

Lianny Sánchez López

**Investigação de Técnicas Eficientes para
Algoritmos Evolutivos Multiobjetivo baseados
em Decomposição**

Belo Horizonte, MG – Brasil

Julho – 2017

Lianny Sánchez López

Investigação de Técnicas Eficientes para Algoritmos Evolutivos Multiobjetivo baseados em Decomposição

Texto submetido à banca examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais para a obtenção do título de Mestre em Engenharia Elétrica.

Universidade Federal de Minas Gerais – UFMG

Departamento de Engenharia Elétrica

Programa de Pós-Graduação em Engenharia Elétrica

Orientador: Lucas de Souza Batista

Belo Horizonte, MG – Brasil

Julho – 2017

Lianny Sánchez López

Investigação de Técnicas Eficientes para Algoritmos Evolutivos Multiobjetivo baseados em Decomposição/ Lianny Sánchez López. – Belo Horizonte, MG – Brasil, Julho – 2017-

66 p. : il. (algumas coloridas) ; 30 cm.

Orientador: Lucas de Souza Batista

Dissertação (Mestrado) – Universidade Federal de Minas Gerais – UFMG

Departamento de Engenharia Elétrica

Programa de Pós-Graduação em Engenharia Elétrica, Julho – 2017.

1. MOEA/D. 2. *Multiple-Layer Simplex-Lattice Design*. 3. Função Tchebycheff Transformada. 4. Computação Evolucionária.

Lianny Sánchez López

Investigação de Técnicas Eficientes para Algoritmos Evolutivos Multiobjetivo baseados em Decomposição

Texto submetido à banca examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais para a obtenção do título de Mestre em Engenharia Elétrica.

Lucas de Souza Batista
Orientador

Prof. PhD. Felipe Campelo

Prof. PhD. Rodney R. Saldanha

Belo Horizonte, MG – Brasil
Julho – 2017

Dedico este trabalho a todas as pessoas que acreditaram em mim, em especial a meu orientador, meus pais, irmão, esposo e meu filho amado.

Agradecimentos

Agradeço ao meu orientador, professor Lucas de Souza Batista, pela oportunidade, apoio, orientação, paciência e dedicação.

Aos meus pais e irmão por serem um exemplo de força, dedicação e amor, representando uma importante referência em minha vida. Agradeço também por terem me ensinado a ser uma pessoa melhor e a caminhar pela vida e, sobretudo, por terem me apoiado em tudo o que fiz.

Ao meu esposo e meu filhinho por serem minha inspiração, por terem me motivado a concluir este trabalho, e por caminharem ao meu lado neste longo percurso que é a vida. Obrigada filho por ser a luz de meus olhos.

Aos amigos e colegas do mestrado, pelo apoio, ajuda e suporte na realização deste trabalho.

Aos professores do Programa de Pós-Graduação em Engenharia Elétrica / UFMG pelo convívio e ensinamentos transmitidos.

Ao CNPq e CAPES pelo apoio financeiro e incentivo à pesquisa.

Resumo

Algoritmos evolucionários (EAs) baseados em decomposição apresentam grande sucesso na otimização de problemas com duas ou três funções mérito e, nos últimos anos, esse potencial tem sido investigado também no contexto de problemas com muitos objetivos. Nesse sentido, esta dissertação investiga duas abordagens promissoras para aumento do desempenho de algoritmos baseados em decomposição, a saber, (i) um modelo sistemático para a geração, uniformemente distribuída, dos vetores de ponderação (vetores referência), e (ii) uma estratégia de escalarização *Tchebycheff* ponderada transformada, a qual promove um controle simples, e independente de parâmetros, tanto da convergência quanto do espalhamento das alternativas aproximadas. Estas técnicas são incorporadas à estrutura geral do *Multiobjective Evolutionary Algorithm Based on Decomposition* (MOEA/D) e o desempenho de cada uma é avaliado frente a outras estratégias populares, considerando-se problemas teste conhecidos, a saber, DTLZ1 a DTLZ4 com 3, 5, 8, 10 e 15 objetivos. Os resultados indicam uma forte competitividade das técnicas sugeridas, principalmente em relação aos indicadores de qualidade *Inverted Generational Distance* (IGD) e *Hypervolume* (HV).

Palavras-chave: MOEA/D. *Multiple-layer simplex-lattice design*. Decomposição *Tchebycheff*. Problemas de otimização com muitos objetivos. Computação evolucionária.

Abstract

Evolutionary algorithms (EAs) based on decomposition have been successfully applied in the optimization of problems with two or three merit functions. Over the last few years, this potential has been also investigated in the context of multi-objective problems. In this sense, this dissertation investigates two promising approaches to increase the performance of algorithms based on decomposition: (i) a systematic model for generating weighted vectors (reference vectors) uniformly distributed; and (ii) a transformed weighted Tchebycheff scalarization strategy, which provides a simple and parameter-free control of both convergence and scattering of the approximated alternatives. These techniques are incorporated into the general structure of the Multiobjective Evolutionary Algorithm Based on Decomposition (MOEA/D) and the performance of each is evaluated against other well-known strategies, considering known benchmark problems, i.e., DTLZ1 to DTLZ4 with 3, 5, 8, 10 and 15 objectives. The results indicate that the proposed techniques are competitive when compared to the other approaches evaluated, mainly in relation to the quality indicators Inverted Generational Distance (IGD) and Hypervolume (HV).

Keywords: MOEA/D. Multiple-layer simplex-lattice design. Transformed Tchebycheff. Optimization problems with many objectives. Evolutionary computation.

Lista de ilustrações

Figura 1 – Ilustração da técnica de agregação Tchebycheff	29
Figura 2 – Ilustração das técnicas PBI e PBI invertida	30
Figura 3 – Ilustração da estratégia <i>multiple-layer simplex-lattice design</i>	42
Figura 4 – Ilustração da técnica de agregação Tchebycheff transformada	43
Figura 5 – Boxplot dos valores médios de d_1 e d_2 para todas as instâncias testadas	51
Figura 6 – Fronteiras estimadas para os problemas com 3 objetivos	54
Figura 7 – Fronteiras estimadas para os problemas com 5 objetivos	55
Figura 8 – Fronteiras estimadas para os problemas com 8 objetivos	56
Figura 9 – Fronteiras estimadas para os problemas com 10 objetivos	57
Figura 10 – Fronteiras estimadas para os problemas com 15 objetivos	58

Lista de tabelas

Tabela 1 – Parâmetros adotados na comparação entre SLD e MLSLD	47
Tabela 2 – Valores de HV e IGD (Melhor, Médio, Pior): SLD <i>versus</i> MLSLD	48
Tabela 3 – Parâmetros adotados na comparação entre TCH_t , TCH e PBI	48
Tabela 4 – Valores de IGD (Melhor, Médio, Pior): TCH_t <i>versus</i> TCH <i>versus</i> PBI .	52
Tabela 5 – Valores de HV (Melhor, Médio, Pior): TCH_t <i>versus</i> TCH <i>versus</i> PBI .	53

Sumário

1	INTRODUÇÃO	19
1.1	Apresentação	19
1.2	Objetivos e Contribuições	20
1.3	Organização do Trabalho	21
2	ESTRUTURAL GERAL DO MOEA/D	23
2.1	Introdução	23
2.2	Problema de Otimização Multiobjetivo	23
2.3	Algoritmos Evolucionários Multiobjetivo	24
2.4	Estrutura do MOEA/D	25
2.4.1	Vetores de Agregação	26
2.4.1.1	<i>Simplex-Lattice Design</i>	26
2.4.1.2	<i>Uniform Design</i>	27
2.4.2	Funções de Agregação Escalar	28
2.4.2.1	Soma Ponderada	28
2.4.2.2	Tchebycheff Ponderada	29
2.4.2.3	<i>Penalty-based Boundary Intersection (PBI)</i>	29
2.4.2.4	PBI Invertida (iPBI)	30
2.4.3	Estrutura de Vizinhança	31
2.4.3.1	Estrutura de Vizinhança baseada no Espaço de Objetivos	31
2.4.3.2	Estrutura de Vizinhança baseada no Espaço de Decisão	31
2.4.4	Seleção para Variação	31
2.4.4.1	Seleção Restrita	32
2.4.4.2	Seleção com Diversidade Probabilística	32
2.4.5	Operadores Evolucionários	32
2.4.5.1	Cruzamento Binário Simulado	32
2.4.5.2	Mutação Polinomial	33
2.4.5.3	Operadores de Evolução Diferencial	34
2.4.6	Estratégias de Atualização	34
2.4.6.1	Atualização de Vizinhança Padrão	34
2.4.6.2	Atualização de Vizinhança Restrita	35
2.4.6.3	Atualização baseada no Melhor Subproblema	35
2.4.6.4	Atualização baseada em Distância	35
2.4.7	Normalização do Domínio de Objetivos	35
2.4.7.1	Normalização Simples	36

2.4.7.2	<i>Corner-Sort Based Normalization</i>	36
2.4.7.3	Normalização baseada no Ponto Extremo	37
2.4.8	Estrutura Geral do MOEA/D	37
2.5	Considerações Finais	39
3	ABORDAGENS PROPOSTAS	41
3.1	Introdução	41
3.2	<i>Multiple-Layer Simplex-Lattice Design</i>	41
3.3	Função Tchebycheff Ponderada Transformada	42
3.4	Considerações Finais	44
4	RESULTADOS	45
4.1	Introdução	45
4.2	Problemas Teste	45
4.3	Indicadores de Qualidade	45
4.3.1	<i>Inverted Generational Distance (IGD)</i>	45
4.3.2	<i>Hypervolume (HV)</i>	46
4.4	Planejamento Experimental e Análise dos Resultados	46
4.4.1	Avaliação da Estratégia <i>Multiple-Layer Simplex-Lattice Design (MLS LD)</i>	46
4.4.2	Avaliação da Estratégia de Agregação Tchebycheff Transformada	47
4.5	Considerações Finais	59
5	CONCLUSÕES	61
5.1	Considerações Finais	61
5.2	Propostas de Continuidade	62
	REFERÊNCIAS	63

1 Introdução

1.1 Apresentação

Em otimização multiobjetivo (MOO), os problemas tratados envolvem funções mérito conflitantes, e a otimização simultânea das mesmas fornece um conjunto de alternativas que refletem diferentes relações de compromisso entre os critérios considerados (COELLO COELLO; LAMONT; van VELDHUIZEN, 2007). Uma vez que obter esse conjunto completo e exato é uma tarefa difícil, e principalmente questionável do ponto de vista prático (devido ao alto custo computacional e à dificuldade imposta ao processo de tomada de decisão *a posteriori*), frequentemente determina-se um conjunto finito e pequeno de alternativas que aproximem as soluções ótimas globais. Algoritmos evolucionários multiobjetivo (MOEAs) têm se mostrado adequados nesse sentido, particularmente para a resolução de problemas de otimização multiobjetivo (MOPs) com dois ou três critérios (COELLO COELLO; LAMONT; van VELDHUIZEN, 2007). Nos últimos anos, entretanto, pesquisas vêm apontando sérias dificuldades dessas abordagens para a otimização de problemas com muitos objetivos, i.e., com quatro ou mais funções mérito (PURSHOUSE; FLEMING, 2007; ISHIBUCHI; TSUKAMOTO; NOJIMA, 2008; LUCKEN; BARÁN; BRIZUELA, 2014). Como diferentes aplicações reais multiobjetivo pertencem a esse domínio (FLEMING; PURSHOUSE; LYGOE, 2005), esta dissertação investiga estratégias promissoras neste contexto.

Os principais desafios enfrentados em problemas de otimização com muitos objetivos (MaOPs) incluem os seguintes aspectos (ISHIBUCHI; TSUKAMOTO; NOJIMA, 2008; LUCKEN; BARÁN; BRIZUELA, 2014): a) À medida que o número de objetivos aumenta, torna-se crescentemente difícil a comparação e discriminação entre as soluções, o que compromete severamente a pressão seletiva. b) A manutenção e promoção de diversidade são dificultadas, resultando na convergência de MOEAs frequentemente para sub-regiões do conjunto solução de interesse. c) As alternativas tendem a apresentar-se muito distantes umas das outras, o que compromete a evolução de suas características. d) Algumas estruturas de dados, sub-rotinas e indicadores de qualidade empregados em MOO com dois ou três objetivos são impraticáveis em altas dimensões, devido ao aumento exponencial da complexidade computacional. e) A visualização das soluções estimadas não é trivial, dificultando o processo de tomada de decisão.

Com o intuito de contornar alguns dos desafios citados, diferentes abordagens foram investigadas e propostas nos últimos anos. Segundo a taxonomia apresentada em (LUCKEN; BARÁN; BRIZUELA, 2014), tais métodos podem ser classificados em dois grupos: i) algoritmos baseados em relações de preferência alternativas (FARINA; AMATO,

2002; SATO; AGUIRRE; TANAKA, 2010; BATISTA et al., 2011), e ii) algoritmos que transformam o problema original com muitos objetivos em um problema relacionado. Algoritmos desta classe são baseados em funções de escalarização (decomposição (HUGHES, 2007; ZHANG; LI, 2007) ou agregação (MURATA; TAKI, 2009)), indicadores de qualidade (BADER; ZITZLER, 2011), técnicas de redução de dimensão (SAXENA et al., 2013), ou estratégias de particionamento de espaço (AGUIRRE; TANAKA, 2009). Nesta dissertação investiga-se duas abordagens úteis para MOEAs baseados em decomposição.

1.2 Objetivos e Contribuições

Essencialmente, propõe-se (i) um gerador de vetores de ponderação (vetores referência), com potencial para viabilizar uma aproximação de soluções uniformemente distribuídas no espaço de objetivos, e (ii) uma função de agregação *Tchebycheff* cujas ponderações são transformadas de forma a alcançar um melhor mapeamento das soluções estimadas.

- A primeira estratégia é inspirada a partir da teoria de projeto e modelagem em experimentos com misturas (CHAN, 2000). Dentre os principais métodos nesta área, *simplex-lattice design* (SCHEFFÉ, 1958) é a abordagem mais comumente adotada no contexto de MOEAs baseados em decomposição. Neste trabalho estuda-se uma extensão desta estratégia, baseada em múltiplas amostras, capaz de promover uma distribuição mais uniforme de vetores de ponderação.
- A segunda abordagem difere-se da decomposição *Tchebycheff* original em dois pontos: assegura a minimização do desvio de uma solução estimada em relação ao vetor de ponderação usado; e possibilita um melhor controle sobre o espalhamento das alternativas estimadas, o que contribui diretamente para o processo de convergência.

As técnicas sugeridas são comparadas a outras abordagens populares na literatura quanto à capacidade de minimização de problemas da família DTLZ (DEB et al., 2005), envolvendo entre três e quinze funções objetivo. A análise dos resultados indica uma forte competitividade das estratégias investigadas, particularmente em relação aos indicadores de qualidade *Inverted Generational Distance* (IGD) (ZITZLER et al., 2003) e *Hypervolume* (HV) (ZITZLER; THIELE, 1999).

Além da proposta dos métodos citados anteriormente, esta dissertação contribui também com a apresentação de uma extensa revisão sobre a estrutura geral do *Multiobjective Evolutionary Algorithm based on Decomposition* (MOEA/D) (ZHANG; LI, 2007). Este algoritmo representa atualmente uma das *framework* mais populares relacionada a MOEAs baseados em decomposição.

Por fim, vale ressaltar que parte dos resultados desta dissertação estão publicados no seguinte artigo:

Lianny Sánchez López e Lucas S. Batista (2015), Uma Estratégia de Decomposição Tchebycheff Transformada para Otimização Evolucionária Multiobjetivo, XII Congresso Brasileiro de Inteligência Computacional (CBIC), Curitiba, v.1, p. 1–6.

