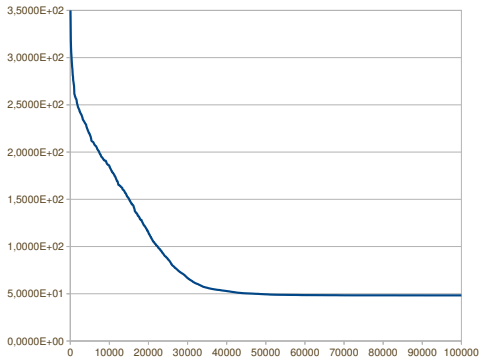
(f) Híbrido do AR-PSO com o CNPSO+

Fig. B.5: Gráficos de convergência dos CNPSOs para a função Salomon.

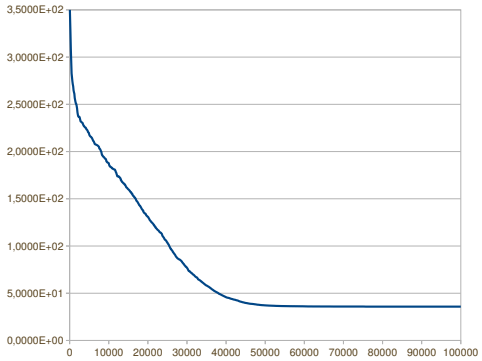
B.0.2 Função Rastrigin



(a) Vizinhaça global

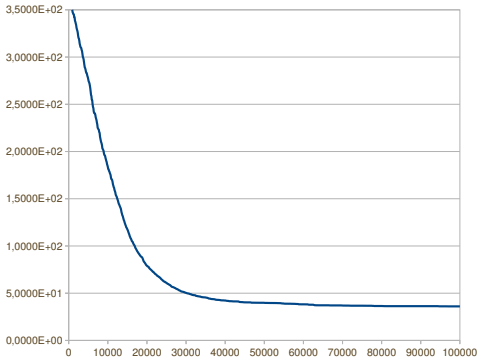


(b) Vizinhança local

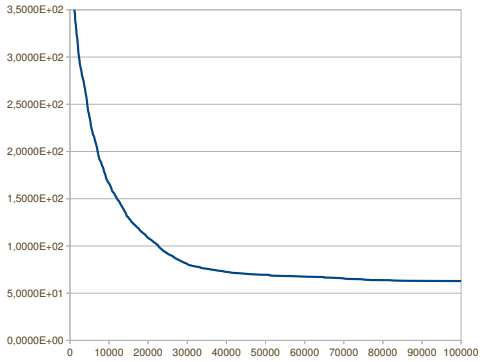


(c) Vizinhança von Neumann

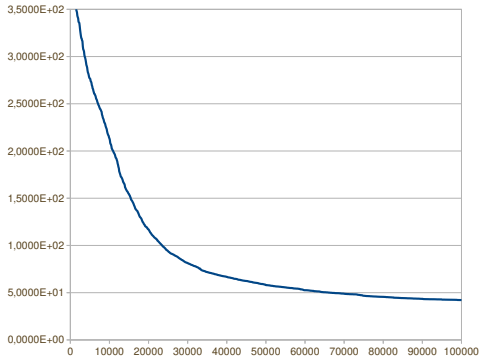
Fig. B.6: Gráficos de convergência do PSO para a função Rastrigin.



(a) Vizinhaça global

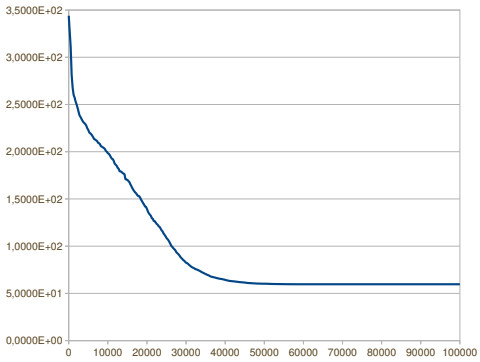


(b) Vizinhança local

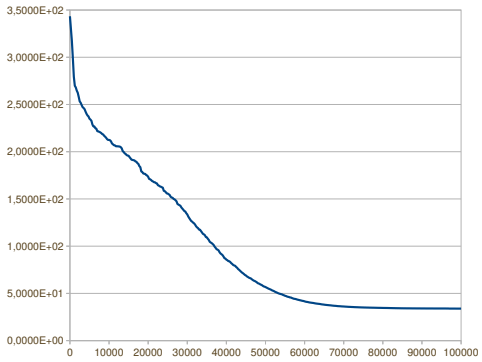


(c) Vizinhança von Neumann

Fig. B.7: Gráficos de convergência do *Bare Bones Particle Swarm* para a função Rastrigin.

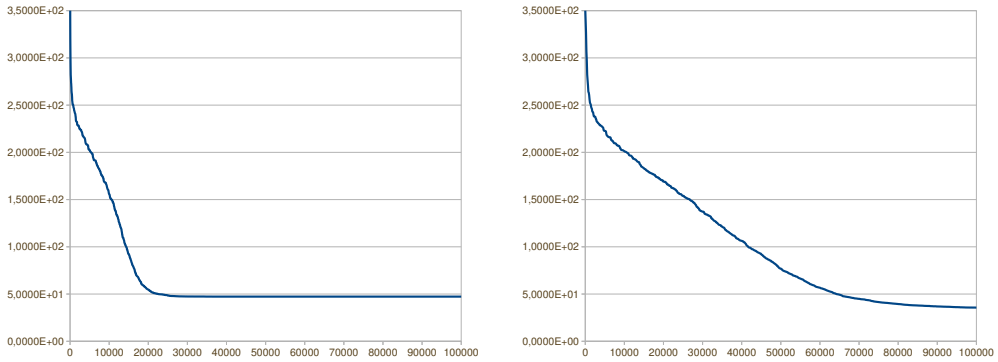


(a) Vizinhaça global



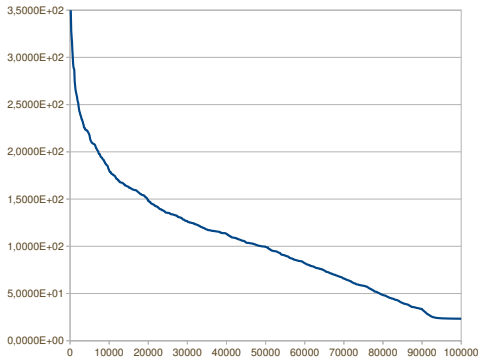
(b) Vizinhança von Neumann

Fig. B.8: Gráficos de convergência do ARPSO para a função Rastrigin.

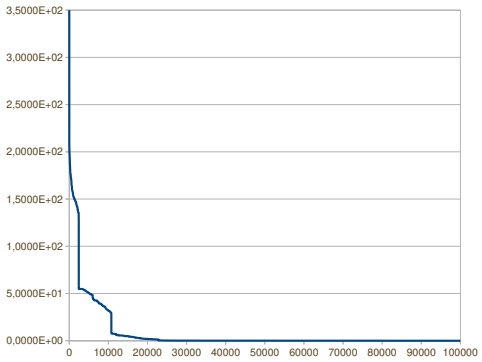


(a) FIPS com vizinhança von Neumann

(b) H-PSO

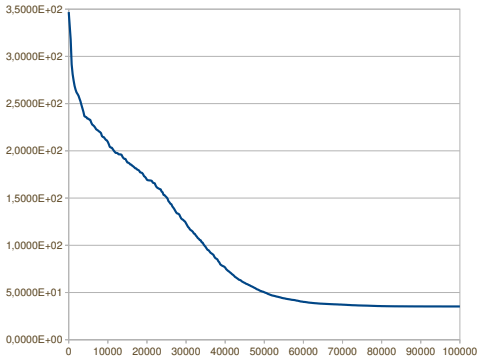


(c) DMS-PSO

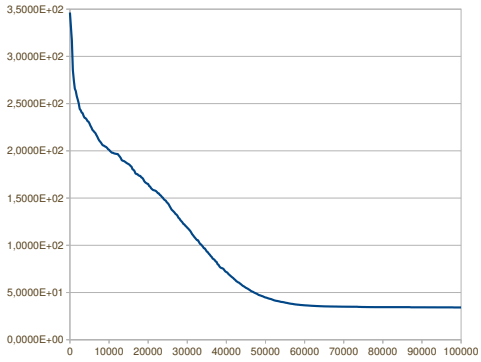


(d) TRIBES

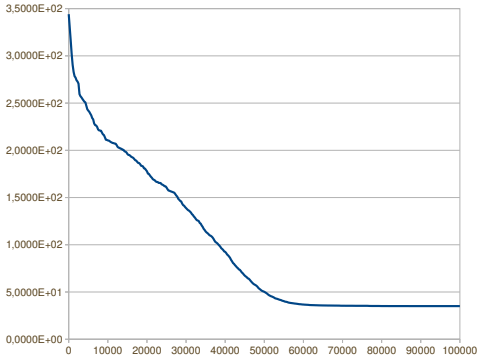
Fig. B.9: Gráficos de convergência dos demais PSOs vistos na literatura para a função Rastrigin.



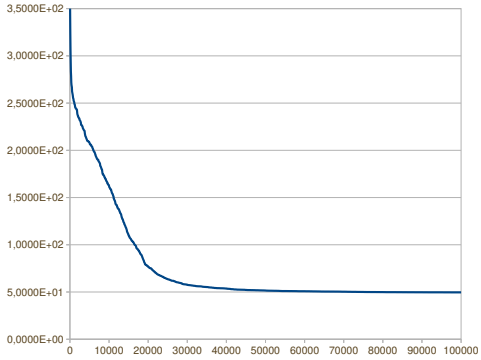
(a) CNPSO sem atualizações topológicas



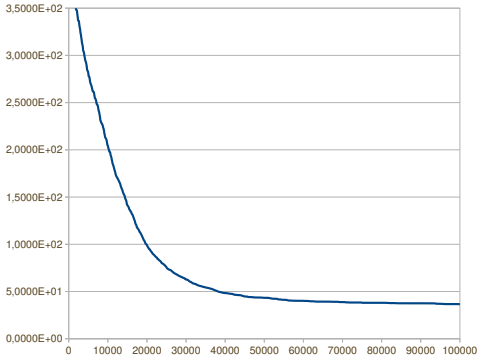
(b) CNPSO-1



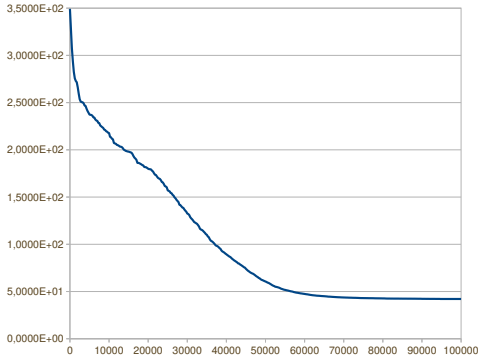
(c) CNPSO+



(d) Híbrido do FIPS com o CNPSO+



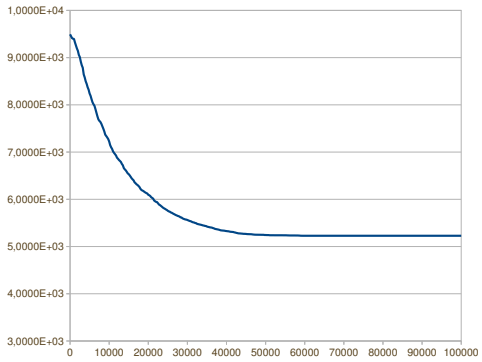
(e) Híbrido do *Bare Bones Particle Swarm* com o CNPSO+



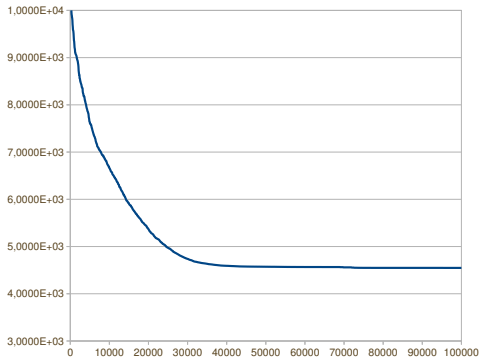
(f) Híbrido do AR-PSO com o CNPSO+

Fig. B.10: Gráficos de convergência dos CNPSOs para a função Rastrigin.

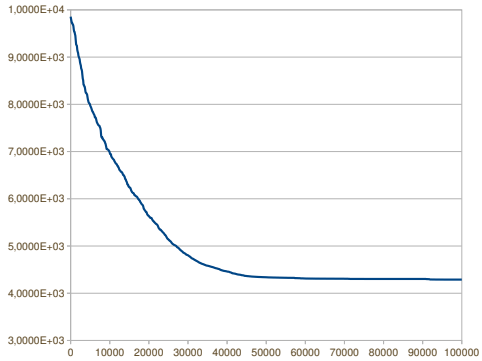
B.0.3 Função Schwefel



(a) Vizinhaça global

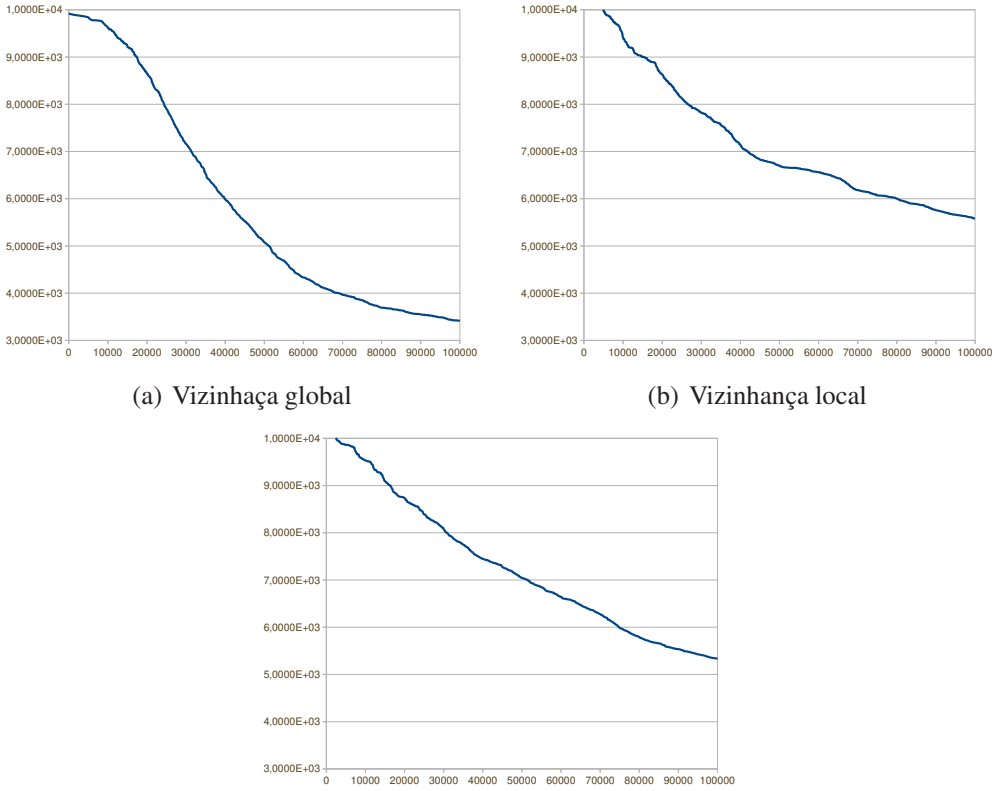


(b) Vizinhança local



(c) Vizinhança von Neumann

Fig. B.11: Gráficos de convergência do PSO para a função Schwefel.

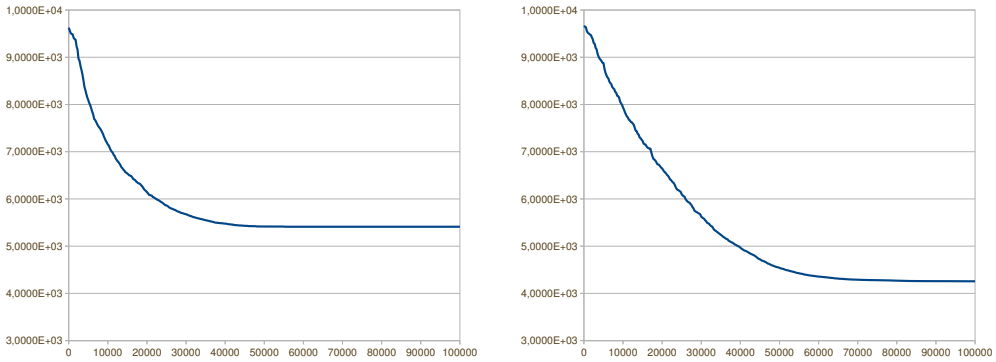


(a) Vizinhaça global

(b) Vizinhança local

(c) Vizinhança von Neumann

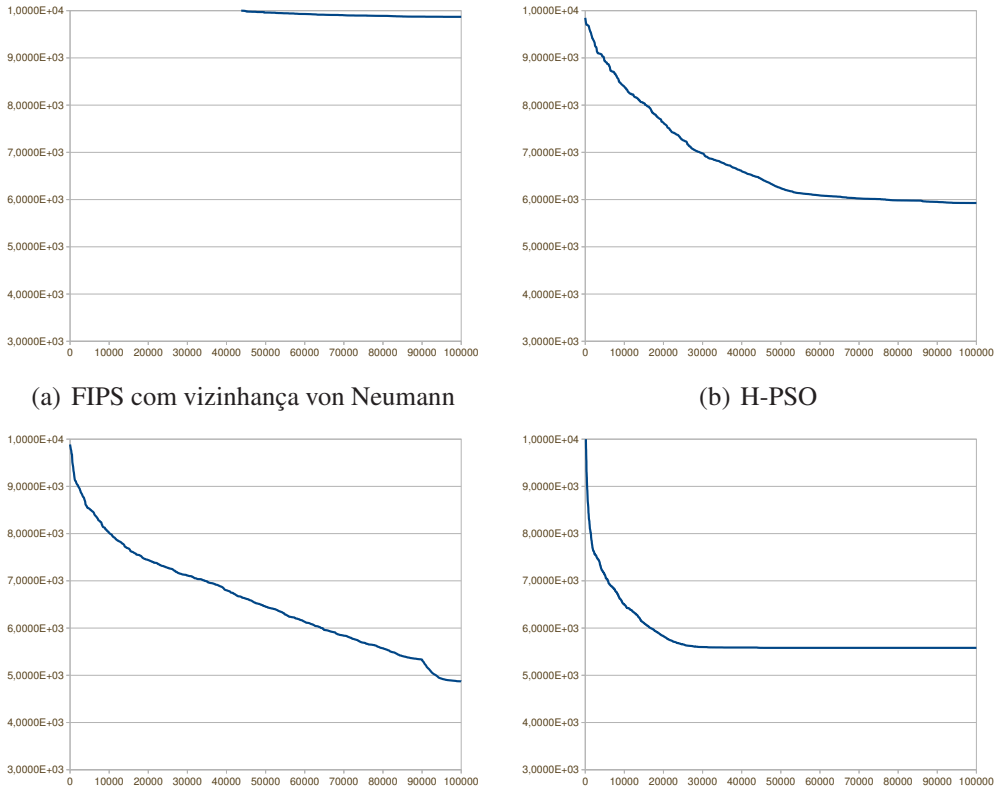
Fig. B.12: Gráficos de convergência do *Bare Bones Particle Swarm* para a função Schwefel.



(a) Vizinhaça global

(b) Vizinhança von Neumann

Fig. B.13: Gráficos de convergência do ARPSO para a função Schwefel.



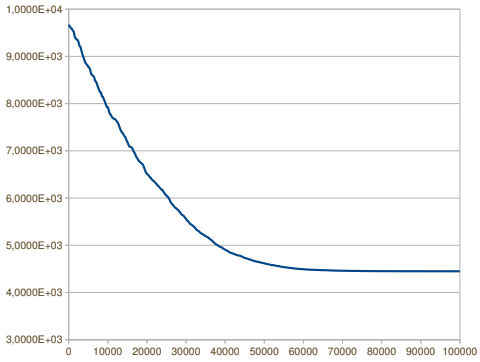
(a) FIPS com vizinhança von Neumann

(b) H-PSO

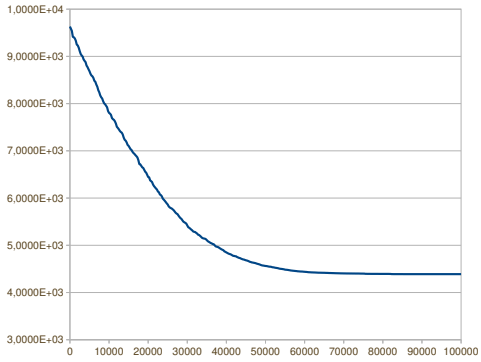
(c) DMS-PSO

(d) TRIBES

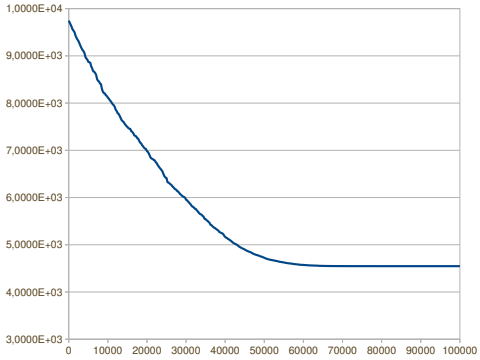
Fig. B.14: Gráficos de convergência dos demais PSOs vistos na literatura para a função Schwefel.