1.3 Organização do Trabalho

Este trabalho está estruturado da seguinte forma. O capítulo 2 define formalmente o problema de otimização multiobjetivo (MOO), enumera diferentes classes de MOEAs, e apresenta a estrutura geral do MOEA/D, discutindo seus principais módulos e artigos relacionados. O capítulo 3 descreve as abordagens propostas. O capítulo 4 apresenta e discute os resultados obtidos. Finalmente, o capítulo 5 conclui o trabalho.

2 Estrutural Geral do MOEA/D

2.1 Introdução

Este capítulo apresenta a estrutura geral que compõe o MOEA/D, discute seus principais módulos e descreve algumas técnicas implementadas na literatura para cada um destes módulos. O principal intuito deste capítulo é enfatizar a modularidade do MOEA/D, como um MOEA geral baseado em decomposição, e evidenciar o desenvolvimento de inúmeros estudos dedicados a cada um desses módulos. Este capítulo está dividido em três partes, em que a primeira apresenta definições básicas relacionadas à otimização multiobjetivo (seção 2.2), a segunda enumera diferentes classes de MOEAs (seção 2.3), e a terceira discorre especificamente sobre a estrutura do MOEA/D (seção 2.4). Uma discussão mais elaborada sobre a estrutura do MOEA/D é apresentada em (CAMPELO; BATISTA; ARANHA, 2017), a qual é acompanhada por um pacote computacional para investigação e desenvolvimento de novos MOEA/Ds. O pacote é implementado na linguagem R e foi publicado no CRAN (*Comprehensive R Archive Network*), em <https://CRAN.R-project.org/package=MOEADr>. Uma versão em desenvolvimento está também disponível no repositório do projeto, <http://github.com/fcampelo/MOEADr>.

2.2 Problema de Otimização Multiobjetivo

Sem qualquer perda de generalidade, um problema de otimização multiobjetivo (MOP) pode ser definido como (MIETTINEN, 1999):

$$\begin{aligned} \min_{\mathbf{x}} \mathbf{f}(\mathbf{x}) &= [f_1(\mathbf{x}), \dots, f_{n_f}(\mathbf{x})]^T \\ \text{sujeito a: } \mathbf{x} &\in \Omega \end{aligned} \quad (2.1)$$

em que $\mathbf{x} = (x_1, \dots, x_{n_v})^T \in \mathbb{R}^{n_v}$ é uma solução candidata, $\mathbf{f}(\cdot) : \mathbb{R}^{n_v} \mapsto \mathbb{R}^{n_f}$ constitui n_f funções objetivo conflitantes, e $\Omega \subset \mathbb{R}^{n_v}$ é o espaço de decisão viável, tal que $\Omega = \{\mathbf{x} \in \mathbb{R}^{n_v} \mid g_i(\mathbf{x}) \leq 0 ; \forall i = 1, \dots, n_g\}$, em que sua imagem $\mathbf{f}(\Omega)$ define o conjunto objetivo alcançável, e $\mathbf{g}(\cdot) : \mathbb{R}^{n_v} \mapsto \mathbb{R}^{n_g}$ representa n_g restrições de desigualdade.

Uma solução viável $\mathbf{x} \in \Omega$ domina outro ponto $\mathbf{x}' \in \Omega$ (representado como $\mathbf{f}(\mathbf{x}) \prec \mathbf{f}(\mathbf{x}')$) se, e somente se, $f_k(\mathbf{x}) \leq f_k(\mathbf{x}')$ para todo $k \in \{1, \dots, n_f\}$ e $\exists k$ tal que $f_k(\mathbf{x}) < f_k(\mathbf{x}')$. Assim sendo, uma solução $\mathbf{x}^* \in \Omega$ é *Pareto-ótima* para (2.1) se não existe nenhuma outra solução $\mathbf{x} \in \Omega$ tal que $\mathbf{f}(\mathbf{x}) \prec \mathbf{f}(\mathbf{x}^*)$, e o conjunto de todas as soluções Pareto-ótimas é chamado *conjunto Pareto-ótimo*, i.e.,

$$\mathcal{PS} = \{\mathbf{x}^* \in \Omega \mid \nexists \mathbf{x} \in \Omega : \mathbf{f}(\mathbf{x}) \prec \mathbf{f}(\mathbf{x}^*)\} . \quad (2.2)$$

A imagem deste conjunto $\mathbf{f}(\mathcal{PS})$ é nomeada *fronteira Pareto-ótima*,

$$\mathcal{PF} = \{\mathbf{f}(\mathbf{x}^*) \mid \mathbf{x}^* \in \mathcal{PS}\}. \quad (2.3)$$

2.3 Algoritmos Evolucionários Multiobjetivo

Os algoritmos evolucionários (AEs) são uma alternativa aos métodos clássicos de otimização, tais como as abordagens de busca global determinísticas e técnicas de programação matemática (DEB, 2001; COELLO COELLO; LAMONT; van VELDHUIZEN, 2007). Algumas dificuldades comumente encontradas por métodos clássicos são exemplificadas a seguir (MIETTINEN, 1999):

- Requerem características específicas sobre as funções que definem o problema, tais como, diferenciabilidade, convexidade, modalidade, etc.
- Apresentam um alto nível de especificidade em relação ao problema abordado, sendo frequentemente necessário alterações consideráveis do método ou modelo para sua aplicação em problemas ligeiramente diferentes;
- Tendem a se estagnar em ótimos locais, frequentemente de baixa qualidade.

Por outro lado, existem muitas vantagens ao se utilizar as técnicas evolutivas. Por exemplo, abre-se mão de características como diferenciabilidade, convexidade, etc.; podem ser facilmente adaptadas para diferentes problemas de otimização, tanto no espaço contínuo quanto discreto; são capazes de fornecer soluções eficientes mesmo para problemas multimodais ou pertencentes à classe NP-difícil e, além disso, evoluem um conjunto de soluções candidatas em paralelo. Conforme observado na literatura, essas características favorecem o uso de algoritmos evolucionários na solução de problemas reais de otimização multiobjetivo (COELLO COELLO; LAMONT; van VELDHUIZEN, 2007).

Segundo a taxonomia apresentada em (LUCKEN; BARÁN; BRIZUELA, 2014), os MOEAs podem ser classificados em dois grupos principais: i) algoritmos baseados em relações de preferência alternativas e ii) algoritmos que transformam o problema original com muitos objetivos em um problema relacionado.

Métodos baseados em relações de preferência são definidos a partir da teoria de conjuntos clássicos (SATO; AGUIRRE; TANAKA, 2010; BATISTA et al., 2011) ou nebulosos (FARINA; AMATO, 2002). Em geral, essas abordagens comparam as soluções, uma a uma, considerando o número de objetivos nos quais uma solução é melhor do que outra (DRECHSLER; DRECHSLER; BECKER, 2001), o tamanho do ganho relativo no domínio de objetivos (FARINA; AMATO, 2002; SULFLOW; DRECHSLER; DRECHSLER, 2007) ou o número de subespaços em que uma alternativa permanece não dominada (di Pierro;

(KHU; SAVIC, 2007). Por outro lado, os métodos que transformam o problema original são baseados em funções de escalarização, indicadores de qualidade, técnicas de redução de dimensão, ou estratégias de particionamento de espaço. Funções de escalarização são empregadas por MOEAs em dois contextos, i.e., via decomposição ou agregação. Métodos baseados em decomposição dividem o problema original em um conjunto de subproblemas escalares que são otimizados simultaneamente e de maneira colaborativa (HUGHES, 2007; ZHANG; LI, 2007), enquanto aqueles definidos em termos de agregação combinam grupos de objetivos com o intuito de tratar um problema de menor dimensão (KRUISSELBRINK et al., 2009; MURATA; TAKI, 2009). Algoritmos baseados em indicadores avaliam as soluções usando uma medida de desempenho escalar, o que converte o problema original em um mono-objetivo cuja função é um indicador de qualidade (ZITZLER; KUNZLI, 2004; BADER; ZITZLER, 2011). Estratégias de redução de dimensão assumem a existência de objetivos redundantes, os quais são identificados e removidos do problema com muitos objetivos (SAXENA et al., 2013).

2.4 Estrutura do MOEA/D

A abordagem MOEA/D (ZHANG; LI, 2007) decompõe um MOP em um número finito de N subproblemas de otimização escalar, os quais são resolvidos simultaneamente usando funções de agregação. Uma vez que a solução ótima de cada subproblema é também Pareto-ótima para o MOP considerado, espera-se que as estimativas encontradas para os subproblemas representem uma aproximação razoável do conjunto Pareto-ótimo. O MOEA/D resolve os subproblemas simultaneamente via a evolução de uma população de soluções, em que somente a melhor alternativa encontrada para cada subproblema é mantida. Como soluções ótimas para dois subproblemas vizinhos podem ser muito similares (no espaço de decisão e/ou no domínio de objetivos), cada subproblema é otimizado considerando somente informações dos subproblemas vizinhos, o que promove um melhor refinamento do espaço de decisão. Entre as características que motivam a aplicação do MOEA/D, ao invés de algoritmos evolucionários multiobjetivo (MOEAs) não baseados em decomposição, é importante destacar que (i) uma vez que somente subproblemas de agregação escalar são otimizados, tanto a atribuição de aptidão (escalar) quanto a manutenção de diversidade são mais fáceis de tratar no MOEA/D do que em MOEAs não baseados em decomposição, (ii) o MOEA/D é capaz de estimar um conjunto de soluções muito bem distribuídas mesmo usando uma pequena população de subproblemas, (iii) o MOEA/D apresenta uma menor complexidade computacional por geração do que outros MOEAs não baseados em decomposição, (iv) técnicas de normalização dos objetivos podem ser incorporadas no MOEA/D para tratar funções mérito com diferentes escalas, dentre outras (ZHANG; LI, 2007; LI; ZHANG, 2009).

A estrutura do MOEA/D é composta principalmente pelos seguintes blocos:

- Definição de um conjunto de subproblemas de otimização escalar, i.e., vetores de referência, tal que seja possível determinar (ou aproximar) um conjunto bem distribuído de soluções candidatas.
- Definição de uma função de agregação escalar apropriada, que permita a estimação de soluções Pareto-ótimas.
- Atribuição de relações de vizinhança entre subproblemas, considerando o espaço de objetivos e/ou o domínio de decisão.
- Definição de uma estratégia de seleção adequada, com o intuito de promover tanto a convergência quanto a diversidade das soluções estimadas.
- Seleção de operadores evolucionários eficientes, para exploração e refinamento do espaço de busca.
- Definição de uma estratégia apropriada de atualização das soluções estimadas, também capaz de promover tanto a convergência quanto a diversidade das soluções estimadas.
- Normalização do domínio de objetivos para lidar com funções de desempenho com diferentes ordens de magnitude (escalas).

Cada um destes blocos são discutidos nas seções seguintes.

2.4.1 Vetores de Agregação

A maneira como os vetores de agregação são gerados na estrutura do MOEA/D é usualmente inspirada a partir da teoria de projeto e modelagem em experimentos com misturas (CHAN, 2000). Entre os principais métodos nesta área, *simplex-lattice design* (SCHEFFÉ, 1958) é a abordagem mais comumente adotada no contexto de MOEA/Ds. Esta estratégia e várias extensões promissoras são apresentadas a seguir.

2.4.1.1 Simplex-Lattice Design

Seja $\{\boldsymbol{\lambda}^1, \dots, \boldsymbol{\lambda}^N\}$ um conjunto de vetores de ponderação, tal que $\boldsymbol{\lambda}^i = (\lambda_1^i, \dots, \lambda_{n_f}^i)^T$, $i = 1, \dots, N$, $\lambda_j^i \geq 0$ para todo $j = 1, \dots, n_f$ e $\sum_{j=1}^{n_f} \lambda_j^i = 1$. Na abordagem *simplex-lattice design* (CHAN, 2000; ZHANG; LI, 2007; LI; ZHANG, 2009), cada vetor de ponderação é gerado tal que

$$\lambda_j^i \in \left\{ \frac{0}{H}, \frac{1}{H}, \frac{2}{H}, \dots, \frac{H}{H} \right\}, \forall i = 1, \dots, N \text{ e } j = 1, \dots, n_f \quad (2.4)$$

em que H é um inteiro positivo (definido pelo usuário) que estabelece o número de divisões consideradas ao longo de cada coordenada dos objetivos. O número total de vetores de

ponderação (igual ao tamanho da população de soluções) é calculado como

$$N = C_{n_f-1}^{H+n_f-1} \quad (2.5)$$

em que C representa a operação de combinação. Por exemplo, para um problema arbitrário com três objetivos ($n_f = 3$), um valor distinto de H produz um conjunto específico de vetores de ponderação, i.e.,

$$\begin{aligned} H = 1, \quad N = C_2^3 = 3, \quad \boldsymbol{\lambda} &\in \{(1,0,0), (0,1,0), (0,0,1)\} \\ H = 2, \quad N = C_2^4 = 6, \quad \boldsymbol{\lambda} &\in \{(1,0,0), (0,1,0), (0,0,1), (0,1/2,1/2), (1/2,0,1/2), (1/2,1/2,0)\} \\ H = 3, \quad N = C_2^5 = 10, \quad \boldsymbol{\lambda} &\in \left\{ \begin{array}{l} (1,0,0), (0,1,0), (0,0,1), (0,1/3,2/3), (1/3,0,2/3), (1/3,2/3,0) \\ (0,2/3,1/3), (2/3,0,1/3), (2/3,1/3,0), (1/3,1/3,1/3) \end{array} \right\}. \end{aligned} \quad (2.6)$$

Dessa forma, um (H, n_f) -simplex-lattice geral pode ser usado para representar $C_{n_f-1}^{H+n_f-1}$ vetores de referência (vetores de ponderação) no domínio de objetivos.

2.4.1.2 Uniform Design

O método *uniform design* (TAN et al., 2012) é basicamente desenvolvido por meio de dois passos. Inicialmente, um conjunto de vetores de discrepância reduzida $\bar{U} = \{\bar{\mathbf{u}}_i = (\bar{u}_{i1}, \dots, \bar{u}_{i(n_f-1)})^T, i = 1, \dots, N\}$ é gerado. O mecanismo *good lattice point* é frequentemente adotado para este propósito (TAN et al., 2012; TAN et al., 2013). Após isso, os vetores coeficientes $\{\boldsymbol{\lambda}^1, \dots, \boldsymbol{\lambda}^N\}$ são obtidos transformando \bar{U} , tal que $\sum_{j=1}^{n_f} \lambda_j^i = 1$ para todo $i = 1, \dots, N$.