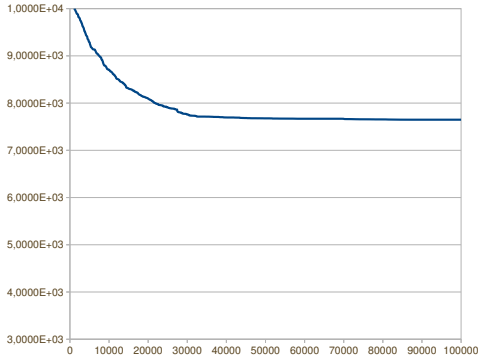
(a) CNPSO sem atualizações topológicas



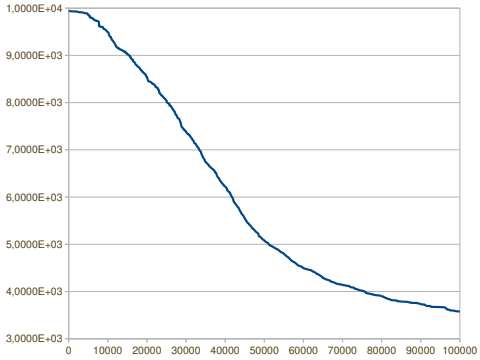
(b) CNPSO-1



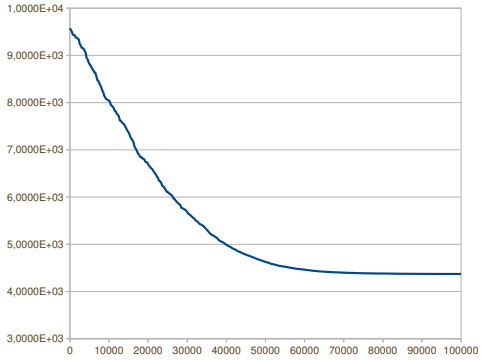
(c) CNPSO+



(d) Híbrido do FIPS com o CNPSO+



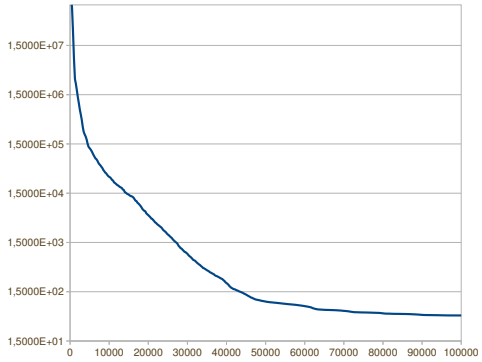
(e) Híbrido do *Bare Bones Particle Swarm* com o CNPSO+



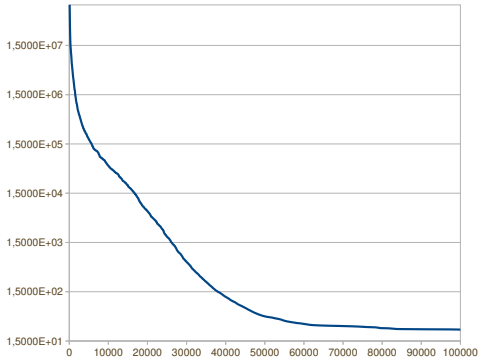
(f) Híbrido do AR-PSO com o CNPSO+

Fig. B.15: Gráficos de convergência dos CNPSOs para a função Schwefel.

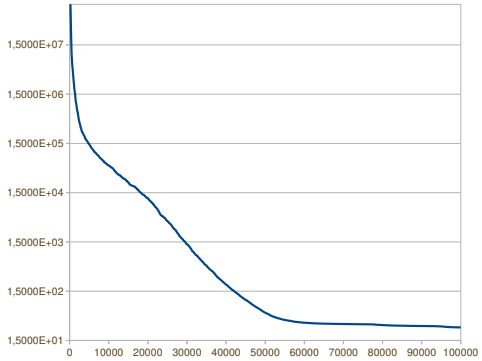
B.0.4 Função Rosenbrock



(a) Vizinhaça global

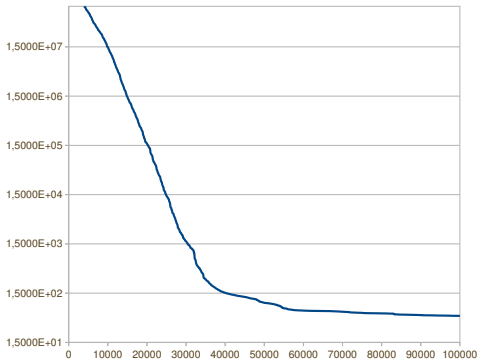


(b) Vizinhança local

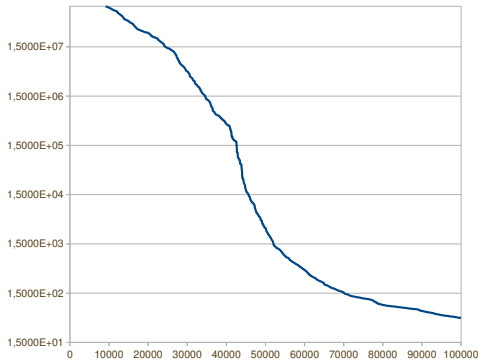


(c) Vizinhança von Neumann

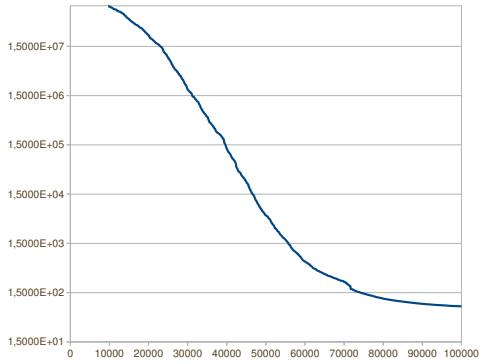
Fig. B.16: Gráficos de convergência do PSO para a função Rosenbrock.



(a) Vizinhaça global

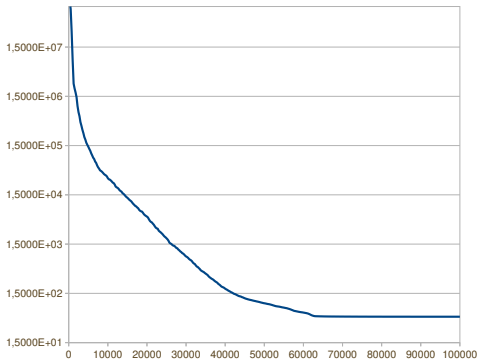


(b) Vizinhança local

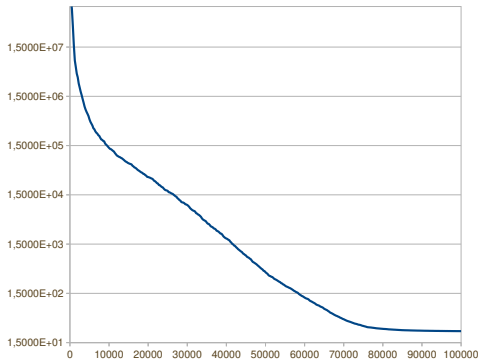


(c) Vizinhança von Neumann

Fig. B.17: Gráficos de convergência do *Bare Bones Particle Swarm* para a função Rosenbrock.

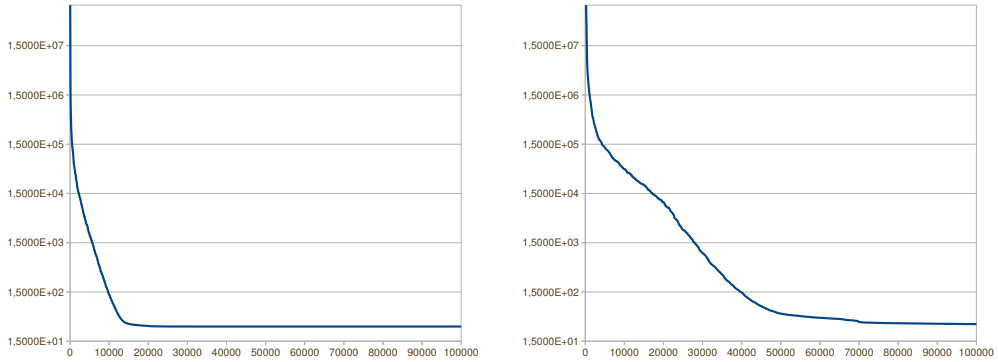


(a) Vizinhaça global



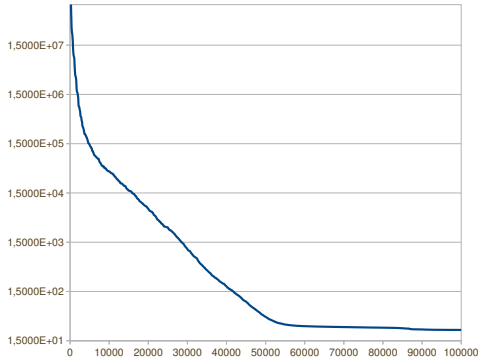
(b) Vizinhança von Neumann

Fig. B.18: Gráficos de convergência do ARPSO para a função Rosenbrock.

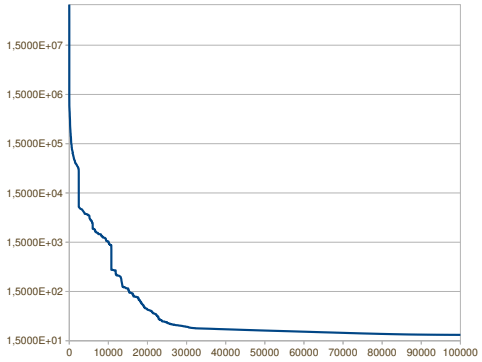


(a) FIPS com vizinhança von Neumann

(b) H-PSO

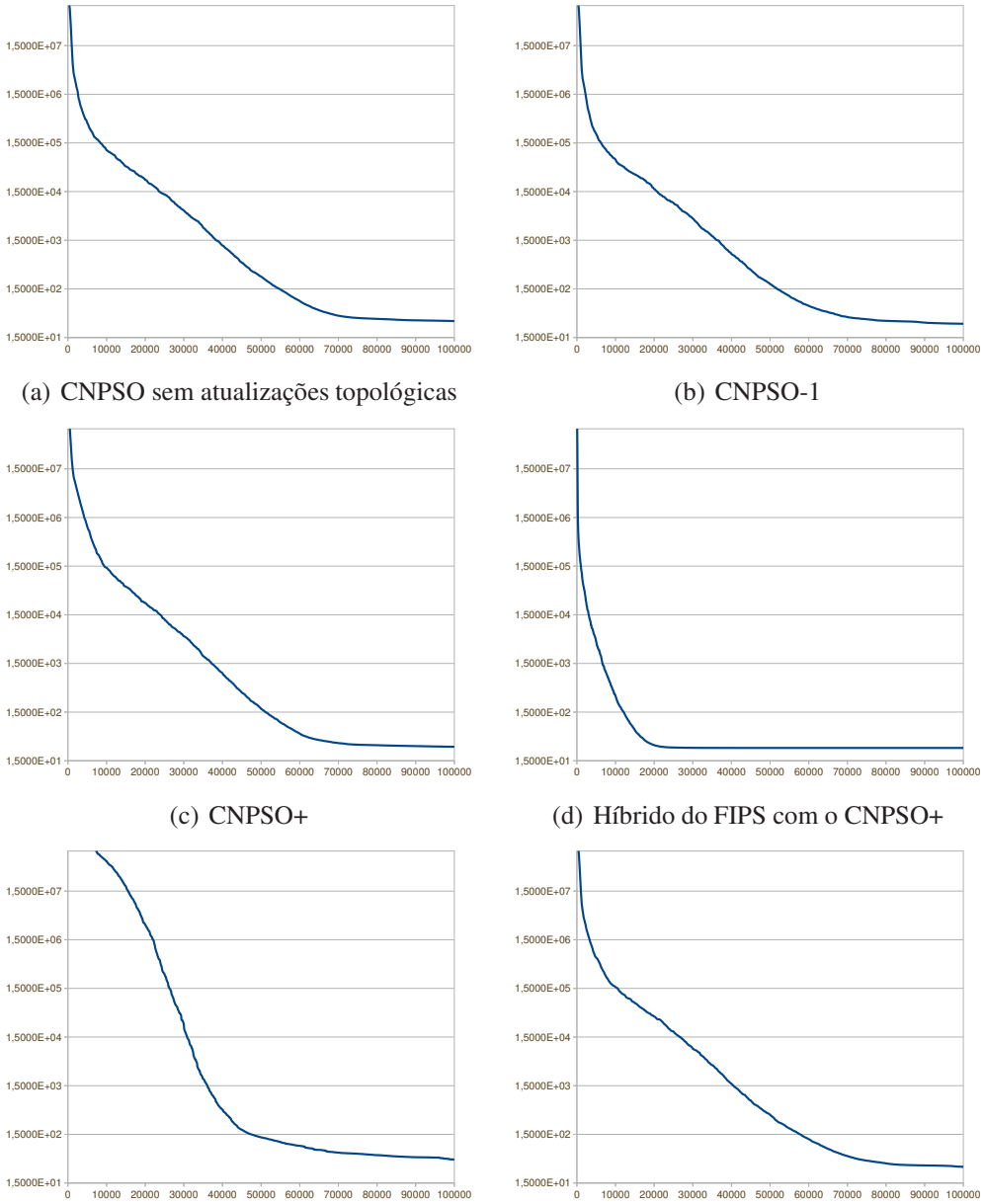


(c) DMS-PSO



(d) TRIBES

Fig. B.19: Gráficos de convergência dos demais PSOs vistos na literatura para a função Rosenbrock.



(a) CNPSO sem atualizações topológicas

(b) CNPSO-1

(c) CNPSO+

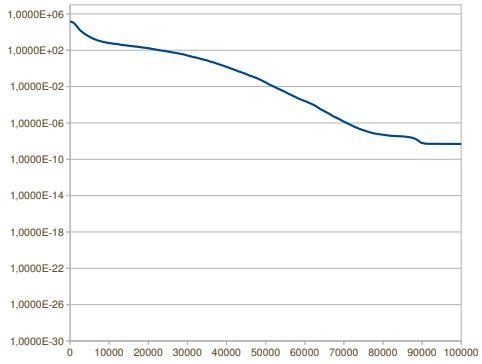
(d) Híbrido do FIPS com o CNPSO+

(e) Híbrido do *Bare Bones Particle Swarm* com o CNPSO+

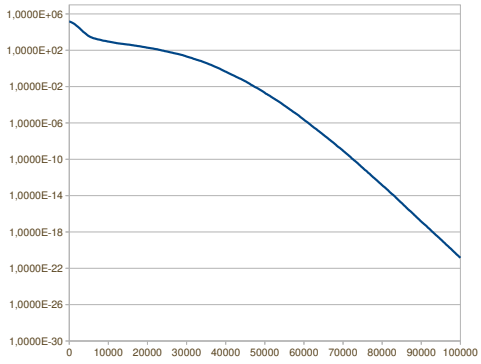
(f) Híbrido do AR-PSO com o CNPSO+

Fig. B.20: Gráficos de convergência dos CNPSOs para a função Rosenbrock.

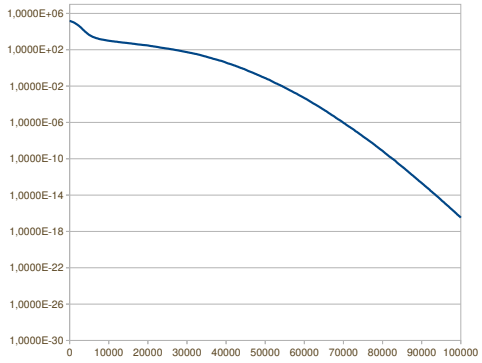
B.0.5 Função Sphere



(a) Vizinhaça global

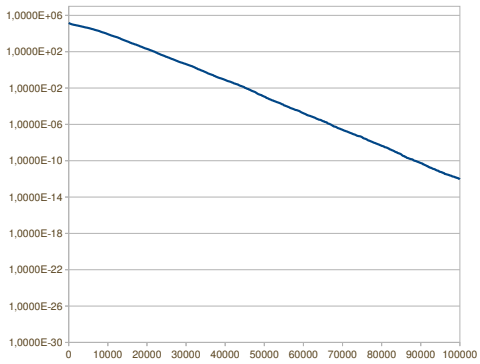


(b) Vizinhança local

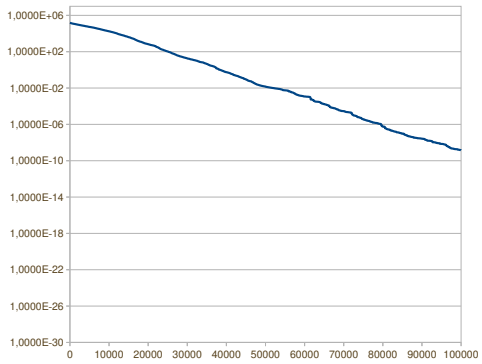


(c) Vizinhança von Neumann

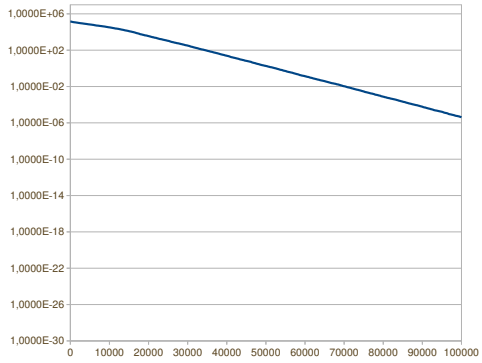
Fig. B.21: Gráficos de convergência do PSO para a função Sphere.



(a) Vizinhaça global

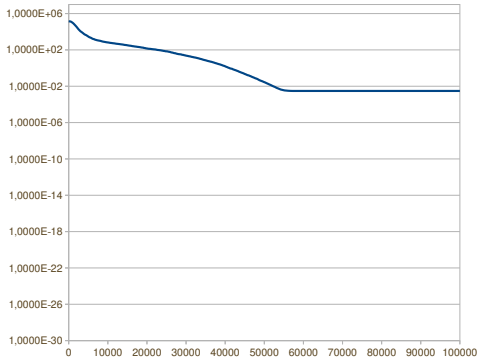


(b) Vizinhança local

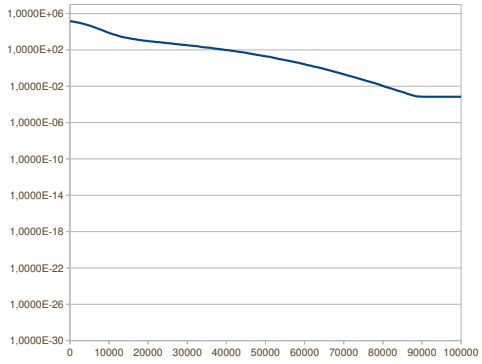


(c) Vizinhança von Neumann

Fig. B.22: Gráficos de convergência do *Bare Bones Particle Swarm* para a função Sphere.

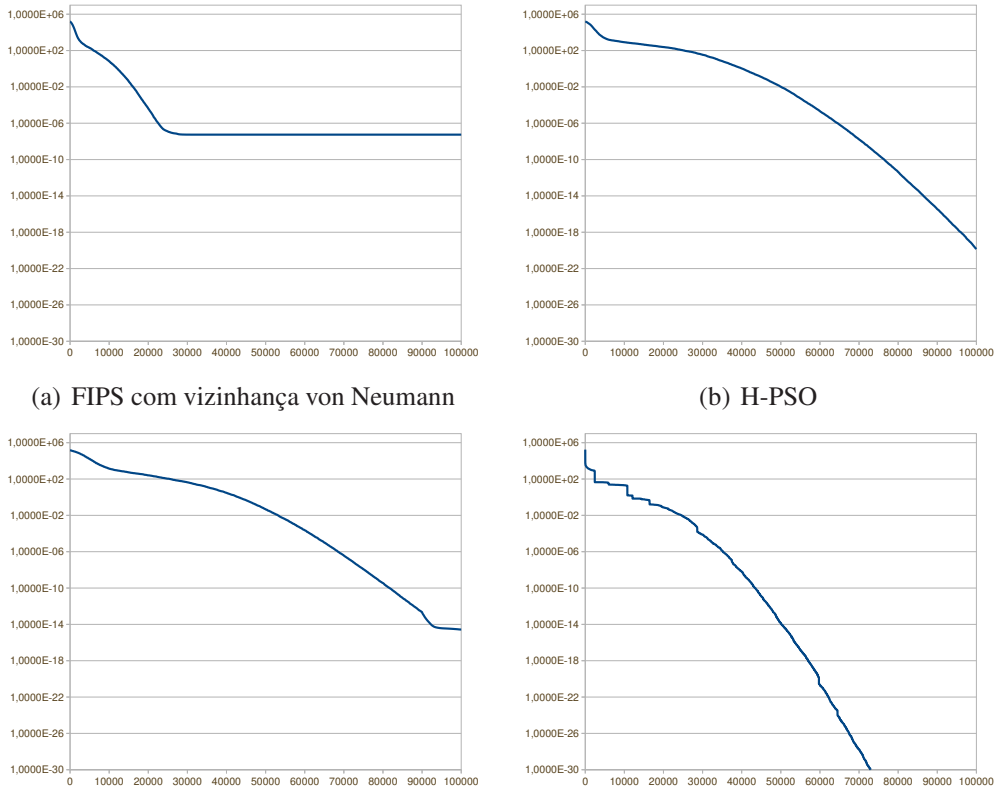


(a) Vizinhaça global



(b) Vizinhança von Neumann

Fig. B.23: Gráficos de convergência do ARPSO para a função Sphere.



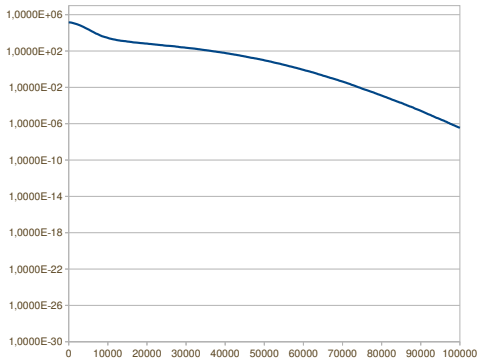
(a) FIPS com vizinhança von Neumann

(b) H-PSO

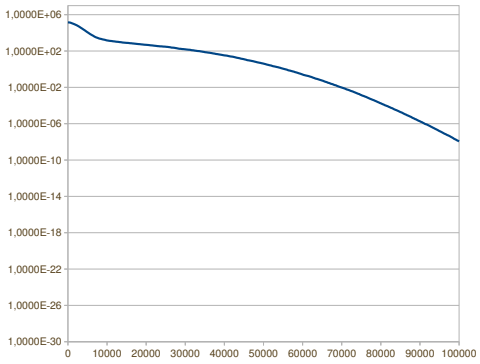
(c) DMS-PSO

(d) TRIBES

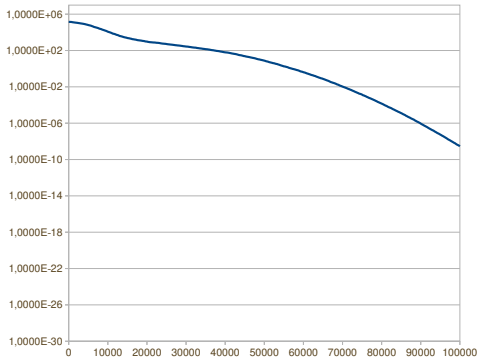
Fig. B.24: Gráficos de convergência dos demais PSOs vistos na literatura para a função Sphere.