Considerando o primeiro passo, um conjunto candidato de inteiros positivos distintos é calculado como:

$$H_N = \{h_j \mid 1 \leq h_j < N, \gcd(N, h_j) = 1, n_f - 1 < N, j = 1, \dots, n_f - 1\} \quad (2.7)$$

em que \gcd define o operador *máximo divisor comum*. Para qualquer $n_f - 1$ elementos distintos de H_N , (h_1, \dots, h_{n_f-1}) , uma matriz $U = (u_{ij}), N \times (n_f - 1)$, em que $u_{ij} = ih_j \bmod N, i = 1, \dots, N, j = 1, \dots, n_f - 1$, é gerada. Assumindo $U(N, \mathbf{h})$ uma matriz específica onde $\mathbf{h} = (h_1, \dots, h_{n_f-1})$ é um vetor gerador de U , denota-se por G_{N, n_f-1} o conjunto de todas as matrizes $U(N, \mathbf{h})$. A partir deste ponto, é possível identificar o vetor gerador $\bar{\mathbf{h}}$, com $U(N, \bar{\mathbf{h}}) \in G_{N, n_f-1}$, que representa a menor discrepância dentre o conjunto G_{N, n_f-1} . A medida de discrepância (CD_2) é a mais comumente usada na literatura (FANG; LIN, 2003), definida como:

$$\begin{aligned} CD_2(U) = & \left(\frac{13}{12}\right)^{n_f-1} - \frac{2}{N} \sum_{i=1}^N \prod_{j=1}^{n_f-1} \left(1 + \frac{1}{2} \left|u_{ij} - \frac{1}{2}\right| - \frac{1}{2} \left|u_{ij} - \frac{1}{2}\right|^2\right) \\ & + \frac{1}{N^2} \sum_{i=1}^N \sum_{k=1}^N \prod_{j=1}^{n_f-1} \left(1 + \frac{1}{2} \left|u_{ij} - \frac{1}{2}\right| + \frac{1}{2} \left|u_{kj} - \frac{1}{2}\right| - \frac{1}{2} |u_{ij} - u_{kj}|\right). \end{aligned} \quad (2.8)$$

Considerando $U(N, \bar{\mathbf{h}}) = (u_{ij}), i = 1, \dots, N$ e $j = 1, \dots, n_f - 1$, obtém-se $\bar{u}_{ij} = (u_{ij} - 0.5)/N$. Por fim, visando satisfazer a restrição $\sum_{j=1}^{n_f} \lambda_j^i = 1$ para todo $i = 1, \dots, N$, a seguinte transformação é aplicada em $\bar{U} = \{\bar{\mathbf{u}}_i = (\bar{u}_{i1}, \dots, \bar{u}_{i(n_f-1)})^T, i = 1, \dots, N\}$:

$$\begin{cases} \lambda_j^i = \left(1 - \bar{u}_{ij}^{\left(\frac{1}{n_f-j}\right)}\right) \prod_{k=1}^{j-1} \bar{u}_{ik}^{\left(\frac{1}{n_f-k}\right)} & \forall j = 1, \dots, n_f - 1 \\ \lambda_{n_f}^i = \prod_{k=1}^{n_f-1} \bar{u}_{ik}^{\left(\frac{1}{n_f-k}\right)}. \end{cases} \quad (2.9)$$

A essência do método *uniform design* é a definição de um conjunto de vetores coeficientes $\{\boldsymbol{\lambda}^1, \dots, \boldsymbol{\lambda}^N\}$ uniformemente distribuídos sobre o simplex $(n_f - 1)$ -dimensional. Esta ideia foi proposta principalmente para minimizar várias das limitações de outras estratégias de *design* existentes. Conforme apontado em (FANG; WANG, 1993), as abordagens *simplex-lattice* (SCHEFFÉ, 1958), *simplex-centroid* (SCHEFFÉ, 1963) e *axial design* (CORNELL, 1973) apresentam pelo menos as seguintes limitações: (i) os pontos de referência não são muito uniformemente distribuídos no domínio de objetivos; e (ii) existem muitos pontos de referência nas bordas do domínio de objetivos. É importante ressaltar que, considerando-se a presente revisão, não encontrou-se aplicações específicas das abordagens *simplex-centroid* e *axial design* no contexto de MOEA/Ds, provavelmente devido as suas limitações.

2.4.2 Funções de Agregação Escalar

Inúmeras estratégias de agregação podem ser aplicadas para converter um MOP em um conjunto de subproblemas de otimização escalar. As abordagens mais populares são definidas na sequência.

2.4.2.1 Soma Ponderada

Considerando $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_{n_f})^T$ um vetor de ponderação, tal que $\lambda_i \geq 0$ para todo $i = 1, \dots, n_f$ e $\sum_{i=1}^{n_f} \lambda_i = 1$, esta técnica desempenha uma combinação convexa das funções objetivo, resultando em um problema de otimização escalar (MIETTINEN, 1999):

$$\begin{aligned} \min f^{wsm}(\mathbf{x} \mid \boldsymbol{\lambda}, \mathbf{z}) &= \sum_{i=1}^{n_f} \lambda_i [f_i(\mathbf{x}) - z_i] \\ \text{sujeito a: } \mathbf{x} &\in \Omega \end{aligned} \quad (2.10)$$

em que $\mathbf{z} = (z_1, \dots, z_{n_f})^T$ é um vetor objetivo de referência, onde $z_i \leq \min\{f_i(\mathbf{x}) \mid \mathbf{x} \in \mathcal{PS}\}$ para cada $i = 1, \dots, n_f$. Um conjunto diverso de soluções Pareto-ótimas pode ser obtido usando diferentes vetores de ponderação $\boldsymbol{\lambda}$ no problema de agregação. Entretanto, devido a natureza convexa dessa definição, somente fronteiras Pareto convexas podem ser adequadamente estimadas no caso de problemas de minimização. Esta limitação pode ser

contornada pela incorporação de uma técnica ϵ -restrito (MIETTINEN, 1999), no entanto, esse tópico não será discutido nesta dissertação.

2.4.2.2 Tchebycheff Ponderada

Nesta abordagem, o problema de otimização escalar é definido como (MIETTINEN, 1999):

$$\begin{aligned} \min f^{tch}(\mathbf{x} \mid \boldsymbol{\lambda}, \mathbf{z}) &= \max_{1 \leq i \leq n_f} \{\lambda_i [f_i(\mathbf{x}) - z_i]\} \\ \text{sujeito a: } \mathbf{x} &\in \Omega . \end{aligned} \quad (2.11)$$

De maneira similar, soluções Pareto-ótimas distintas são obtidas modificando o vetor de ponderação $\boldsymbol{\lambda}$. Esta estratégia de agregação, em contraste com a soma ponderada, não é sensível à convexidade da fronteira Pareto, entretanto, ela frequentemente promove a determinação de soluções muito distantes do vetor de ponderação empregado, o que dificulta o controle de diversidade sobre o conjunto de soluções encontrado. Esta limitação é ilustrada na Fig. 1.

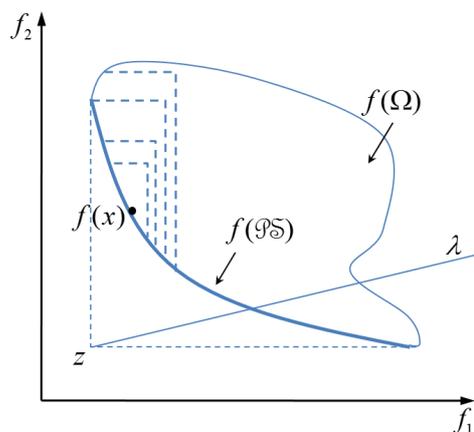


Figura 1 – Curvas de nível de um subproblema escalar e a relação geométrica entre seu vetor de ponderação e a solução ótima correspondente, considerando-se a estratégia de agregação Tchebycheff ponderada.

2.4.2.3 Penalty-based Boundary Intersection (PBI)

A função de agregação PBI (ZHANG; LI, 2007) é uma extensão da técnica *boundary intersection* (DAS; DENNIS, 1998). Essencialmente, ela visa encontrar um ponto de interseção entre o vetor referência (vetor de ponderação) e a fronteira Pareto do MOP. O problema de otimização escalar definido pela estratégia PBI é dado por:

$$\begin{aligned} \min f^{pbi}(\mathbf{x} \mid \boldsymbol{\lambda}, \mathbf{z}) &= d_1 + \theta^{pbi} d_2 \\ \text{sujeito a: } \mathbf{x} &\in \Omega \end{aligned} \quad (2.12)$$

$$d_1 = |(\mathbf{f}(\mathbf{x}) - \mathbf{z})^T \boldsymbol{\lambda}| / \|\boldsymbol{\lambda}\|$$

$$d_2 = \|\mathbf{f}(\mathbf{x}) - (\mathbf{z} + d_1 \bar{\boldsymbol{\lambda}})\| ; \bar{\boldsymbol{\lambda}} = \boldsymbol{\lambda} / \|\boldsymbol{\lambda}\|$$

em que $\theta^{pbi} \geq 0$ é um parâmetro de penalidade definido pelo usuário. A Fig. 2a ilustra as medidas d_1 e d_2 considerando um vetor objetivo $\mathbf{f}(\mathbf{x})$ e vetor de ponderação $\boldsymbol{\lambda}$. Note que d_1 permite avaliar a convergência de $\mathbf{f}(\mathbf{x})$ para a fronteira Pareto $\mathbf{f}(\mathcal{PS})$, enquanto d_2 provê uma maneira de controlar a diversidade das soluções. Por isso, esta função de agregação possibilita impor um compromisso entre convergência e diversidade das soluções obtidas, por meio de um ajuste *a priori* de θ^{pbi} , o qual influencia diretamente o desempenho da abordagem.

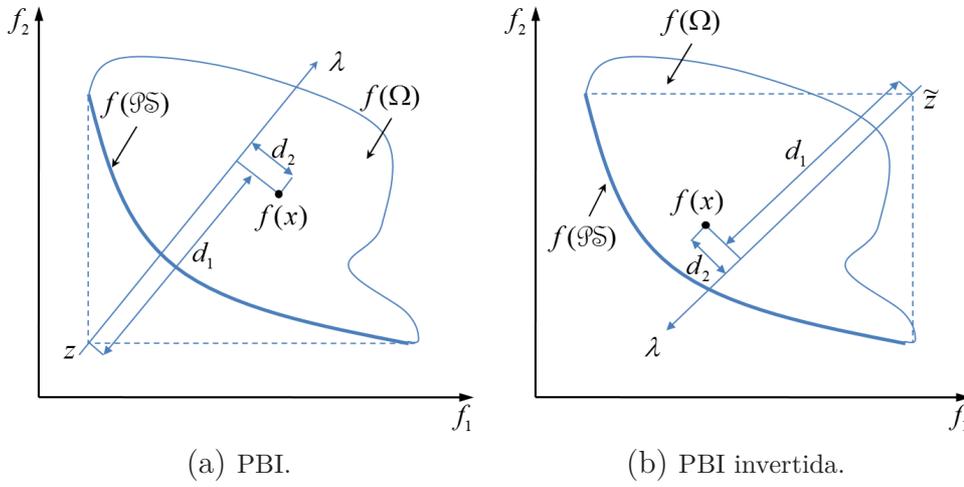


Figura 2 – Ilustração das distâncias d_1 e d_2 considerando um vetor referência arbitrário $\boldsymbol{\lambda}$ e um vetor objetivo $\mathbf{f}(\mathbf{x})$.

2.4.2.4 PBI Invertida (iPBI)

Na função PBI invertida, em contraste à estratégia PBI convencional, a busca evolui a partir de um ponto de referência *nadir* $\tilde{\mathbf{z}}$, $\tilde{z}_i \geq \max\{f_i(\mathbf{x}) \mid \mathbf{x} \in \mathcal{PS}\}$ para todo $i = 1, \dots, n_f$, por meio da maximização do problema de otimização escalar definido como (SATO, 2014):

$$\max f^{ipbi}(\mathbf{x} \mid \boldsymbol{\lambda}, \tilde{\mathbf{z}}) = d_1 - \theta^{ipbi} d_2$$

(2.13)

sujeito a: $\mathbf{x} \in \Omega$

$$d_1 = |(\mathbf{f}(\mathbf{x}) - \tilde{\mathbf{z}})^T \boldsymbol{\lambda}| / \|\boldsymbol{\lambda}\|$$

$$d_2 = \|\mathbf{f}(\mathbf{x}) - (\tilde{\mathbf{z}} - d_1 \bar{\boldsymbol{\lambda}})\| ; \bar{\boldsymbol{\lambda}} = \boldsymbol{\lambda} / \|\boldsymbol{\lambda}\|$$

o qual pode ser reescrito como um problema de minimização (d_1 e d_2 não são alterados):

$$\min f^{ipbi}(\mathbf{x} \mid \boldsymbol{\lambda}, \tilde{\mathbf{z}}) = \theta^{ipbi} d_2 - d_1$$

(2.14)

sujeito a: $\mathbf{x} \in \Omega$

em que $\theta^{ipbi} \geq 0$ é um parâmetro livre (definido pelo usuário) que controla o balanço entre d_1 e d_2 . Note, entretanto, que um d_1 elevado e um d_2 reduzido representam a situação mais desejada, em que ambos os critérios convergência e diversidade são enfatizados, respectivamente. Esta função escalar é ilustrada na Fig. 2b.

2.4.3 Estrutura de Vizinhança

No MOEA/D, uma definição apropriada das relações de vizinhança entre os subproblemas é crucial. Em geral, estas relações são usadas tanto para a seleção de soluções empregadas em operadores de variação quanto para a substituição (atualização) das soluções dos subproblemas. Baseado nessa definição, é possível controlar o compromisso entre diversidade e convergência das soluções estimadas usando diferentes tamanhos de vizinhança ou mesmo definições distintas para seleção e atualização na estrutura de vizinhança (ISHIBUCHI et al., 2009; ISHIBUCHI; AKEDO; NOJIMA, 2013). A seguir são apresentadas duas possíveis definições de vizinhança adotadas na estrutura do MOEA/D.

2.4.3.1 Estrutura de Vizinhança baseada no Espaço de Objetivos

No caso usual, a estrutura de vizinhança segue a definição padrão (ZHANG; LI, 2007; LI; ZHANG, 2009). Essencialmente, distâncias Euclidianas são calculadas entre dois vetores de ponderação quaisquer no espaço de objetivos e, para cada vetor referência $i = 1, \dots, N$, define-se a vizinhança $B(i) = \{i_1, \dots, i_T\}$, onde $\lambda^{i_1}, \dots, \lambda^{i_T}$ são os T vetores de ponderação mais próximos a λ^i .

2.4.3.2 Estrutura de Vizinhança baseada no Espaço de Decisão

Uma abordagem alternativa calcula as distâncias Euclidianas entre as soluções dos subproblemas mapeadas no domínio de decisão (CHIANG; LAI, 2011). De maneira similar, define-se $B(i) = \{i_1, \dots, i_T\}$, onde $\lambda^{i_1}, \dots, \lambda^{i_T}$ são os T vetores de ponderação cujas soluções no domínio de decisão são as mais próximas da alternativa correspondente ao vetor λ^i . Uma vez que as soluções candidatas evoluem durante a otimização dos subproblemas, as relações de vizinhança precisam ser recalculadas de tempo em tempo nesta estratégia, o que representa um custo computacional adicional com relação à abordagem anterior. Entretanto, em problemas onde soluções vizinhas no espaço de decisão não são próximas (similares) no domínio de objetivos, esta definição melhorada de vizinhança pode contribuir para o desempenho do algoritmo, o que justifica esse aumento de custo.

2.4.4 Seleção para Variação

A definição de operadores de seleção para variação representa um passo fundamental na estrutura do MOEA/D. Conforme discutido a seguir, este ajuste influencia tanto a

convergência quanto a diversidade no algoritmo.

2.4.4.1 Seleção Restrita

No MOEA/D original (ZHANG; LI, 2007), as soluções “pais” são sempre selecionadas a partir de uma vizinhança restrita $B(i)$. Basicamente, assume-se que a seleção entre soluções similares é mais útil para melhorar o desempenho do algoritmo do que a seleção entre soluções de qualidade muito diferentes. Entretanto, como efeito colateral, esta estratégia induz uma perda de diversidade, o que inviabiliza uma exploração efetiva do espaço de decisão, principalmente nos primeiros estágios da otimização (LI; ZHANG, 2009).