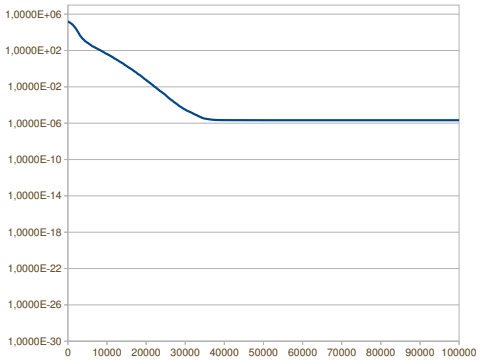
(a) CNPSO sem atualizações topológicas



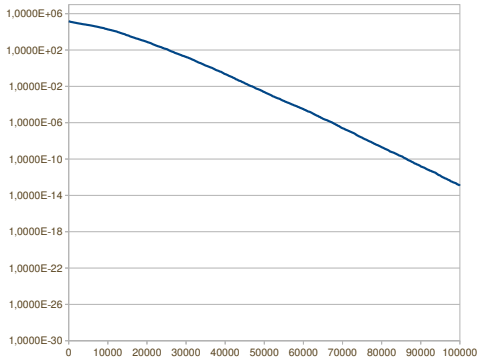
(b) CNPSO-1



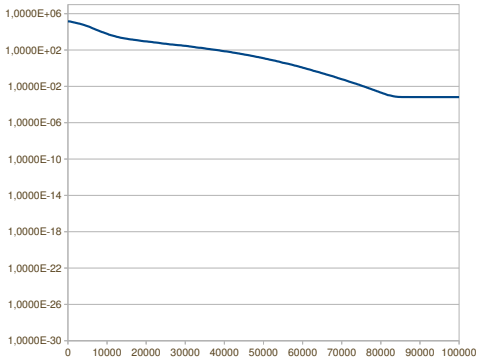
(c) CNPSO+



(d) Híbrido do FIPS com o CNPSO+



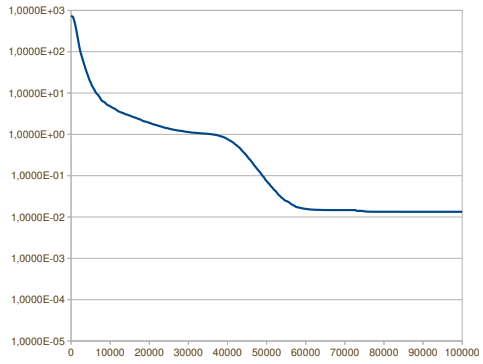
(e) Híbrido do *Bare Bones Particle Swarm* com o CNPSO+



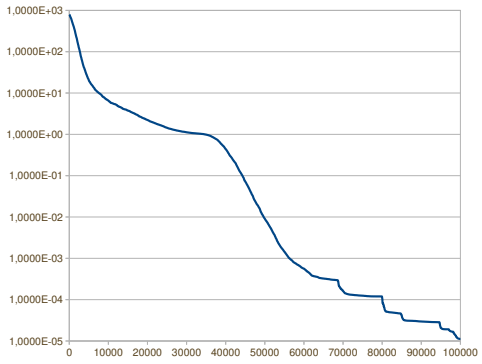
(f) Híbrido do AR-PSO com o CNPSO+

Fig. B.25: Gráficos de convergência dos CNPSOs para a função Sphere.

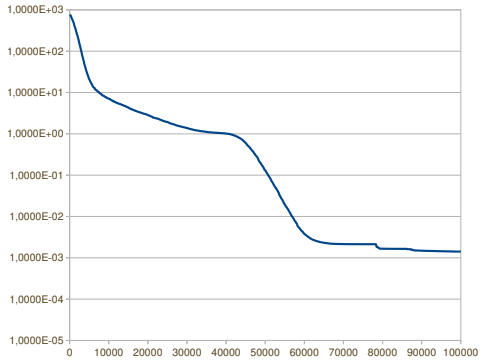
B.0.6 Função Griewank



(a) Vizinhaça global

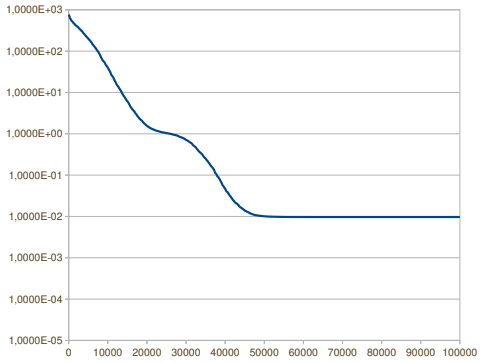


(b) Vizinhança local

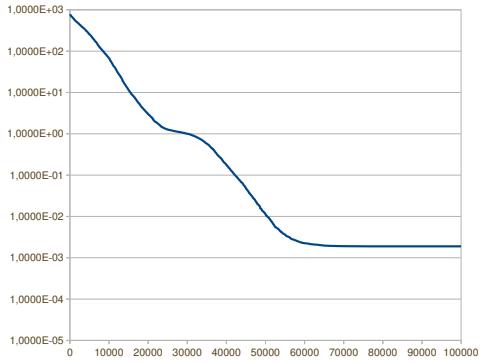


(c) Vizinhança von Neumann

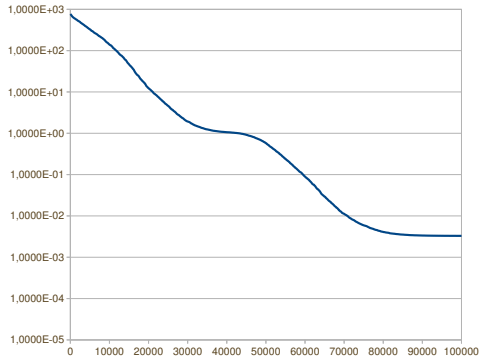
Fig. B.26: Gráficos de convergência do PSO para a função Griewank.



(a) Vizinhaça global

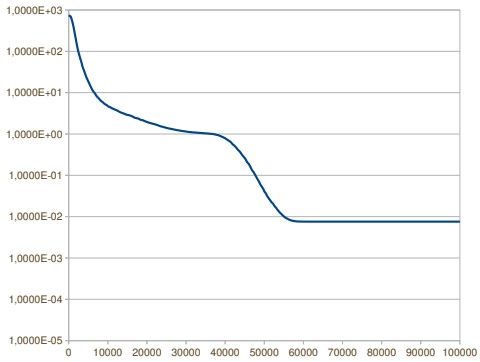


(b) Vizinhança local

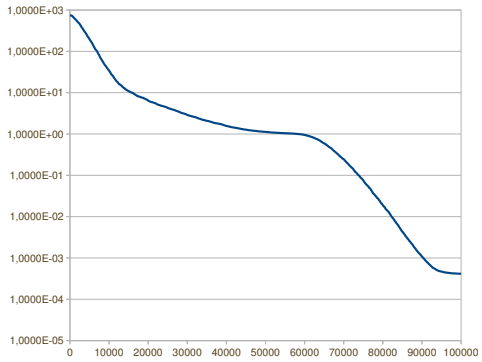


(c) Vizinhança von Neumann

Fig. B.27: Gráficos de convergência do *Bare Bones Particle Swarm* para a função Griewank.

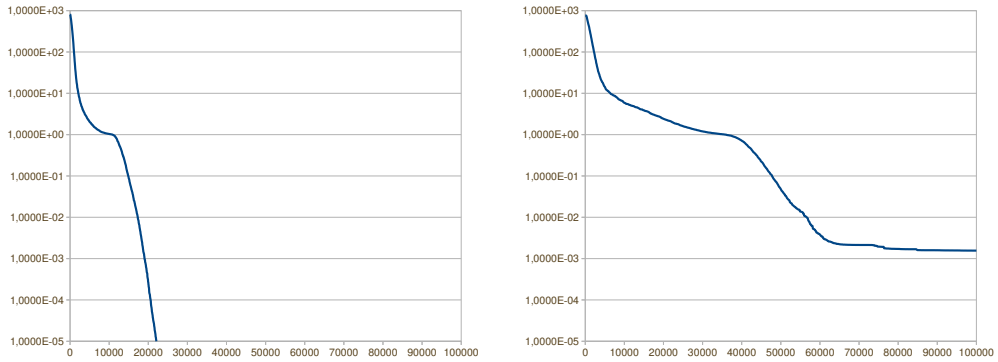


(a) Vizinhaça global



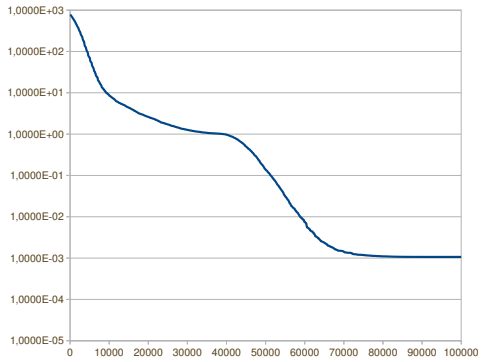
(b) Vizinhança von Neumann

Fig. B.28: Gráficos de convergência do ARPSO para a função Griewank.

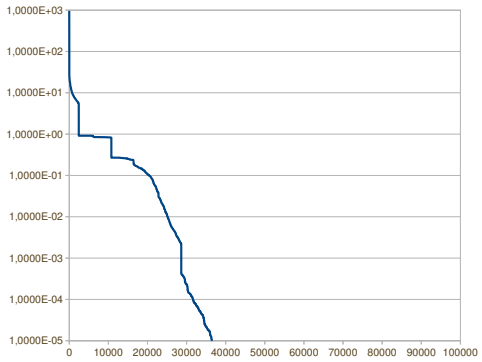


(a) FIPS com vizinhança von Neumann

(b) H-PSO

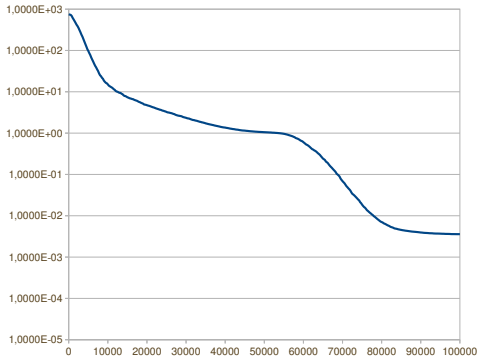


(c) DMS-PSO

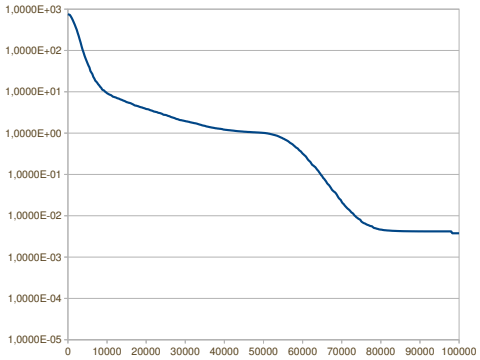


(d) TRIBES

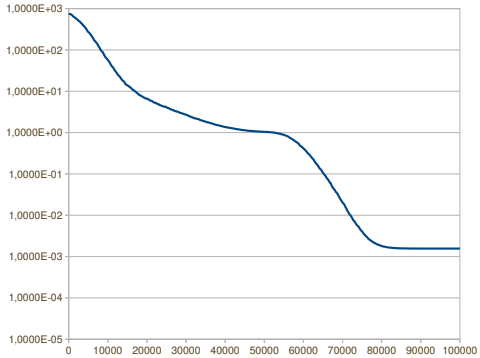
Fig. B.29: Gráficos de convergência dos demais PSOs vistos na literatura para a função Griewank.



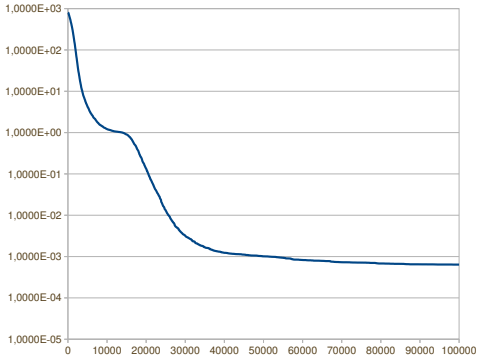
(a) CNPSO sem atualizações topológicas



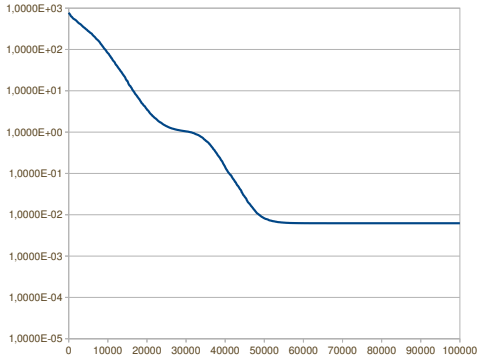
(b) CNPSO-1



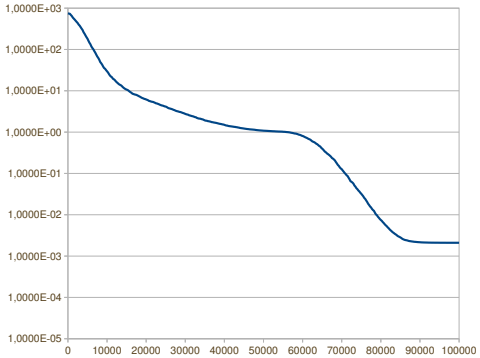
(c) CNPSO+



(d) Híbrido do FIPS com o CNPSO+



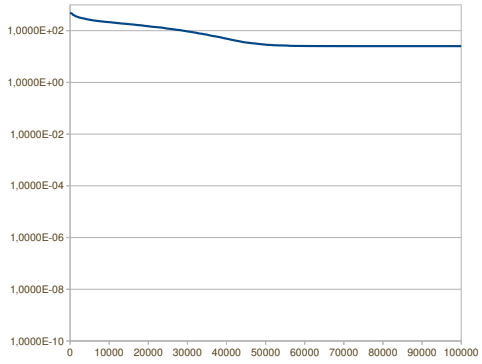
(e) Híbrido do *Bare Bones Particle Swarm* com o CNPSO+



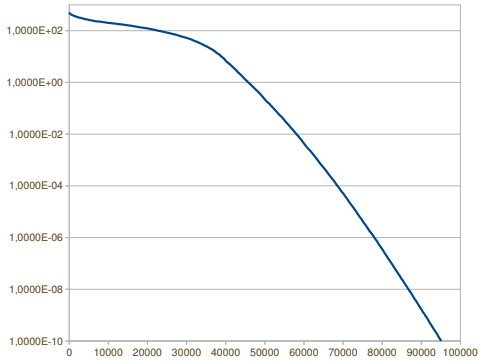
(f) Híbrido do AR-PSO com o CNPSO+

Fig. B.30: Gráficos de convergência dos CNPSOs para a função Griewank.

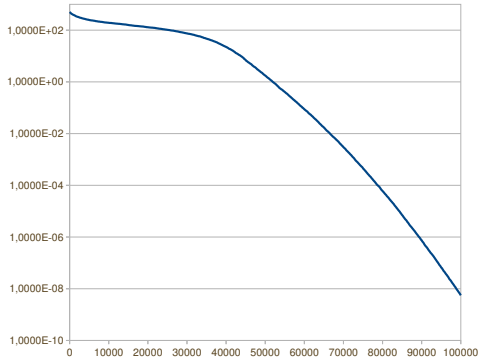
B.0.7 Função Ackley



(a) Vizinhaça global

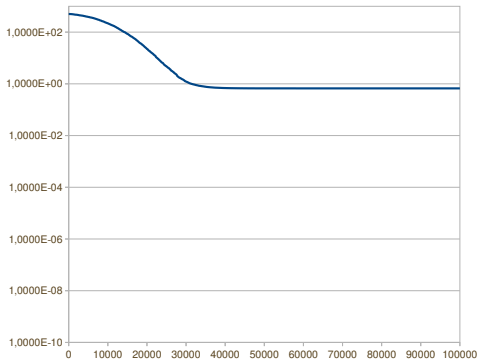


(b) Vizinhança local

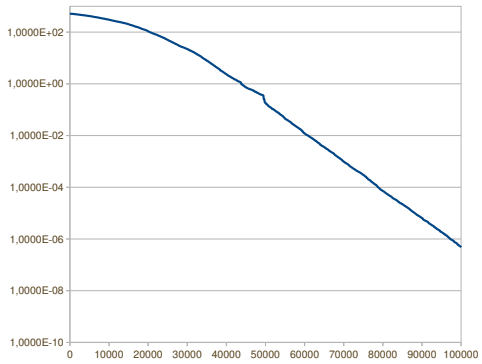


(c) Vizinhança von Neumann

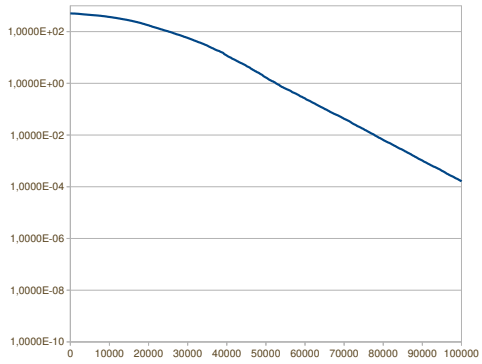
Fig. B.31: Gráficos de convergência do PSO para a função Ackley.



(a) Vizinhaça global

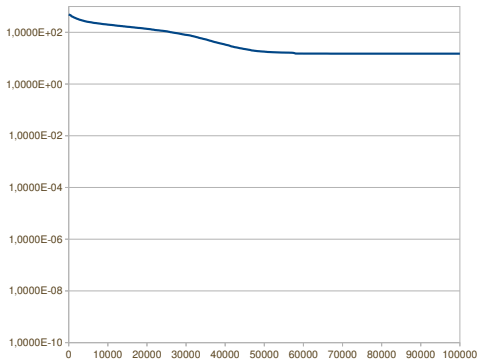


(b) Vizinhança local

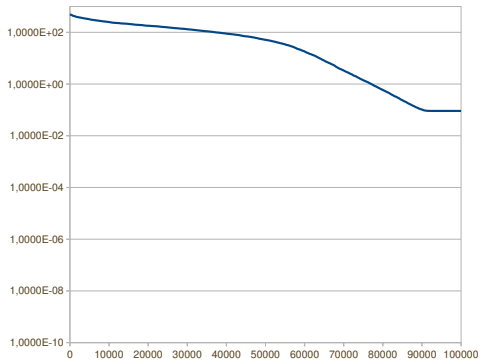


(c) Vizinhança von Neumann

Fig. B.32: Gráficos de convergência do *Bare Bones Particle Swarm* para a função Ackley.

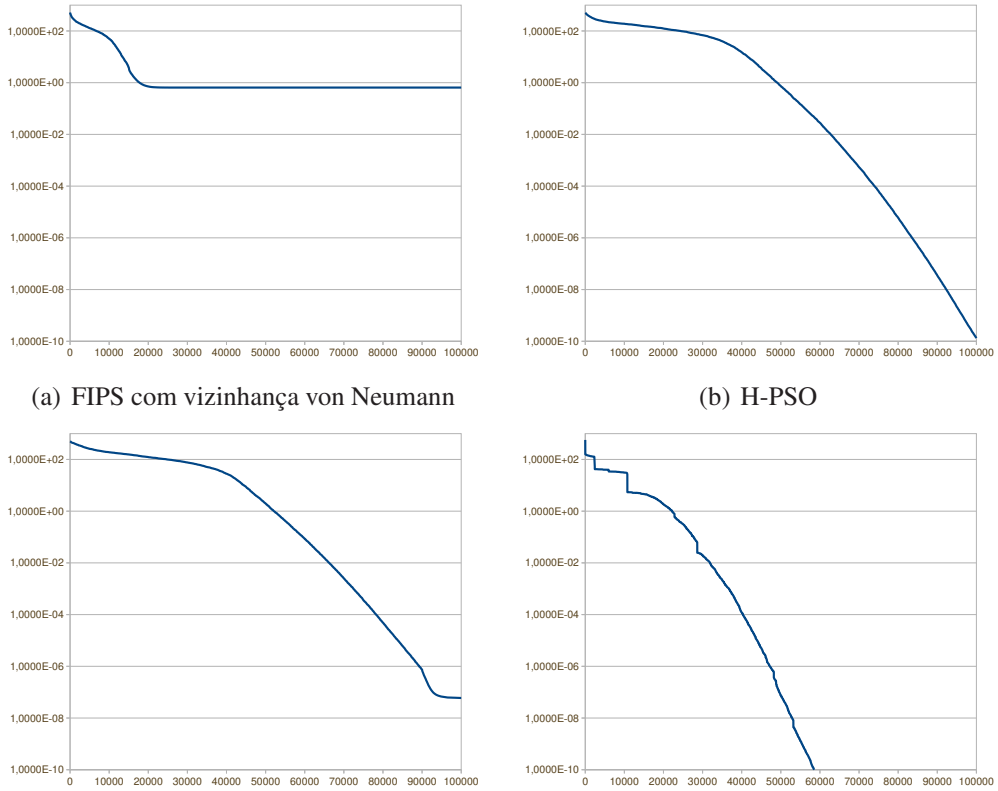


(a) Vizinhaça global



(b) Vizinhança von Neumann

Fig. B.33: Gráficos de convergência do ARPSO para a função Ackley.