2.4.4.2 Seleção com Diversidade Probabilística

Com o objetivo de contornar a limitação anterior, uma estratégia melhorada permite que soluções “pais” sejam selecionadas a partir de toda a população considerando-se uma baixa probabilidade de ocorrência $1 - \delta$ (LI; ZHANG, 2009; CHIANG; LAI, 2011). Nesse sentido, as diferenças entre soluções “pais” podem promover a geração de uma grande variedade de novas soluções, melhorando dessa forma a habilidade de exploração do algoritmo e a diversidade da população. Note que esta abordagem converge para a seleção restrita fazendo $\delta = 1$.

2.4.5 Operadores Evolucionários

No MOEA/D, um operador de recombinação gera uma nova solução, a qual frequentemente é modificada por um operador de mutação. Na literatura, é muito comum a aplicação do operador cruzamento binário simulado seguido por uma mutação polinomial (ZHANG; LI, 2007; ASAFUDDOULA; RAY; SARKER, 2014; LI et al., 2014) ou mesmo os operadores do método evolução diferencial seguidos por uma mutação polinomial (LI; ZHANG, 2009; TAN et al., 2012). Estas estratégias são brevemente definidas a seguir.

2.4.5.1 Cruzamento Binário Simulado

Sejam \mathbf{x}_1 e \mathbf{x}_2 duas soluções “pais”, x_j^{min} e x_j^{max} os valores mínimo e máximo da j -ésima coordenada, respectivamente, considerando as alternativas “pais”, x_j^{lb} e x_j^{ub} os limites inferior e superior, respectivamente, da j -ésima variável de decisão, r_j e u_j dois valores aleatórios uniformemente distribuídos no intervalo $[0,1]$, e η_x um índice de distribuição do cruzamento. Uma nova solução candidata $\mathbf{c} = \{c_1, \dots, c_j, \dots, c_{n_v}\}$ gerada pela aplicação do operador cruzamento binário simulado (SBX) é obtida por (2.15)–(2.17)

(DEB; BEYER, 2001):

$$\beta_j = \begin{cases} 1 + 2 \frac{x_j^{min} - x_j^{lb}}{x_j^{max} - x_j^{min}} & \text{se } r_j \leq 0.5 \\ 1 + 2 \frac{x_j^{ub} - x_j^{max}}{x_j^{max} - x_j^{min}} & \text{caso contrário.} \end{cases} \quad (2.15)$$

$$\alpha_j = 2 - \beta_j^{-(\eta_x + 1)}$$

$$\gamma_j = \begin{cases} \frac{1}{(\alpha_j u_j)^{\eta_x + 1}} & \text{se } u_j \leq 1/\alpha_j \\ \left(2(1 - \alpha_j u_j)\right)^{-\frac{1}{\eta_x + 1}} & \text{caso contrário.} \end{cases} \quad (2.16)$$

$$c_j = \begin{cases} 0.5 \left[(x_j^{min} + x_j^{max}) - \gamma_j (x_j^{max} - x_j^{min}) \right] & \text{se } r_j \leq 0.5 \\ 0.5 \left[(x_j^{min} + x_j^{max}) + \gamma_j (x_j^{max} - x_j^{min}) \right] & \text{caso contrário.} \end{cases} \quad (2.17)$$

Como o nome sugere, o operador SBX simula o princípio de funcionamento da operação de cruzamento com corte único em uma cadeia binária. Essencialmente, a estratégia SBX possui duas propriedades principais: para um dado η_x fixo, a nova solução terá um “espalhamento” proporcional ao das soluções “pais”; e uma solução próxima às soluções “pais” é monotonicamente mais provável de ser escolhida como solução “filha” do que alternativas distantes das soluções “pais” (DEB; BEYER, 2001).

2.4.5.2 Mutação Polinomial

Seja \mathbf{x} uma solução arbitrária, x_j^{lb} e x_j^{ub} os limites inferior e superior, respectivamente, da j -ésima variável de decisão, u_j um valor aleatório uniformemente distribuído no intervalo $[0,1]$, $j = \{1, \dots, n_v\}$, e η_m um índice de distribuição da mutação. Uma solução candidata mutante \mathbf{c} gerada pela aplicação do operador mutação polinomial é obtida por (2.18) e (2.19) (DEB, 2014):

$$\begin{aligned} \text{se } u_j \leq 0.5 & \begin{cases} \beta_j = 1 - \frac{x_j - x_j^{lb}}{x_j^{ub} - x_j^{lb}} \\ \alpha_j = 2u_j + (1 - 2u_j)\beta_j^{(\eta_m + 1)} \\ \gamma_j = \alpha_j^{1/(\eta_m + 1)} - 1 \end{cases} \\ \text{caso contrário} & \begin{cases} \beta_j = 1 - \frac{x_j^{ub} - x_j}{x_j^{ub} - x_j^{lb}} \\ \alpha_j = 2(1 - u_j) + 2(u_j - 0.5)\beta_j^{(\eta_m + 1)} \\ \gamma_j = 1 - \alpha_j^{1/(\eta_m + 1)} \end{cases} \end{aligned} \quad (2.18)$$

$$c_j = x_j + \gamma_j (x_j^{ub} - x_j^{lb}) . \quad (2.19)$$

Nesta estratégia, uma distribuição de probabilidade polinomial é usada para perturbar uma solução na vizinhança das soluções “pais”, em que η_m induz um efeito de uma perturbação $\mathcal{O}((x_j^{ub} - x_j^{lb})/\eta_m)$ para todas as variáveis $j = \{1, \dots, n_v\}$.

2.4.5.3 Operadores de Evolução Diferencial

Estratégias de mutação diferencial são invariantes a qualquer rotação ortogonal de coordenadas, o que de fato explica sua grande aplicação no contexto de MOEAs baseados em decomposição (LI; ZHANG, 2009; TAN et al., 2012; LI; ZHOU; ZHANG, 2014). Assumindo x_j a j -ésima variável de decisão de uma solução arbitrária \mathbf{x} , as técnicas de mutação diferencial mais populares são definidas como (PRICE; STORN; LAMPINEN, 2005):

$$\begin{cases} DE/rand/1 : & \tilde{c}_j = x_j^{r_1} + F(x_j^{r_2} - x_j^{r_3}) \\ DE/best/1 : & \tilde{c}_j = x_j^{r_i} + F(x_j^{r_1} - x_j^{r_2}) \\ DE/best/2 : & \tilde{c}_j = x_j^{r_i} + u_j(x_j^{r_1} - x_j^{r_2}) + F(x_j^{r_3} - x_j^{r_4}) \\ DE/rand/2 : & \tilde{c}_j = x_j^{r_1} + u_j(x_j^{r_2} - x_j^{r_3}) + F(x_j^{r_4} - x_j^{r_5}) \end{cases} \quad (2.20)$$

onde a solução mutante $\tilde{\mathbf{c}}$ é sujeita a uma recombinação binomial conforme:

$$c_j = \begin{cases} \tilde{c}_j & \text{com probabilidade } CR \\ x_j^{r_i} & \text{com probabilidade } 1 - CR \end{cases} \quad (2.21)$$

em que u_j é um número aleatório uniformemente distribuído em $[0,1]$, $j = \{1, \dots, n_v\}$, \mathbf{x}^{r_i} é a solução corrente do i -ésimo subproblema, e r_i, r_1, r_2, r_3, r_4 e r_5 são índices mutuamente distintos aleatoriamente escolhidos no conjunto de soluções candidatas. A taxa de cruzamento $CR \in [0,1]$ e o fator de escala $F \in [0,1]$ são parâmetros de controle definidos pelo usuário.

2.4.6 Estratégias de Atualização

Inúmeras abordagens distintas para atualização da população foram propostas para a estrutura do MOEA/D na literatura (ZHANG; LI, 2007; ASAFUDDOULA; RAY; SARKER, 2014; WANG et al., 2014). As mais populares são apresentadas a seguir.

2.4.6.1 Atualização de Vizinhança Padrão

Esta estratégia representa o mecanismo de atualização original proposto para o MOEA/D (ZHANG; LI, 2007). Basicamente, assumindo \mathbf{x}' uma nova solução candidata gerada para o i -ésimo subproblema, ela substitui todos os seus vizinhos \mathbf{x}^j , $j \in B(i)$, em que $f^{agg}(\mathbf{x}' | \boldsymbol{\lambda}^j, \mathbf{z})$ é melhor que $f^{agg}(\mathbf{x}^j | \boldsymbol{\lambda}^j, \mathbf{z})$, definindo então $\mathbf{x}^j = \mathbf{x}'$ e $\mathbf{f}^j = \mathbf{f}(\mathbf{x}')$.

2.4.6.2 Atualização de Vizinhança Restrita

No MOEA/D original, o máximo número de soluções substituídas por uma solução “filha” pode ser tão grande quanto o tamanho da vizinhança, T , de acordo com a definição anterior. Dessa forma, uma solução “filha” de alta qualidade pode substituir a maioria das soluções correntes dos subproblemas vizinhos, promovendo assim uma redução significativa de diversidade na população. Para contornar essa limitação, o máximo número de soluções substituídas por uma solução “filha” pode ser limitado por η_r , cujo valor pode ser definido muito inferior a T (LI; ZHANG, 2009). Obviamente, esta técnica converge para a abordagem de atualização padrão quando $\eta_r = T$.

2.4.6.3 Atualização baseada no Melhor Subproblema

Nas estratégias anteriores, uma nova solução candidata \mathbf{x}' do i -ésimo subproblema é considerada para a atualização das soluções associadas aos seus subproblemas vizinhos $B(i)$. Entretanto, é incerto se a nova alternativa \mathbf{x}' é de fato a mais adequada para estes subproblemas. Com o intuito de aperfeiçoar o mecanismo de atualização, pode ser melhor identificar o subproblema j mais apropriado para \mathbf{x}' , o qual pode ser encontrado como (WANG et al., 2014):

$$j = \arg \min_{1 \leq k \leq N} f^{agg}(\mathbf{x}' | \boldsymbol{\lambda}^k, \mathbf{z}) .$$

Na sequência, a estratégia de atualização restrita é aplicada considerando a vizinhança mais apropriada $B(j)$, de tamanho T , e um número máximo de substituições, η_r .

2.4.6.4 Atualização baseada em Distância

Em (ASAFUDDOULA; RAY; SARKER, 2014), os autores propuseram uma técnica de atualização alternativa baseada nas distâncias $\{d_1, d_2\}$ ilustradas na Fig. 2a. Essencialmente, assumindo as distâncias $\{d_1^j, d_2^j\}$ para a j -ésima solução de $B(i)$, e $\{d_{1c}, d_{2c}\}$ as distâncias para uma solução “filha” do i -ésimo subproblema, esta nova alternativa candidata é considerada melhor se d_{2c} é menor do que d_2^j . Além disso, se d_{2c} for igual a d_2^j , a solução “filha” é considerada melhor se d_{1c} for menor do que d_1^j .

Neste mesmo contexto, note que a estratégia baseada no melhor subproblema, discutida anteriormente, pode ser aplicada de forma diferente considerando-se o cálculo da distância entre $\mathbf{f}(\mathbf{x}')$ e os vetores de ponderação, com o intuito de identificar o subproblema mais apropriado j para $\mathbf{f}(\mathbf{x}')$.

2.4.7 Normalização do Domínio de Objetivos

Dado que as funções objetivo de muitos MOPs estão em diferentes ordens de magnitude, e várias estratégias de agregação escalar usadas pelo MOEA/D dependem do cálculo de distâncias Euclidianas, uma normalização apropriada do domínio de objetivos

pode ser necessária para melhorar o desempenho do método. Vários esforços foram feitos nesse sentido, e alguns mecanismos populares são apresentados na sequência.

2.4.7.1 Normalização Simples

Assumindo que $\mathbf{z} = (z_1, \dots, z_{n_f})^T$ e $\tilde{\mathbf{z}} = (\tilde{z}_1, \dots, \tilde{z}_{n_f})^T$ sejam estimativas dos vetores objetivo ideal (utópico) e nadir, considerando a população corrente, uma forma simples de normalização do domínio de objetivos é através da substituição de cada função original f_j por (MIETTINEN, 1999; ZHANG; LI, 2007):

$$\bar{f}_j(\mathbf{x}) = \frac{f_j(\mathbf{x}) - z_j}{\tilde{z}_j - z_j}, \quad \forall j = 1, \dots, n_f \quad (2.22)$$

garantindo assim que a escala de cada objetivo pertença a $[0,1]$.

2.4.7.2 Corner-Sort Based Normalization

Baseando-se no *corner-sort ranking procedure* (SINGH; ISAACS; RAY, 2011), a normalização das funções desempenho é obtida a partir de $2n_f$ vetores objetivo de referência não-dominados, em que os primeiros n_f correspondem ao mínimo de cada objetivo, enquanto os seguintes n_f correspondem ao mínimo das normas Euclidianas de $n_f - 1$ objetivos. Dessa forma, para um problema de minimização tri-objetivo arbitrário, estes $2n_f$ vetores são aqueles associados ao mínimo de cada função no conjunto $\{f_1, f_2, f_3, f_2^2 + f_3^2, f_1^2 + f_3^2, f_1^2 + f_2^2\}$. Uma vez que vetores idênticos devem ser evitados, se uma solução já tiver sido selecionada, então a próxima melhor alternativa (de acordo com esta função) deve ser escolhida.

Para este conjunto de $2n_f$ soluções, o máximo de cada objetivo é identificado e as soluções correspondentes são selecionadas e referidas como pontos extremos. Estes são usados para criar um hiperplano e para calcular suas interseções com os eixos dos objetivos. Considerando que as interseções do hiperplano com os eixos dos objetivos são indicados por a_1, a_2, \dots, a_{n_f} , o sistema geral de um hiperplano definido por estes pontos pode ser representado como:

$$c_1 f_1 + c_2 f_2 + \dots + c_{n_f} f_{n_f} = 1 \quad (2.23)$$

onde c_1, c_2, \dots, c_{n_f} são a norma unitária do plano, e as interseções do plano com os eixos são dadas por $a_1 = 1/c_1, a_2 = 1/c_2, \dots, a_{n_f} = 1/c_{n_f}$.

É importante ter em mente que a minimização da norma Euclidiana minimizará os valores dos objetivos somente se estes forem positivos. Se alguma função objetivo pode assumir valores negativos, então uma quantidade positiva (constante) pode ser adicionada a cada objetivo com o intuito de assegurar somente valores positivos. Além disso, caso o número de pontos extremos seja menor que n_f , o máximo valor do objetivo (f_j^{max}) é

usado como o valor de interseção. Cada solução na população é então escalada conforme:

$$\bar{f}_j(\mathbf{x}) = \frac{f_j(\mathbf{x}) - z_j}{a_j - z_j}, \quad \forall j = 1, \dots, n_f \quad (2.24)$$

em que $\mathbf{z} = (z_1, \dots, z_{n_f})^T$ representa o ponto ideal da população corrente.