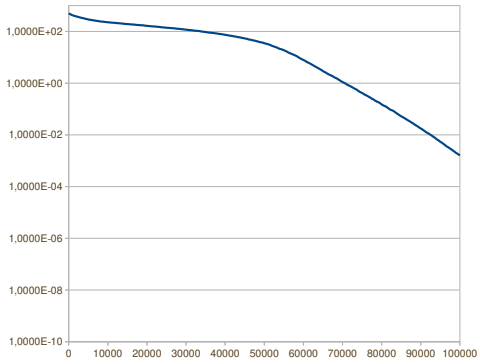
(a) FIPS com vizinhança von Neumann

(b) H-PSO

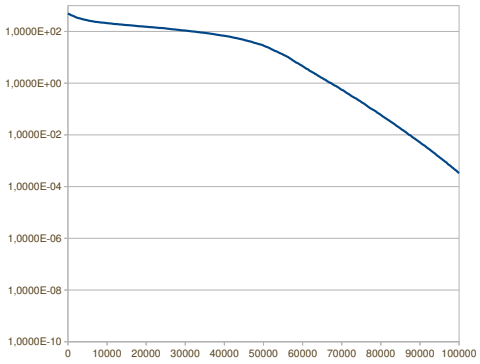
(c) DMS-PSO

(d) TRIBES

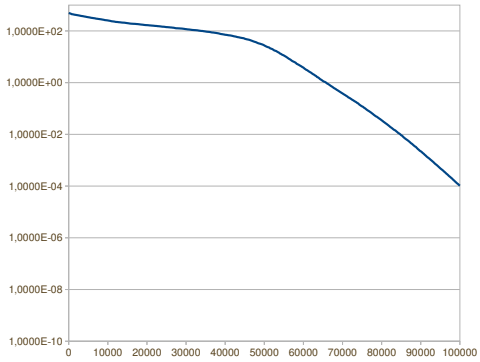
Fig. B.34: Gráficos de convergência dos demais PSOs vistos na literatura para a função Ackley.



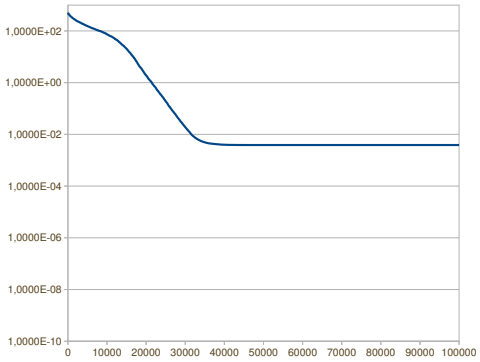
(a) CNPSO sem atualizações topológicas



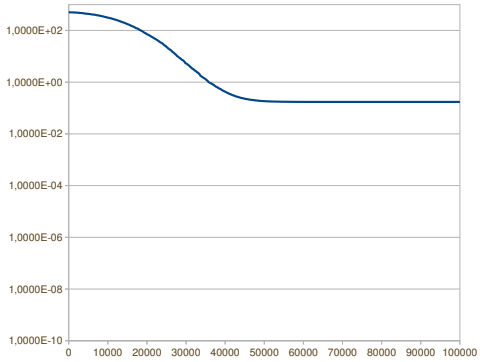
(b) CNPSO-1



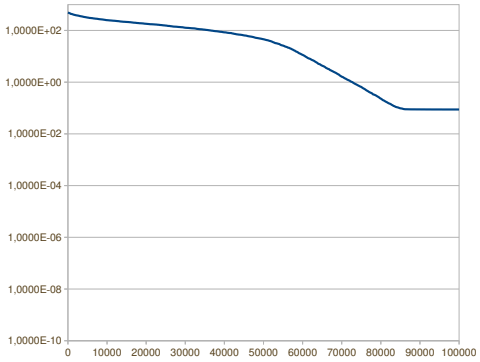
(c) CNPSO+



(d) Híbrido do FIPS com o CNPSO+



(e) Híbrido do *Bare Bones Particle Swarm* com o CNPSO+



(f) Híbrido do AR-PSO com o CNPSO+

Fig. B.35: Gráficos de convergência dos CNPSOs para a função Ackley.

Apêndice C

Valores e desvios-padrão

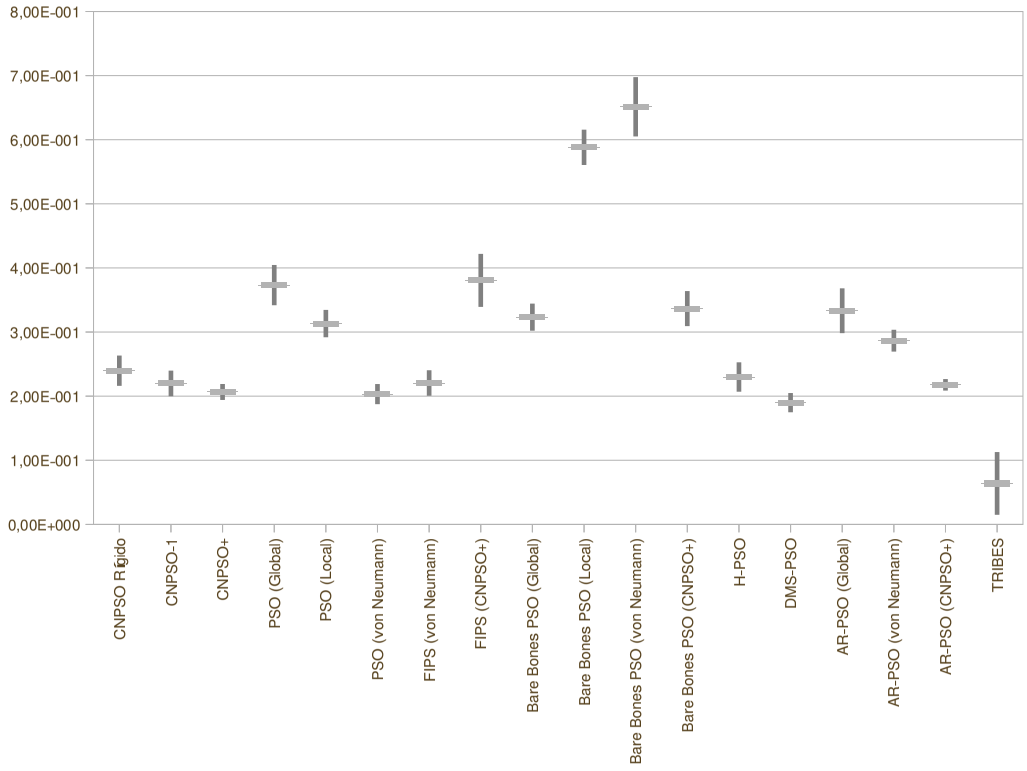


Fig. C.1: Valores e desvios-padrão obtidos por cada algoritmo para a função Salomon.

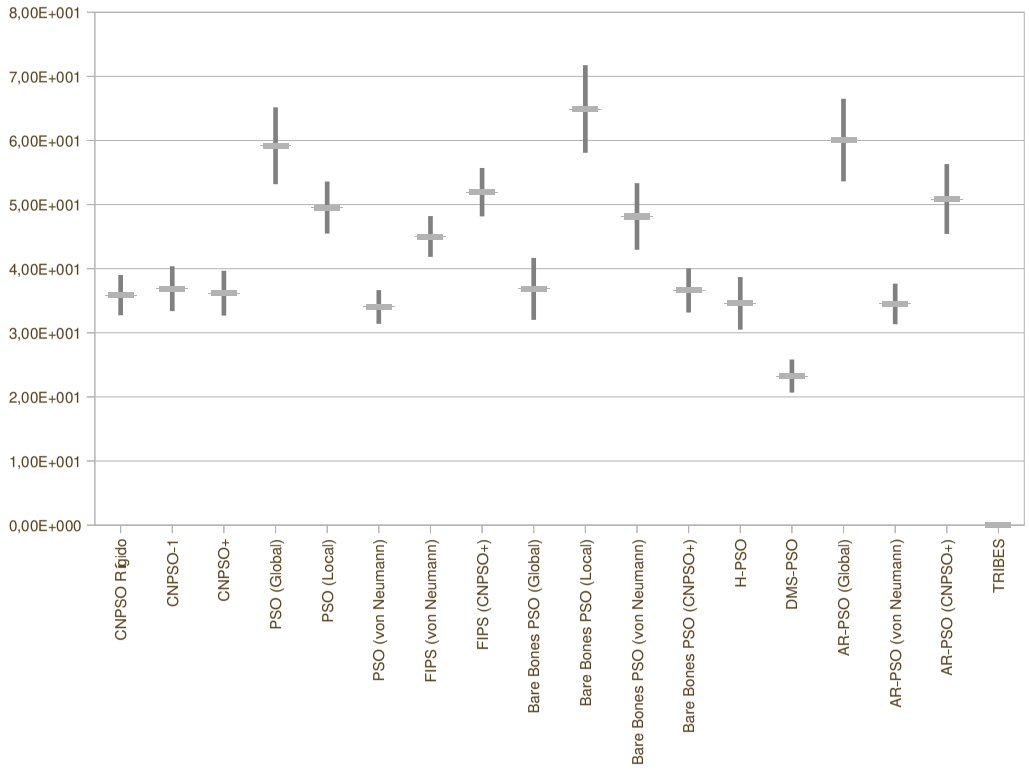


Fig. C.2: Valores e desvios-padrão obtidos por cada algoritmo para a função Rastrigin.

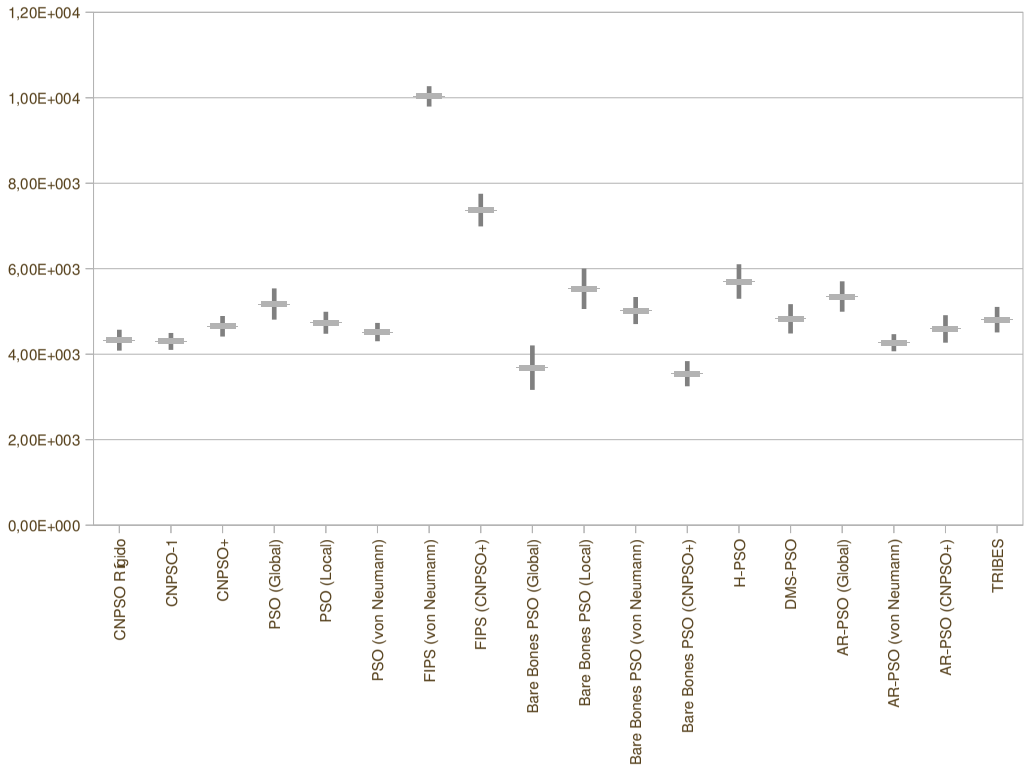


Fig. C.3: Valores e desvios-padrão obtidos por cada algoritmo para a função Schwefel.