2.4.7.3 Normalização baseada no Ponto Extremo

Esta estratégia de normalização é dividida em dois passos (DEB; JAIN, 2014). Primeiramente, cada função objetivo f_j é modificada para obter $f'_j(\mathbf{x}) = f_j(\mathbf{x}) - z_j$, para todo $j = 1, \dots, n_f$, em que $\mathbf{z} = (z_1, \dots, z_{n_f})^T$ representa uma estimativa do ponto ideal considerando a população corrente. Feito isto, um ponto extremo $\tilde{\mathbf{z}}_j$ para cada coordenada objetivo $j = 1, \dots, n_f$ é identificada pela determinação da solução \mathbf{x}_j que otimiza a função de agregação adotada formada com $f'_j(\mathbf{x})$ e um vetor de ponderação próximo à j -ésima coordenada objetivo, i.e.,

$$\mathbf{x}_j = \min_{\mathbf{x}} f^{agg}(f'_j(\mathbf{x}) \mid \boldsymbol{\lambda}_j) \quad \text{e} \quad \tilde{\mathbf{z}}_j = \mathbf{f}'(\mathbf{x}_j) .$$

Este conjunto de n_f vetores objetivo extremos são então usados para compor um hiperplano n_f -dimensional por meio da resolução do mesmo sistema em (2.23). As funções objetivo são então normalizadas como:

$$\bar{f}_j(\mathbf{x}) = \frac{f'_j(\mathbf{x})}{a_j}, \quad \forall j = 1, \dots, n_f . \quad (2.25)$$

2.4.8 Estrutura Geral do MOEA/D

A estrutura geral do MOEA/D é apresentada no Algoritmo 1. De acordo com as definições anteriores, seja N o número de subproblemas considerados no MOEA/D, T o tamanho da vizinhança, δ a probabilidade das soluções “pais” serem selecionadas a partir da vizinhança, e η_r o máximo número de soluções substituídas por cada solução “filha”. Além disso, \mathbf{z} e $\tilde{\mathbf{z}}$ são definidos como os pontos ideal e nadir atuais, respectivamente, obtidos a partir da população de soluções corrente. O uso de um arquivo externo (linha 1), uma heurística de refinamento (linha 11), e uma estratégia de reparo (linha 12), embora frequentemente muito úteis para aumentar o desempenho do algoritmo, são opcionais, e não são abordadas nessa dissertação.

Algoritmo 1 Estrutura Geral do MOEA/D

Entrada funções objetivo $\mathbf{f}(\cdot)$, restrições $\mathbf{g}(\cdot)$, número de vetores de ponderação N , tamanho da vizinhança T , probabilidade δ , número de soluções substituídas η_r .

▷ **Passo 1: Inicialização**

- 1: Defina a população externa $EP = \emptyset$.
- 2: Gere os vetores de ponderação $\lambda^1, \dots, \lambda^N$ de acordo com uma técnica específica.
- 3: Gere uma população inicial $\mathbf{x}^1, \dots, \mathbf{x}^N$ aleatoriamente em Ω ou usando um método específico. Defina $\mathbf{f}^i = \mathbf{f}(\mathbf{x}^i)$.
- 4: Defina a vizinhança $B(i) = \{i_1, \dots, i_T\}$ para cada $i = 1, \dots, N$, em que $\lambda^{i_1}, \dots, \lambda^{i_T}$ são os T vetores de ponderação vizinhos a λ^i .
- 5: Atribua cada solução \mathbf{x}^i a uma única direção de referência λ^i aleatoriamente.
- 6: Inicialize $\mathbf{z} = (z_1, \dots, z_{n_f})^T$ fazendo $z_j = \min_{1 \leq i \leq N} f_j(\mathbf{x}^i)$.
- 7: Inicialize $\tilde{\mathbf{z}} = (\tilde{z}_1, \dots, \tilde{z}_{n_f})^T$ fazendo $\tilde{z}_j = \max_{1 \leq i \leq N} f_j(\mathbf{x}^i)$.

▷ **Passo 2: Atualização**

- 8: **Para** $i = 1, \dots, N$ **faça**
- 9: Gere um número aleatório $rand$ em $[0,1]$, e então defina ▷ vizinhança

$$P = \begin{cases} B(i) & \text{se } rand < \delta \\ \{1, \dots, N\} & \text{caso contrário.} \end{cases}$$

- 10: Selecione aleatoriamente um subconjunto de índices $I \subseteq P$, e gere uma nova solução \mathbf{y} a partir de \mathbf{x}^k , $k \in I$, usando operadores evolutivos. ▷ reprodução
- 11: Aplique uma heurística de refinamento em \mathbf{y} para produzir \mathbf{y}' . ▷ refinamento
- 12: Repare a solução \mathbf{y}' se ela violar os limites de Ω . ▷ reparo
- 13: Para cada $j = 1, \dots, n_f$, se $z_j > f_j(\mathbf{y}')$, então defina $z_j = f_j(\mathbf{y}')$. ▷ atualização de \mathbf{z}
- 14: Para cada $j = 1, \dots, n_f$, se $\tilde{z}_j < f_j(\mathbf{y}')$, então defina $\tilde{z}_j = f_j(\mathbf{y}')$. ▷ atualização de $\tilde{\mathbf{z}}$
- 15: Defina $r = 0$ e faça: ▷ substituição
 - (1) Se $r = \eta_r$ ou P estiver vazio, então **continue**. Caso contrário, escolha um índice apropriado $j \in P$.
 - (2) Se $f^{agg}(\mathbf{y}' | \lambda^j, \mathbf{z}) \leq f^{agg}(\mathbf{x}^j | \lambda^j, \mathbf{z})$, defina $\mathbf{x}^j = \mathbf{y}'$, $\mathbf{f}^j = \mathbf{f}(\mathbf{y}')$ e $r = r + 1$.
 - (3) Remova j de P e retorne para (1).
- 16: Atualize a vizinhança $B(i) = \{i_1, \dots, i_T\}$ para cada $i = 1, \dots, N$. ▷ atualização de B
- 17: Atualize a população externa EP . ▷ atualização de EP
- 18: **Fim Para**

▷ **Passo 3: Critérios de Parada**

- 19: Se os critérios de parada são satisfeitos, termine o algoritmo e retorne $\{\mathbf{x}^1, \dots, \mathbf{x}^N\}$ e $\{\mathbf{f}(\mathbf{x}^1), \dots, \mathbf{f}(\mathbf{x}^N)\}$. Caso contrário, retorne para **Passo 2**.

2.5 Considerações Finais

Este capítulo enfatizou a grande flexibilidade existente na estrutura do MOEA/D e abordou inúmeros trabalhos correlatos desenvolvidos na área. No capítulo seguinte apresenta-se duas novas abordagens promissoras para o aumento de desempenho do MOEA/D, sendo a primeira associada ao grupo de estratégias geradoras de vetores de ponderação e a segunda uma função de agregação escalar.

3 Abordagens Propostas

3.1 Introdução

Conforme evidenciado no capítulo anterior, o MOEA/D representa uma estrutura geral de MOEAs baseados em decomposição, com grande flexibilidade associada a cada um dos seus módulos. Em função disso, percebe-se claramente o interesse da comunidade científica envolvida com Computação Evolucionária em desenvolver novas estratégias no âmbito dessa metodologia. Nesse contexto, a autora desta dissertação investiga uma nova metodologia para uma geração mais uniformemente distribuída dos vetores de ponderação (seção 3.2) e também uma estratégia de agregação Tchebycheff transformada, com potencial para promover resultados mais interessantes quando comparada a outras técnicas presentes na literatura (seção 3.3).

3.2 *Multiple-Layer Simplex-Lattice Design*

De acordo com Deb e Jain (2014), com o intuito de obter vetores de ponderação intermediários no interior do simplex n_f -dimensional usando o *simplex-lattice design*, é preciso considerar $H \geq n_f$. Entretanto, em um domínio objetivo de alta dimensão, uma grande quantidade de vetores de ponderação é gerada mesmo quando $H = n_f$. Por exemplo, para um problema com sete objetivos, $H = n_f = 7$ resulta em $C_6^{13} = 1716$ vetores de ponderação, o que obviamente agrava o custo computacional do MOEA/D. Por outro lado, se esta limitação é remediada pela redução de H (e.g., ajustando $H < n_f$), todos os vetores de ponderação ficarão distribuídos ao longo da borda do simplex, prejudicando assim a diversidade das soluções.

Para evitar essa situação, pode-se considerar um método gerador de vetores de ponderação de múltiplas camadas. Nesse sentido, esta estratégia gera k camadas de pontos usando valores de H não necessariamente distintos. Cada camada pode ser gerada preferencialmente a partir de valores $H < n_f$, o que é particularmente interessante quando $n_f > 3$. Entretanto, em baixas dimensões do domínio de objetivos, i.e., $n_f \leq 3$, valores $H \geq n_f$ continuam praticáveis. Nesta abordagem, cada conjunto (camada) $L_i = \{\lambda_1^i, \dots, \lambda_{N_i}^i\}$, $i = 1, \dots, k$, é usada para compor o conjunto total de $N = N_1 + \dots + N_k$ vetores de ponderação. Para assegurar uma melhor diversidade entre as soluções, as coordenadas dos vetores gerados são contraídas no interior do simplex por meio de uma transformação de coordenadas. Essencialmente, supondo um vetor $\lambda^i = (\lambda_1^i, \dots, \lambda_{n_f}^i)$, seu j -ésimo

componente é mapeado como:

$$\lambda_j^i = \tau^i \lambda_j^i + (1 - \tau^i)/n_f \quad (3.1)$$

em que $j \in \{1, \dots, n_f\}$ e $\tau^i \in [0,1]$ é um fator de contração. Cada conjunto L_i precisa ser associado a um fator distinto τ^i , em que $\tau^i = 1$ mantém o sistema de coordenadas original e $\tau^i \rightarrow 0$ promove uma contração máxima. A Fig. 3 ilustra um conjunto de vetores de ponderação obtidos para um problema de minimização tri-objetivo, considerando-se $k = 3$ camadas e $H_1 = 2$, $H_2 = 1$ e $H_3 = 2$.

Esta abordagem pode ser vista como uma extensão da estratégia proposta em (LI et al., 2014), onde um *simplex-lattice design* de duas camadas é considerado. A generalização para k camadas é intuitiva e visivelmente pode melhorar a distribuição dos vetores de ponderação no domínio de objetivos (ver Fig. 3).

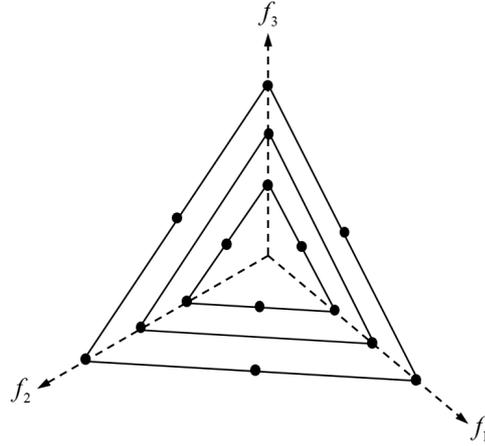


Figura 3 – Ilustração de um conjunto de vetores de ponderação obtidos para um problema tri-objetivo, considerando-se $k = 3$ camadas e $H_1 = 2$, $H_2 = 1$ e $H_3 = 2$.

3.3 Função Tchebycheff Ponderada Transformada

A técnica de decomposição proposta representa uma função Tchebycheff ponderada transformada, definida como:

$$\begin{aligned} \min f^{tchz}(\mathbf{x} \mid \boldsymbol{\rho}, \mathbf{z}) &= \max_{1 \leq i \leq n_f} \{\rho_i [f_i(\mathbf{x}) - z_i]\} \\ \text{sujeito a: } \mathbf{x} &\in \Omega \end{aligned} \quad (3.2)$$

em que $\mathbf{z} = (z_1, \dots, z_{n_f})^T$ é um vetor objetivo de referência, onde $z_i \leq \min\{f_i(\mathbf{x}) \mid \mathbf{x} \in \mathcal{PS}\}$ para todo $i = 1, \dots, n_f$, e $\boldsymbol{\rho} = (\rho_1, \dots, \rho_{n_f})^T$ é tal que

$$\rho_i = w_i / \sum_{i=1}^{n_f} w_i \quad \text{e} \quad w_i = 1/\lambda_i.$$

Note que o vetor de referência λ , empregado pela função Tchebycheff ponderada usual, é substituído pela normalização do seu inverso, i.e., por ρ , e, como efeito, além de continuar assegurando a minimização do máximo desvio $|f_i(\mathbf{x}) - z_i|$, $i = 1, \dots, n_f$, também promove a minimização da distância entre um vetor objetivo $\mathbf{f}(\mathbf{x})$ e seu vetor de ponderação correspondente ρ .

A Fig. 4 ilustra os contornos das funções escalares f^{tch} e f^{tcht} considerando um vetor de referência $\lambda \approx (0.7, 0.3)$. A partir da Fig. 4a, observa-se que o cone-negativo gerado pela função f^{tch} é deslocado em relação ao vetor de ponderação, apresentando sua origem sobre esse vetor somente quando $\lambda_i = 0.5$ para todo $i = \{1, 2\}$. Este fato dificulta o controle sobre a convergência e diversidade das soluções aproximadas, sendo possivelmente mais crítico em problemas com muitos objetivos. Por outro lado, a função Tchebycheff transformada f^{tcht} reposiciona a origem do cone-negativo sobre o vetor de referência, para qualquer ponderação considerada (veja a Fig. 4b), contornando a dificuldade ressaltada. Dessa forma, essa estratégia desempenha um papel similar ao da abordagem PBI, porém sem a necessidade de se definir parâmetros adicionais. De fato, a definição do cone-negativo sobre o vetor de ponderação ρ promove tanto a convergência da solução quanto a sua aproximação deste vetor referência, o que também é observado ao se minimizar as distâncias d_1 e d_2 na estratégia PBI, respectivamente. Vale destacar que a decomposição f^{pbi} converge para a função f^{tcht} quando $\theta = 1.0$ e $\lambda_i = \lambda \forall i$, i.e., somente para ponderações idênticas. Para $\theta > 1$ a função f^{pbi} possibilitará a estimação de soluções mais próximas ao vetor referência, entretanto, como a pressão seletiva será maior, i.e., o cone-negativo será mais fechado, poderá apresentar um desempenho degradado em problemas onde a promoção de diversidade é crítica.

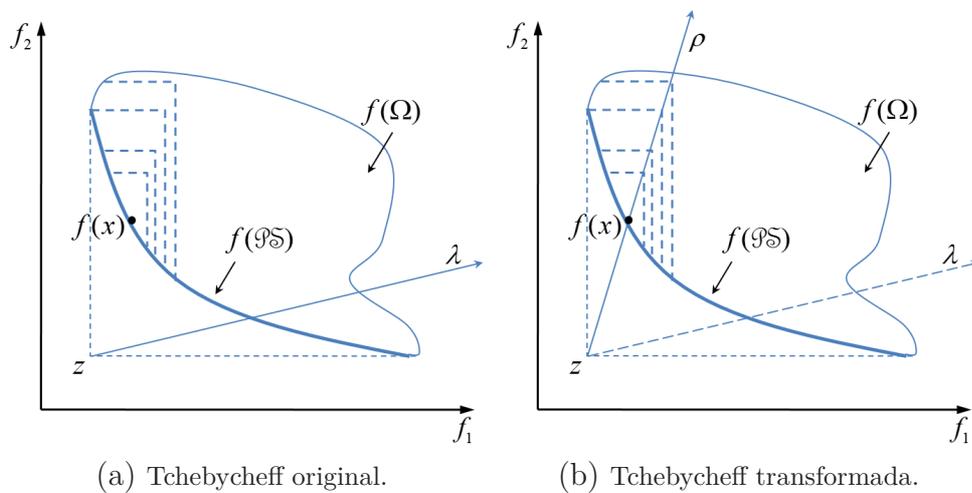


Figura 4 – Curvas de nível de um subproblema escalar e a relação geométrica entre seu vetor de ponderação e a solução ótima correspondente, considerando-se diferentes estratégias de agregação Tchebycheff.

Finalmente, ressalta-se que a estratégia Tchebycheff transformada não foi proposta originalmente por esta autora. Observou-se, ao longo do desenvolvimento deste trabalho, que um outro grupo de pesquisadores havia proposto esta mesma ideia em um artigo publicado na *Evolutionary Computation* (QI et al., 2013). Entretanto, é importante frisar que o desenvolvimento desta abordagem neste trabalho ocorreu de forma independente.

3.4 Considerações Finais

Este capítulo apresentou uma metodologia, baseada em camadas, para uma geração uniformemente distribuída de vetores de ponderação e também uma função de agregação escalar Tchebycheff transformada. Estas abordagens são particularmente úteis no contexto de MOEAs baseados em decomposição, as quais são avaliadas no capítulo seguinte considerando-se um conjunto de funções teste.

4 Resultados

4.1 Introdução

Neste capítulo avalia-se a aplicação das abordagens propostas, *multiple-layer simplex-lattice design* e função de agregação Tchebycheff transformada, considerando-se problemas testes populares, i.e., DTLZ1 a DTLZ4 com 3, 5, 8, 10 e 15 objetivos, e indicadores de qualidade usuais, *Inverted Generational Distance* (IGD) e *Hypervolume* (HV). Os resultados obtidos indicam uma forte competitividade das técnicas sugeridas, tanto para a geração de vetores de ponderação melhor distribuídos no espaço de objetivos quanto para a estimação de fronteiras com convergências mais atrativas.

4.2 Problemas Teste

Os problemas DTLZ1 a DTLZ4 (DEB et al., 2005) têm sido frequentemente abordados na literatura no contexto de otimização com muitos objetivos (TAN et al., 2013; DEB; JAIN, 2014). Neste trabalho estas instâncias são avaliadas considerando $n_f \in \{3,5,8,10,15\}$. O número de variáveis de decisão é definido como $n_v = n_f + r - 1$, em que $r = 5$ para DTLZ1 e $r = 10$ para DTLZ2–4 (DEB et al., 2005).

Os demais problemas da família DTLZ não são considerados pois apresentam fronteira Pareto degenerada (DTLZ5 e DTLZ6), desconexa (DTLZ7), ou compõem um MOP restrito (DTLZ8 e DTLZ9) (DEB et al., 2005). Obviamente é importante a investigação de problemas com essas características no contexto de MOPs com muitos objetivos, entretanto, dado o grande desafio atual para tratar problemas simples, e.g., quadráticos, com muitos objetivos, a literatura tem frequentemente ignorado tais problemas (TAN et al., 2013; DEB; JAIN, 2014).

4.3 Indicadores de Qualidade

4.3.1 *Inverted Generational Distance* (IGD)

O indicador de qualidade IGD (ZITZLER et al., 2003) fornece uma medida de distância entre os pontos de uma amostra detalhada da fronteira Pareto-ótima (\mathcal{PF}) e as soluções estimadas (\mathcal{S}) obtidas por um algoritmo. O valor IGD de um conjunto \mathcal{S} é dado por:

$$IGD(\mathcal{PF}, \mathcal{S}) = \left(\sum_{\mathbf{y} \in \mathcal{PF}} d(\mathbf{y}, \mathcal{S}) \right) / |\mathcal{PF}| \quad (4.1)$$

em que $d(\mathbf{y}, \mathcal{S})$ é a distância Euclidiana entre um ponto $\mathbf{y} \in \mathcal{PF}$ e seu vizinho mais próximo em \mathcal{S} , e $|\mathcal{PF}|$ é a cardinalidade de \mathcal{PF} . Assim, quanto menor o valor de IGD, melhor é a aproximação de \mathcal{S} em relação a \mathcal{PF} .

Nesta dissertação, o conjunto \mathcal{PF} é obtido considerando os N vetores de ponderação e informações analíticas sobre os problemas DTLZ1–4 (DEB et al., 2005). De forma geral, obtém-se as soluções ótimas que interceptam os vetores de ponderação, logo $|\mathcal{PF}| = |\mathcal{S}|$. As coordenadas de uma solução ótima para o problema DTLZ1, associada a um vetor ponderado $\boldsymbol{\lambda}$, são dadas por (LI et al., 2014):

$$f_i(\mathbf{x}^*) = 0.5\lambda_i / \sum_{j=1}^{n_f} \lambda_j, \quad i \in \{1, \dots, n_f\}.$$

De maneira similar, para os problemas DTLZ2–4, tem-se (LI et al., 2014):

$$f_i(\mathbf{x}^*) = \lambda_i / \sqrt{\sum_{j=1}^{n_f} \lambda_j^2}, \quad i \in \{1, \dots, n_f\}.$$

4.3.2 Hypervolume (HV)

O indicador de qualidade HV (ZITZLER; THIELE, 1999) calcula o hipervolume envolto pela fronteira de soluções estimadas \mathcal{S} e um ponto de referência \mathbf{y}_r dominado por todas as soluções de \mathcal{S} . Formalmente, esse indicador é descrito como a medida Lebesgue Λ da união dos hipercubos h_i definidos por um ponto não-dominado \mathbf{y}_i e \mathbf{y}_r :

$$HV(\mathcal{S}) = \Lambda \left(\left\{ \bigcup_i h_i \mid \mathbf{y}_i \in \mathcal{S} \right\} \right). \quad (4.2)$$

Supondo a minimização das funções de um MOP, quanto maior o volume dominado, melhor a característica da fronteira aproximada. Para o problema DTLZ1, usou-se $\mathbf{y}_r = (1.0, \dots, 1.0)$, e para DTLZ2–4 considerou-se $\mathbf{y}_r = (2.0, \dots, 2.0)$. O hipervolume exato foi calculado para $n_f = 3$, enquanto para $n_f > 3$ obteve-se uma aproximação do hipervolume via simulação de Monte Carlo (BADER; ZITZLER, 2011). Para o cálculo aproximado de HV são empregados 10,000 pontos na simulação de Monte Carlo.

4.4 Planejamento Experimental e Análise dos Resultados

4.4.1 Avaliação da Estratégia *Multiple-Layer Simplex-Lattice Design* (MLSLD)

Com o intuito de avaliar a metodologia proposta para uma geração mais uniforme dos vetores de ponderação (baseada em múltiplas camadas), esta é comparada com a técnica *simplex-lattice design* de uma única camada. Para simplificar o estudo, os problemas teste DTLZ1–4 foram considerados apenas com $n_f \in \{3, 5\}$. A Tabela 1 apresenta o parâmetro H empregado nas técnicas *simplex-lattice design* (SLD) e *multiple-layer simplex-lattice design* (MLSLD). Os valores de H foram escolhidos de forma que o tamanho da população de soluções (N) fosse aproximadamente o mesmo para ambas abordagens.

Tabela 1 – Parâmetros adotados na comparação entre SLD e MLSLD: tamanho da amostra H e tamanho da população N .

n_f	<i>simplex-lattice design</i> (SLD)		<i>multiple-layer SLD</i> (MLSLD)	
	Tamanho da mostra (H)	População (N)	Tamanho da amostra (H)	População (N)
3	$H = 12$	91	$H_1 = 9, H_2 = 7$	91
5	$H = 6$	210	$H_1 = 5, H_2 = 4, H_3 = 2$	211

Os fatores de contração para $\mathbf{H} = (9, 7)$ são $\boldsymbol{\tau} = (1.0, 0.5)$, respectivamente, e para $\mathbf{H} = (5, 4, 2)$ usou-se $\boldsymbol{\tau} = (1.0, 0.8, 0.5)$. Os valores adotados para τ são arbitrários (não empregou-se nenhuma estratégia para ajuste ótimo destes parâmetros).

As estratégias SLD e MLSLD foram incorporadas ao MOEA/D com a função de escalarização Tchebycheff usual. Os operadores evolutivos do MOEA/D aplicado incluem a variação $DE/rand/1/bin$ ($F = 0.5$ e $CR = 0.5$) seguida de uma mutação polinomial ($p_m = 1/n_v$ e $\eta_m = 20$) (ZHANG; LI, 2007). Os algoritmos são definidos com tamanho de vizinhança fixo, i.e., $T = 20$ (ZHANG; LI, 2007; DEB; JAIN, 2014); e considera-se uma probabilidade $\delta = 0.1$ de que a seleção para variação envolva todos os vetores referência (LI; ZHANG, 2009). Os algoritmos MOEA/D-SLD e MOEA/D-MLSLD foram executados 30 vezes em cada uma das instâncias.

Os resultados obtidos são apresentadas na Tabela 2. Observa-se um indício de que o gerador de vetores de ponderação baseado na abordagem MLSLD promove, de fato, uma distribuição mais uniforme destes vetores. De acordo com os indicadores de qualidade adotados, os resultados sugerem que o MOEA/D-MLSLD é capaz de estimar soluções Pareto-ótimas com valores de HV e, principalmente, IGD mais interessantes do que aqueles obtidos via MOEA/D-SLD. Nota-se inclusive que os valores de HV são mais atraentes nos problemas com $n_f = 5$, o que sugere um maior potencial da estratégia MLSLD para problemas com muitos objetivos. Essas observações certamente motivam a utilização da técnica MLSLD em substituição à SLD.

4.4.2 Avaliação da Estratégia de Agregação Tchebycheff Transformada

Para avaliar o desempenho da função de agregação proposta, a mesma é comparada com as escalarizações Tchebycheff padrão e PBI. Essas estratégias são incorporadas ao método MOEA/D (ZHANG; LI, 2007), resultando nos algoritmos MOEA/D-tch, MOEA/D-pbi e MOEA/D-tch_t¹.

Empregou-se as mesmas definições para o MOEA/D apresentado anteriormente, entretanto, usou-se tamanho da amostra H e tamanho de população N de acordo com a Tabela 3. Além disso, na estratégia PBI empregou-se um fator de penalidade $\theta = 5$ (ZHANG; LI, 2007; DEB; JAIN, 2014). Novamente, os algoritmos foram executados 30

¹ Os algoritmos e indicadores de qualidade implementados estão disponíveis em <http://www.ppgee.ufmg.br/~lusoba/files/cbic2015.zip>.

Tabela 2 – Valores de HV e IGD (Melhor, Médio, Pior): SLD *versus* MLSLD. As medidas em **negrito** sugerem o método que alcançou o melhor desempenho.

DTLZ	Indicador	Obj. (MaxGen)	MOEA/D-SLD	MOEA/D-MLSLD
1	HV	3 (400)	(9.6773e-01 9.6751e-01 9.6733e-01)	(9.6531e-01 9.6508e-01 9.6454e-01)
		5 (600)	(9.9780e-01 9.9683e-01 9.9540e-01)	(9.9860e-01 9.9784e-01 9.9710e-01)
	IGD	3 (400)	(3.2431e-02 3.2879e-02 3.3501e-02)	(3.0170e-02 3.0785e-02 3.1157e-02)
		5 (600)	(1.1214e-01 1.1257e-01 1.1276e-01)	(7.9058e-02 7.9438e-02 7.9930e-02)
2	HV	3 (250)	(7.3750e+00 7.3719e+00 7.3695e+00)	(7.3764e+00 7.3725e+00 7.3693e+00)
		5 (350)	(3.1595e+01 3.1525e+01 3.1425e+01)	(3.1632e+01 3.1538e+01 3.1434e+01)
	IGD	3 (250)	(7.3165e-02 7.4889e-02 7.6523e-02)	(7.1316e-02 7.3521e-02 7.6100e-02)
		5 (350)	(2.9283e-01 2.9445e-01 2.9715e-01)	(2.8699e-01 2.8786e-01 2.8869e-01)
3	HV	3 (1000)	(7.3734e+00 7.3706e+00 7.3689e+00)	(7.3698e+00 7.3674e+00 7.3666e+00)
		5 (1000)	(3.1594e+01 3.1533e+01 3.1414e+01)	(3.1642e+01 3.1561e+01 3.1494e+01)
	IGD	3 (1000)	(7.3935e-02 7.5605e-02 7.7022e-02)	(7.3412e-02 7.4631e-02 7.6008e-02)
		5 (1000)	(2.9358e-01 2.9464e-01 2.9537e-01)	(2.3797e-01 2.4114e-01 2.4470e-01)
4	HV	3 (600)	(7.3840e+00 6.9076e+00 4.0000e+00)	(7.3708e+00 6.6943e+00 4.0000e+00)
		5 (1000)	(3.1613e+01 3.1481e+01 3.0816e+01)	(3.1632e+01 3.1511e+01 3.0861e+01)
	IGD	3 (600)	(7.2782e-02 2.2320e-01 9.5033e-01)	(7.3510e-02 2.4839e-01 9.4355e-01)
		5 (1000)	(2.9280e-01 3.0996e-01 4.7755e-01)	(2.4144e-01 2.5947e-01 4.4545e-01)

vezes em cada um dos problemas.

Tabela 3 – Parâmetros adotados na comparação entre TCH_t , TCH e PBI: tamanho da amostra H e tamanho da população N .

n_f	Tamanho da amostra (H)	Tamanho da população (N)
3	$H = 12$	91
5	$H = 6$	210
8	$H_1 = 3, H_2 = 2$	156
10	$H_1 = 3, H_2 = 2$	275
15	$H_1 = 2, H_2 = 2, H_3 = 1$	255

Os fatores de contração para $\mathbf{H} = (3, 2)$ são $\boldsymbol{\tau} = (1.0, 0.5)$, respectivamente, e para $\mathbf{H} = (2, 2, 1)$ usou-se $\boldsymbol{\tau} = (1.0, 0.8, 0.5)$. Os valores adotados para $\boldsymbol{\tau}$ são arbitrários (não empregou-se nenhuma estratégia para ajuste ótimo destes parâmetros).

Inicialmente, avalia-se a capacidade das estratégias em estimar soluções que promovam a minimização das distâncias d_1 e d_2 (veja Fig. 2a) em relação aos vetores de ponderação gerados. A Fig. 5 apresenta gráficos em que as técnicas TCH_t , TCH e PBI são comparadas considerando as médias das distâncias para todas as fronteiras estimadas. De maneira geral, a técnica TCH obteve o pior desempenho em todos os casos avaliados, o que é um resultado direto do efeito de deslocamento do cone-negativo em relação ao vetor de referência. Nessa situação, a medida d_2 pode assumir valores elevados para alguns vetores de ponderação. Por outro lado, as estratégias TCH_t e PBI foram capazes de aproximar adequadamente as medidas ótimas em praticamente todos os problemas.

Embora não esteja evidente pelo gráfico, como a decomposição PBI exerce uma maior pressão seletiva, essa técnica alcançou em geral menores valores de d_2 . Entretanto, essa mesma característica pode explicar sua maior dificuldade de convergência no problema DTLZ3, uma vez que uma pressão seletiva muito forte pode comprometer a diversidade das soluções estimadas.

A Tabela 4 apresenta os valores de IGD (melhor, médio e pior) calculados com base nas soluções de \mathcal{PF} . Essa medida está diretamente relacionada aos valores médios de d_2 encontrados na análise anterior e, conforme esperado, a abordagem PBI obteve resultados mais atraentes em um maior número de instâncias. A técnica TCH_t , entretanto, superou a decomposição TCH em todos os casos, e encontrou em geral valores de IGD competitivos com aqueles estimados pela PBI.

Observa-se pela Tabela 5 que a decomposição TCH_t dominou a técnica TCH em todos os testes e, além disso, apresentou muitos valores de HV competitivos ou mais interessantes do que aqueles obtidos pela PBI. De fato, como a TCH_t não aproximou tão bem as soluções ótimas pertencentes a \mathcal{PF} , esse maior *espalhamento* das soluções estimadas pode ter contribuído para a determinação de maiores valores de HV. Ainda nesse sentido, é possível que a técnica TCH_t consiga melhorar seu indicador IGD, caso um conjunto maior de soluções eficientes seja considerado em substituição a \mathcal{PF} . Essa análise, entretanto, não é explorada neste trabalho.