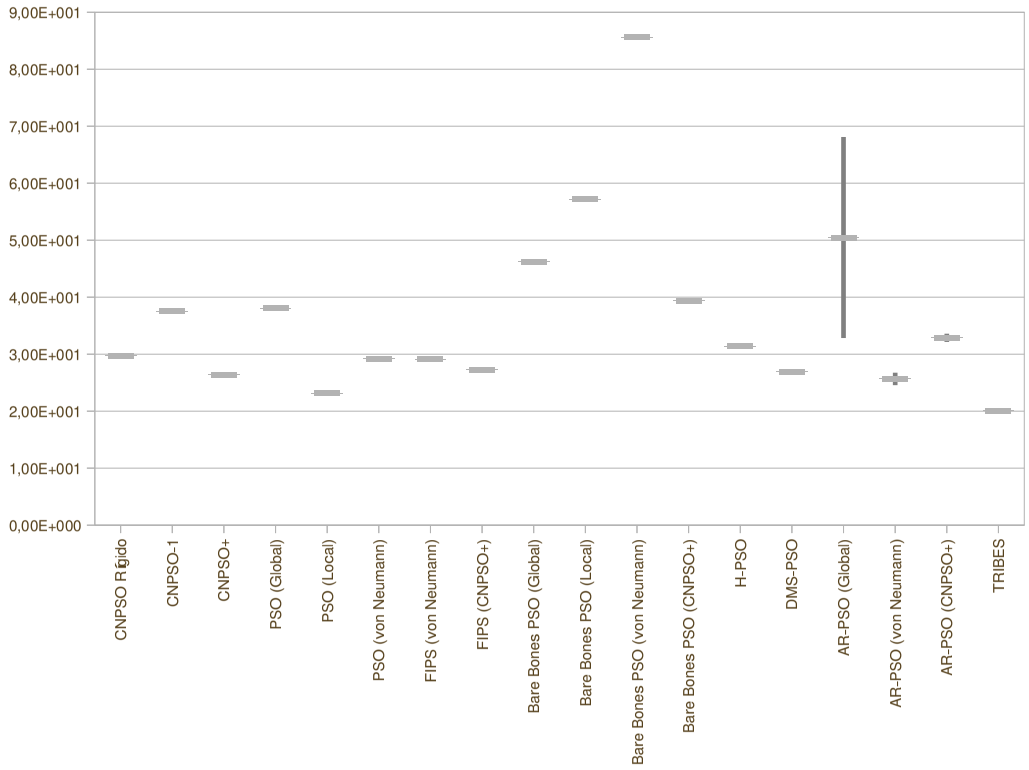


Fig. C.4: Valores e desvios-padrão obtidos por cada algoritmo para a função Rosenbrock.

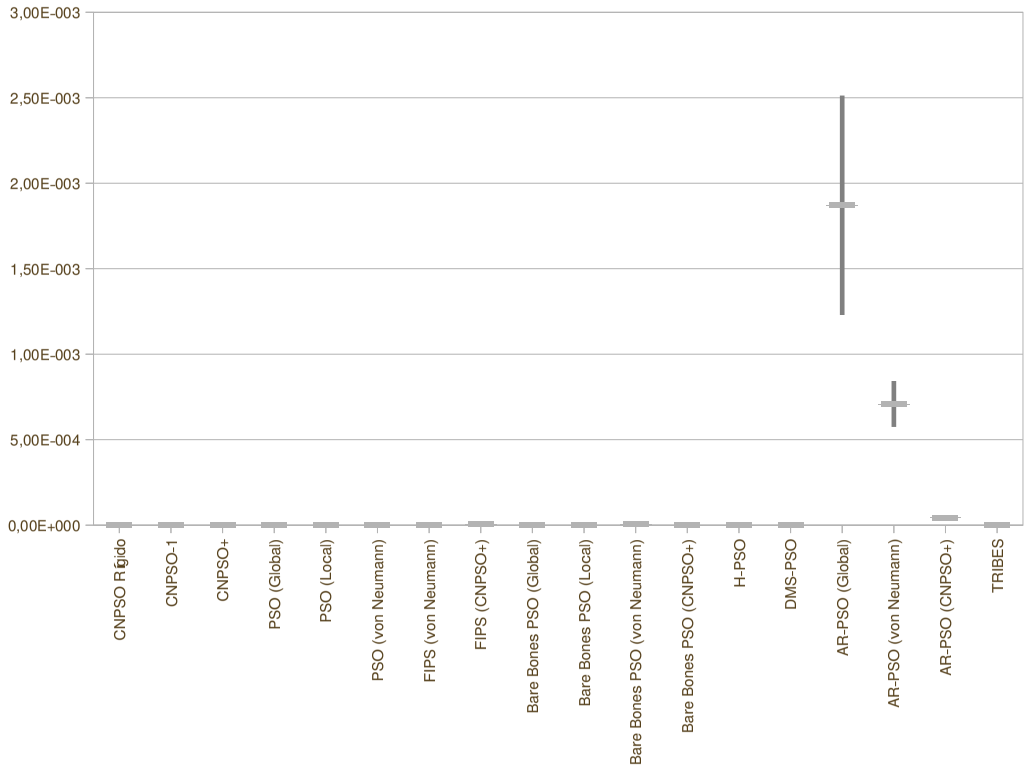


Fig. C.5: Valores e desvios-padrão obtidos por cada algoritmo para a função Sphere.

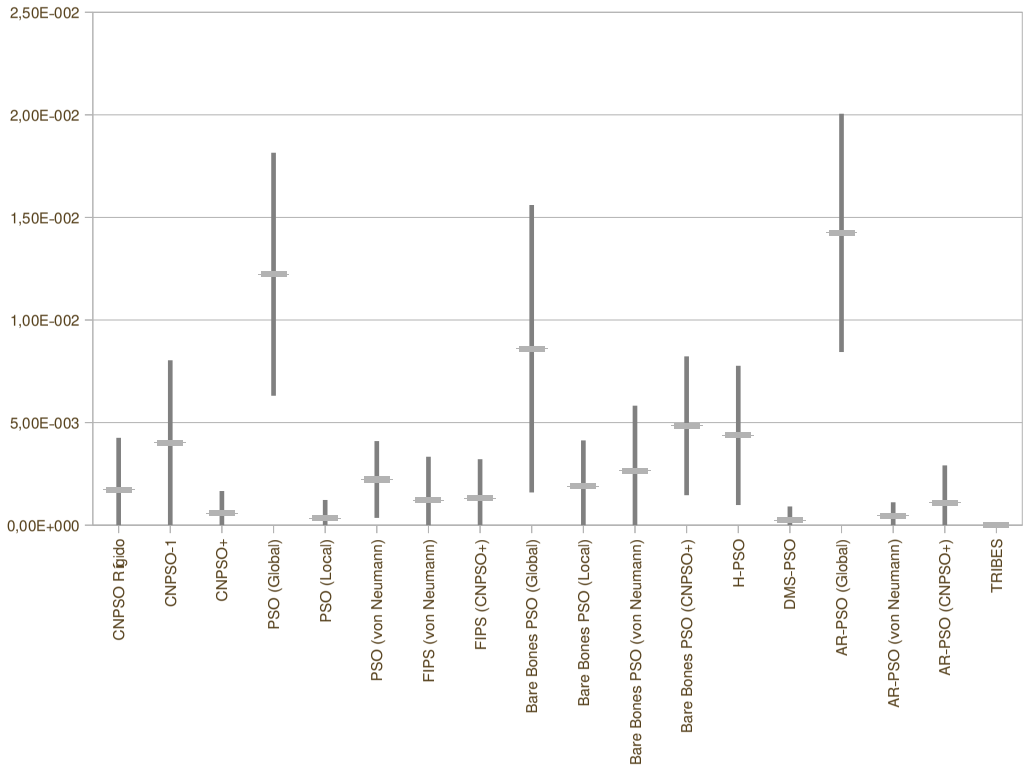


Fig. C.6: Valores e desvios-padrão obtidos por cada algoritmo para a função Griewank.

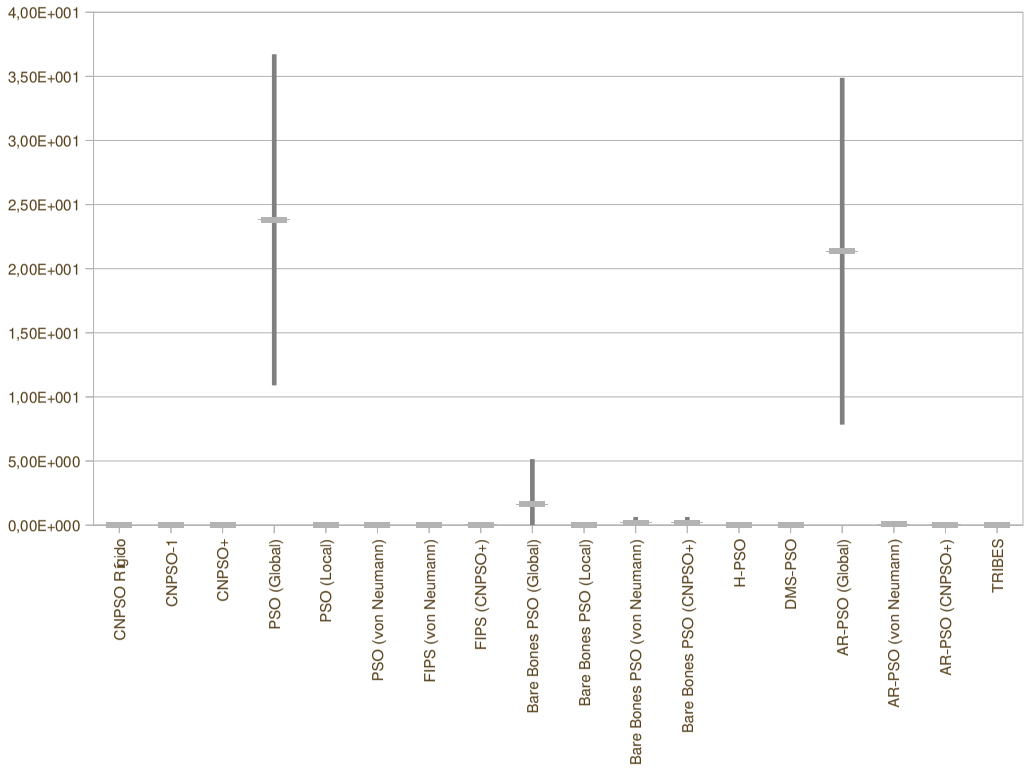


Fig. C.7: Valores e desvios-padrão obtidos por cada algoritmo para a função Ackley.