Por fim, as Figs. 6–10 apresentam a sobreposição de todas as 30 fronteiras estimadas pelos algoritmos comparados, para $n_f \in \{3,5,8,10,15\}$, respectivamente. A fronteira Pareto real destes problemas é tal que $f_i = 0.5, \forall i = 1, \dots, n_f$, para DTLZ1 e $f_i = 1, \forall i = 1, \dots, n_f$, para DTLZ2 a DTLZ4.

Considerando-se a visualização destas sobreposições, ressalta-se que o MOEA/D- tch_t aproximou a fronteira Pareto-ótima em todas as execuções realizadas, e para todos os problemas teste. Os métodos MOEA/D- tch e MOEA/D- pbi , no entanto, apresentaram falha de convergência, inclusive em problemas mais simples, com $n_f = 3$. Por exemplo, o MOEA/D- tch obteve diversidade/espalhamento aparentemente pobre em relação ao MOEA/D- tch_t , para todos os problemas onde $n_f = 3$; além disso, o MOEA/D- pbi apresentou dificuldades de convergência nos problemas DTLZ1 e DTLZ3 quando $n_f = 3$ (Fig. 6).

Nas Figs. 7–10 as fronteiras estimadas são apresentadas por meio de gráficos de coordenadas paralelas, em que cada coordenada representa um objetivo específico. Nos problemas em que $n_f = 5$ (Fig. 7), o MOEA/D- tch_t obteve resultados similares ao MOEA/D- pbi , entretanto este último não convergiu para a fronteira real do DTLZ3 em todas as execuções. O MOEA/D- tch convergiu adequadamente para a fronteira Pareto, entretanto, está evidente sua dificuldade para obter diversidade/espalhamento. Este fato pode ser constatado pois a faixa de variação dos eixos coordenados, f_i , é muito descon-

tínua, indicando que algumas regiões da fronteira real não foram estimadas. Nos demais problemas, em que $n_f \in \{8,10,15\}$, o MOEA/D-tch não conseguiu aproximar toda a extensão da fronteira real em nenhuma das execuções (Figs. 8–10). Por outro lado, estas figuras sugerem uma forte similaridade dos resultados obtidos pelos métodos MOEA/D-tch_t e MOEA/D-pbi, no entanto, o MOEA/D-pbi enfrentou dificuldades de convergência no DTLZ3 com $n_f = 8$. Além disso, este método não apresentou um espalhamento superior ao do MOEA/D-tch_t nos problemas DTLZ1 a DTLZ3 com $n_f = 15$.

De forma geral, considerando os resultados apresentados nessa seção, observa-se que as fronteiras aproximadas pelo MOEA/D-tch_t são competitivas em ambos os critérios convergência e diversidade.

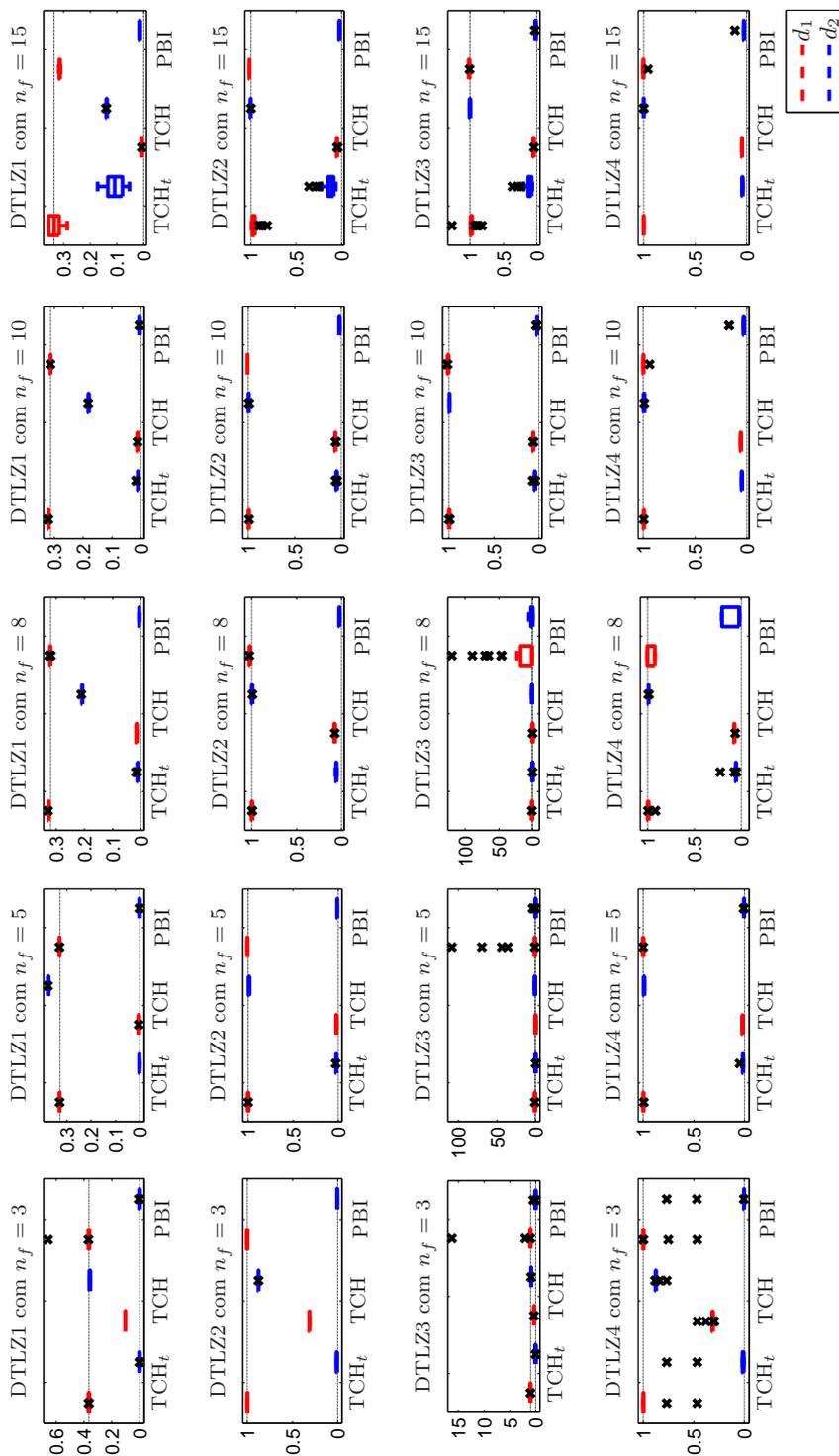


Figura 5 – Boxplot das medidas médias das distâncias d_1 e d_2 considerando todas as fronteiras estimadas e todas as instâncias testadas. Em cada problema, o valor ótimo de d_1 deve tender à curva pontilhada superior (definida em 0.3643, 0.3288, 0.3201, 0.3131 e 0.3358 para o problema DTLZ1 com $n_f \in \{3, 5, 8, 10, 15\}$, respectivamente, e definida em 1.0 para os problemas DTLZ2-4), e o valor ótimo de d_2 deve tender à zero (curva pontilhada inferior).

Tabela 4 – Valores de IGD (Melhor, Médio, Pior): TCH_t versus TCH versus PBI. As medidas em **negrito** sugerem o método que alcançou o melhor desempenho.

DTLZ	Obj. (MaxGen)	MOEA/D-tch _t	MOEA/D-tch	MOEA/D-pbi
1	3 (400)	(8.013e-04 1.240e-03 2.813e-03)	(3.253e-02 3.288e-02 3.335e-02)	(5.125e-04 6.939e-03 1.738e-01)
	5 (600)	(1.168e-03 1.615e-03 1.991e-03)	(1.123e-01 1.125e-01 1.129e-01)	(7.340e-04 1.275e-03 2.508e-03)
	8 (750)	(1.155e-02 1.371e-02 2.002e-02)	(2.178e-01 2.228e-01 2.263e-01)	(6.041e-03 7.486e-03 1.026e-02)
	10 (1000)	(1.071e-02 1.292e-02 1.888e-02)	(2.366e-01 2.407e-01 2.435e-01)	(4.884e-03 5.394e-03 6.390e-03)
	15 (1500)	(3.102e-02 9.206e-02 1.727e-01)	(2.943e-01 2.994e-01 3.014e-01)	(2.489e-02 2.726e-02 2.955e-02)
2	3 (250)	(6.839e-03 9.122e-03 1.083e-02)	(7.316e-02 7.489e-02 7.652e-02)	(3.533e-03 4.169e-03 4.799e-03)
	5 (350)	(1.676e-02 1.968e-02 2.629e-02)	(2.928e-01 2.944e-01 2.971e-01)	(1.169e-02 1.394e-02 1.551e-02)
	8 (500)	(5.212e-02 5.681e-02 6.434e-02)	(6.735e-01 7.461e-01 7.692e-01)	(2.929e-02 3.230e-02 3.791e-02)
	10 (750)	(4.504e-02 4.944e-02 5.742e-02)	(7.573e-01 8.421e-01 8.597e-01)	(2.018e-02 2.223e-02 2.411e-02)
	15 (1000)	(4.370e-02 9.522e-02 3.694e-01)	(1.022e+00 1.045e+00 1.091e+00)	(2.783e-02 2.956e-02 3.180e-02)
3	3 (1000)	(4.126e-03 5.600e-03 7.728e-03)	(7.440e-02 7.540e-02 7.619e-02)	(2.421e-03 4.351e-01 1.175e+01)
	5 (1000)	(1.071e-02 1.387e-02 1.773e-02)	(2.931e-01 2.945e-01 2.954e-01)	(9.410e-03 5.187e+00 7.065e+01)
	8 (1000)	(5.041e-02 5.813e-02 7.498e-02)	(7.076e-01 7.398e-01 7.633e-01)	(3.385e-02 1.081e+01 8.290e+01)
	10 (1500)	(4.104e-02 4.653e-02 6.724e-02)	(8.069e-01 8.472e-01 8.605e-01)	(1.881e-02 2.324e-02 3.990e-02)
	15 (2000)	(4.213e-02 8.628e-02 3.914e-01)	(1.012e+00 1.048e+00 1.082e+00)	(2.341e-02 2.531e-02 3.207e-02)
4	3 (600)	(5.686e-03 1.999e-01 9.503e-01)	(7.294e-02 1.925e-01 9.503e-01)	(1.854e-03 8.328e-02 9.503e-01)
	5 (1000)	(8.966e-03 1.504e-02 4.699e-02)	(2.932e-01 2.953e-01 2.975e-01)	(3.496e-03 4.320e-03 5.891e-03)
	8 (1250)	(5.236e-02 7.010e-02 2.431e-01)	(7.105e-01 7.589e-01 7.877e-01)	(1.957e-02 9.438e-02 2.264e-01)
	10 (2000)	(4.801e-02 4.997e-02 5.200e-02)	(8.319e-01 8.572e-01 8.983e-01)	(2.844e-02 4.703e-02 1.867e-01)
	15 (3000)	(4.135e-02 4.213e-02 4.286e-02)	(1.020e+00 1.056e+00 1.085e+00)	(2.694e-02 3.830e-02 1.266e-01)

Tabela 5 – Valores de HV (Melhor, Médio, Pior): TCH_t versus TCH versus PBI. As medidas em **negrito** sugerem o método que alcançou o melhor desempenho.

DTLZ	Obj. (MaxGen)	MOEA/D-tch _t	MOEA/D-tch	MOEA/D-pbi
1	3 (400)	(9.737e-01 9.736e-01 9.735e-01)	(9.677e-01 9.675e-01 9.673e-01)	(9.736e-01 9.699e-01 8.641e-01)
	5 (600)	(9.994e-01 9.990e-01 9.983e-01)	(9.980e-01 9.970e-01 9.961e-01)	(9.995e-01 9.989e-01 9.985e-01)
	8 (750)	(1.000e+00 1.000e+00 9.999e-01)	(9.967e-01 9.958e-01 9.945e-01)	(1.000e+00 1.000e+00 9.998e-01)
	10 (1000)	(1.000e+00 1.000e+00 1.000e+00)	(9.973e-01 9.964e-01 9.947e-01)	(1.000e+00 1.000e+00 1.000e+00)
	15 (1500)	(1.000e+00 1.000e+00 9.999e-01)	(9.872e-01 9.832e-01 9.799e-01)	(1.000e+00 1.000e+00 9.999e-01)
2	3 (250)	(7.413e+00 7.413e+00 7.412e+00)	(7.375e+00 7.372e+00 7.369e+00)	(7.411e+00 7.410e+00 7.409e+00)
	5 (350)	(3.176e+01 3.169e+01 3.165e+01)	(3.159e+01 3.152e+01 3.142e+01)	(3.175e+01 3.169e+01 3.162e+01)
	8 (500)	(2.560e+02 2.558e+02 2.556e+02)	(2.443e+02 2.363e+02 2.347e+02)	(2.559e+02 2.558e+02 2.557e+02)
	10 (750)	(1.024e+03 1.024e+03 1.024e+03)	(9.699e+02 9.257e+02 9.157e+02)	(1.024e+03 1.024e+03 1.024e+03)
	15 (1000)	(3.277e+04 3.277e+04 3.277e+04)	(2.755e+04 2.266e+04 2.157e+04)	(3.277e+04 3.277e+04 3.277e+04)
3	3 (1000)	(7.414e+00 7.413e+00 7.413e+00)	(7.373e+00 7.370e+00 7.369e+00)	(7.413e+00 4.711e+00 0)
	5 (1000)	(3.173e+01 3.170e+01 3.165e+01)	(3.161e+01 3.152e+01 3.143e+01)	(3.172e+01 2.743e+01 0)
	8 (1000)	(2.559e+02 2.558e+02 2.557e+02)	(2.382e+02 2.361e+02 2.341e+02)	(2.559e+02 1.758e+02 0)
	10 (1500)	(1.024e+03 1.024e+03 1.024e+03)	(9.330e+02 9.213e+02 9.140e+02)	(1.024e+03 1.024e+03 1.024e+03)
	15 (2000)	(3.277e+04 3.277e+04 3.276e+04)	(2.427e+04 2.239e+04 2.140e+04)	(3.277e+04 3.277e+04 3.276e+04)
4	3 (600)	(7.414e+00 6.777e+00 4.000e+00)	(7.380e+00 6.972e+00 4.000e+00)	(7.413e+00 7.152e+00 4.000e+00)
	5 (1000)	(3.176e+01 3.170e+01 3.165e+01)	(3.158e+01 3.153e+01 3.147e+01)	(3.175e+01 3.169e+01 3.163e+01)
	8 (1250)	(2.560e+02 2.559e+02 2.556e+02)	(2.441e+02 2.377e+02 2.352e+02)	(2.559e+02 2.558e+02 2.554e+02)
	10 (2000)	(1.024e+03 1.024e+03 1.024e+03)	(9.401e+02 9.270e+02 9.147e+02)	(1.024e+03 1.024e+03 1.024e+03)
	15 (3000)	(3.277e+04 3.277e+04 3.276e+04)	(2.467e+04 2.318e+04 2.192e+04)	(3.277e+04 3.277e+04 3.277e+04)

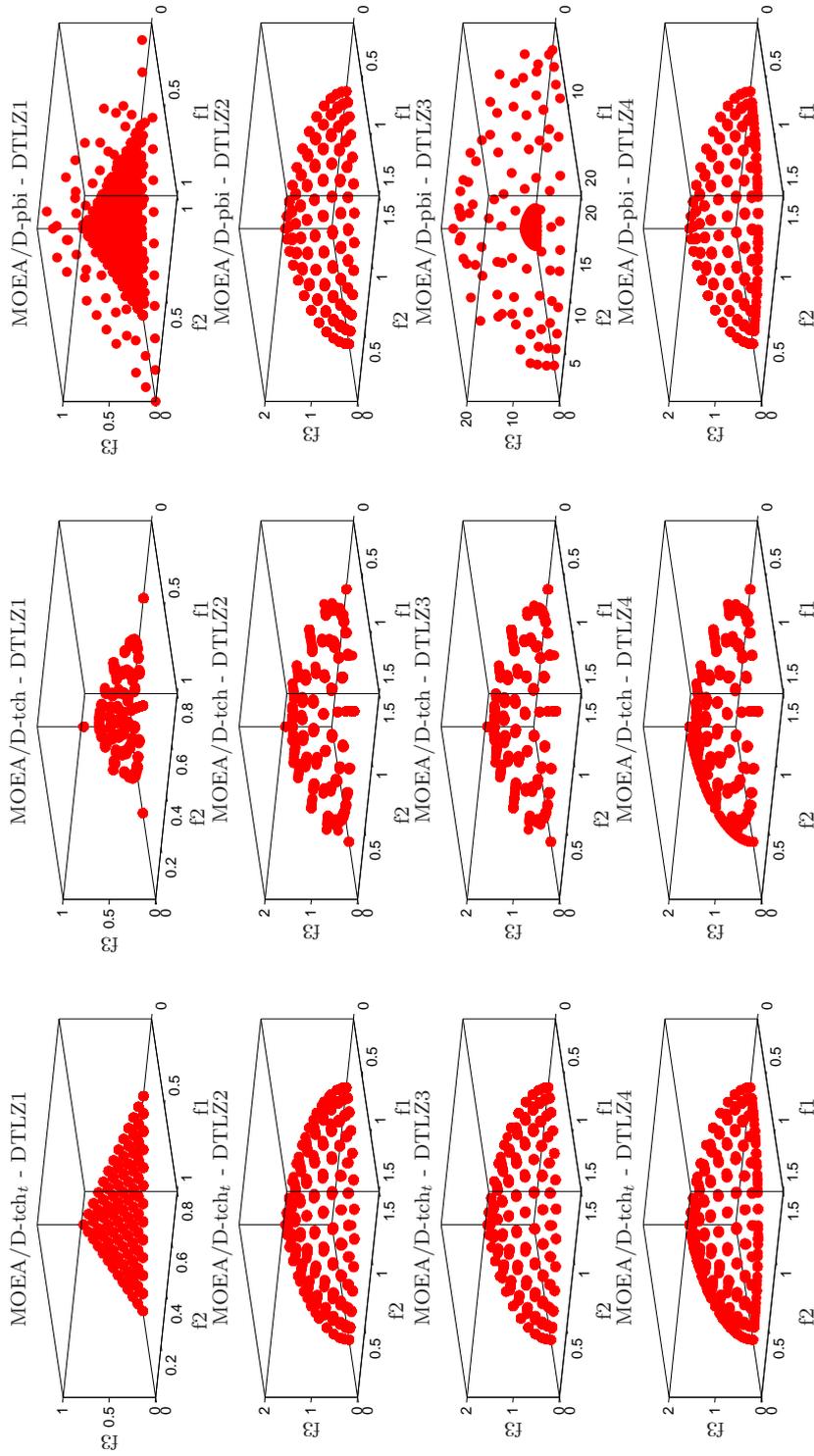


Figura 6 – Gráficos tridimensionais considerando todas as fronteiras estimadas e problemas com 3 objetivos.

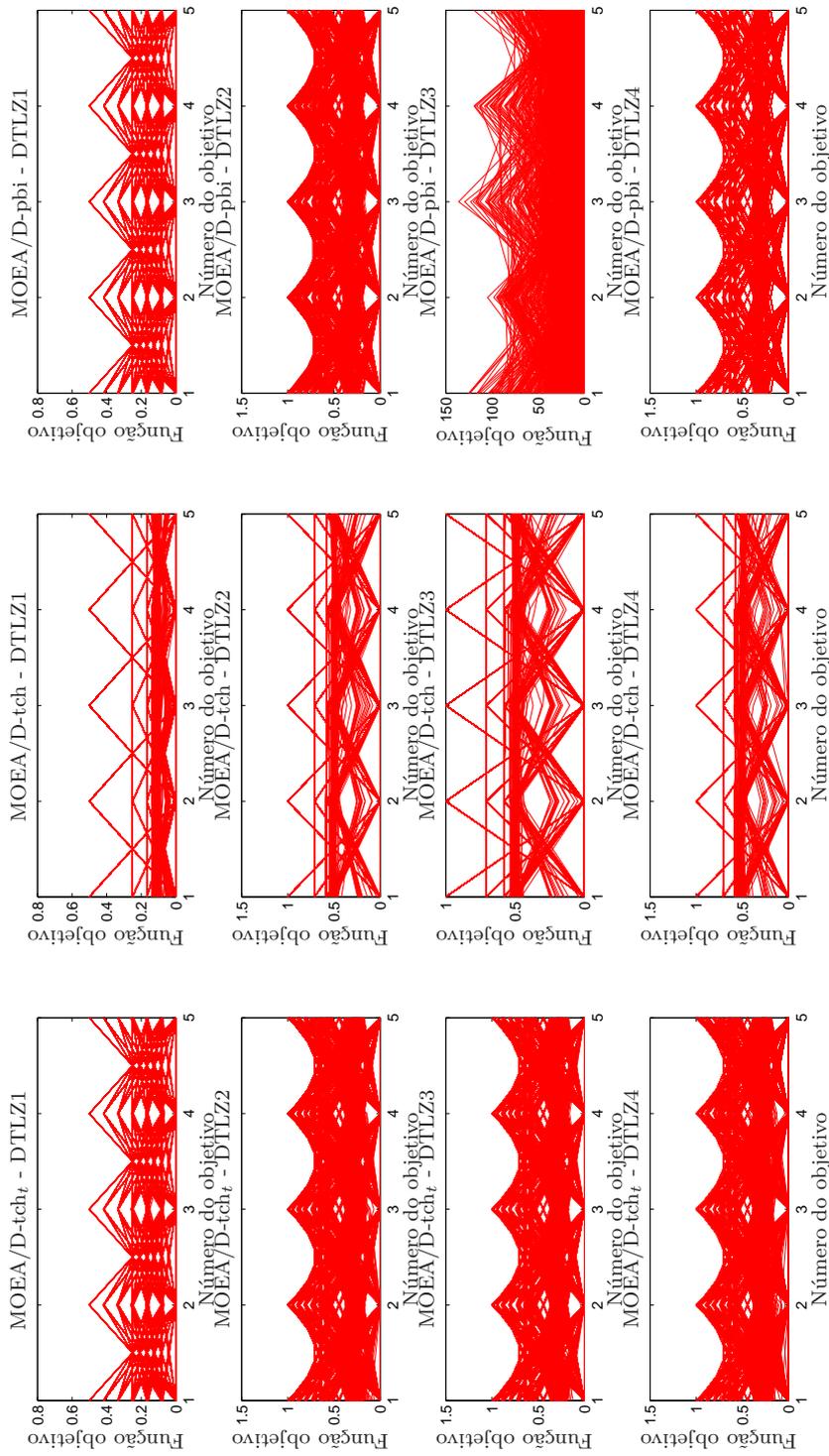


Figura 7 – Gráficos de coordenadas paralelas considerando todas as fronteiras estimadas e problemas com 5 objetivos.

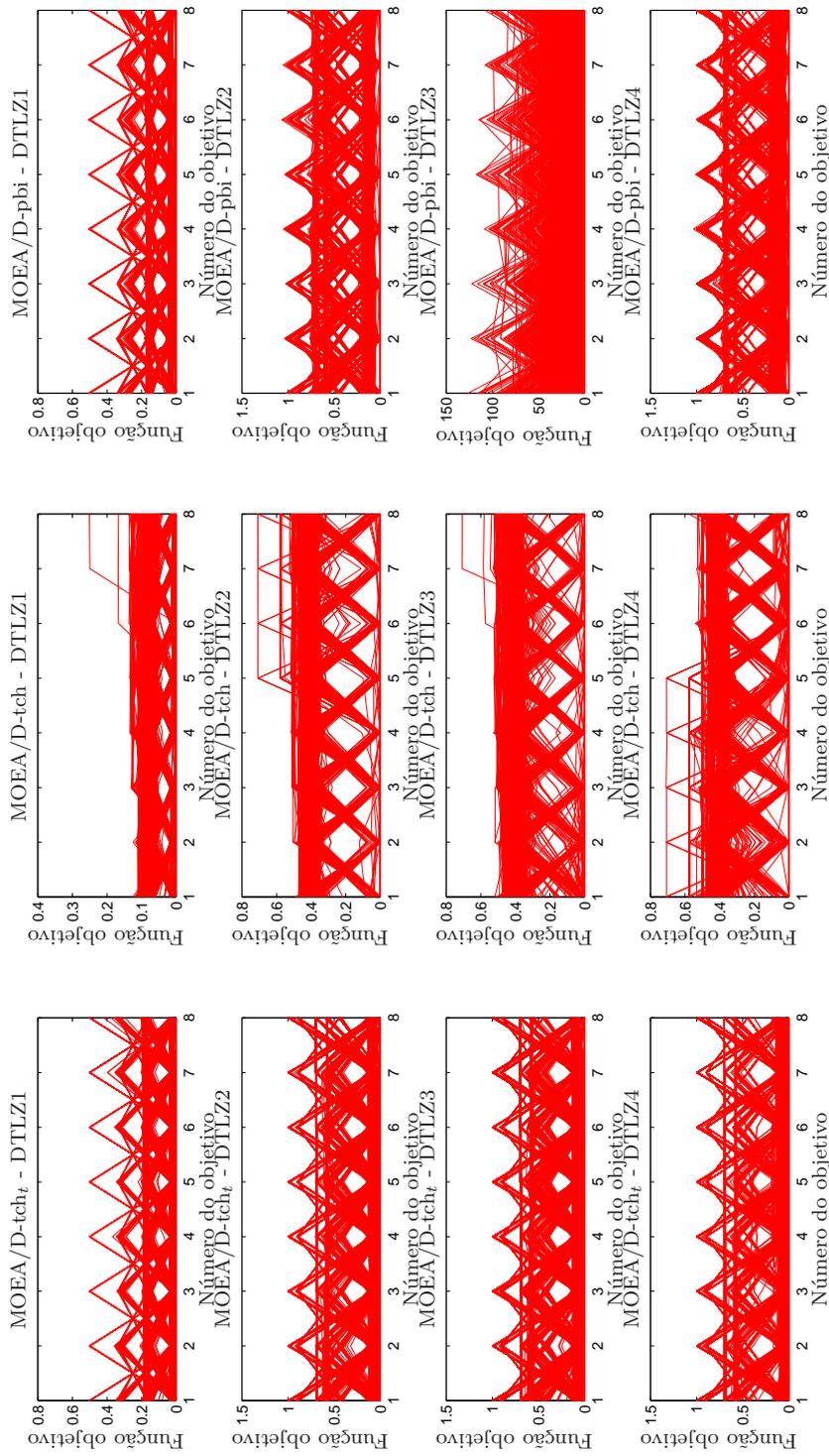


Figura 8 – Gráficos de coordenadas paralelas considerando todas as fronteiras estimadas e problemas com 8 objetivos.

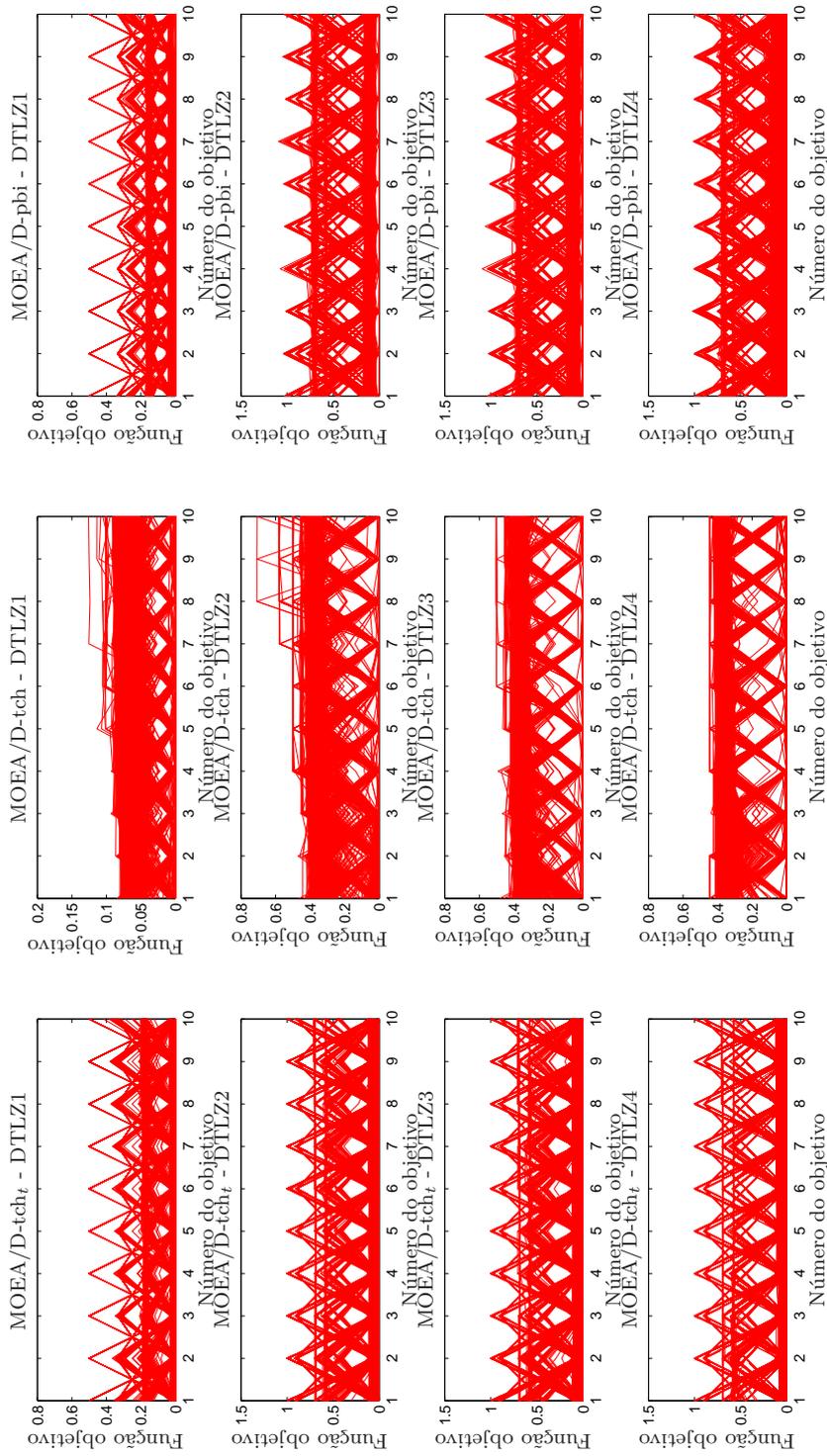


Figura 9 – Gráficos de coordenadas paralelas considerando todas as fronteiras estimadas e problemas com 10 objetivos.

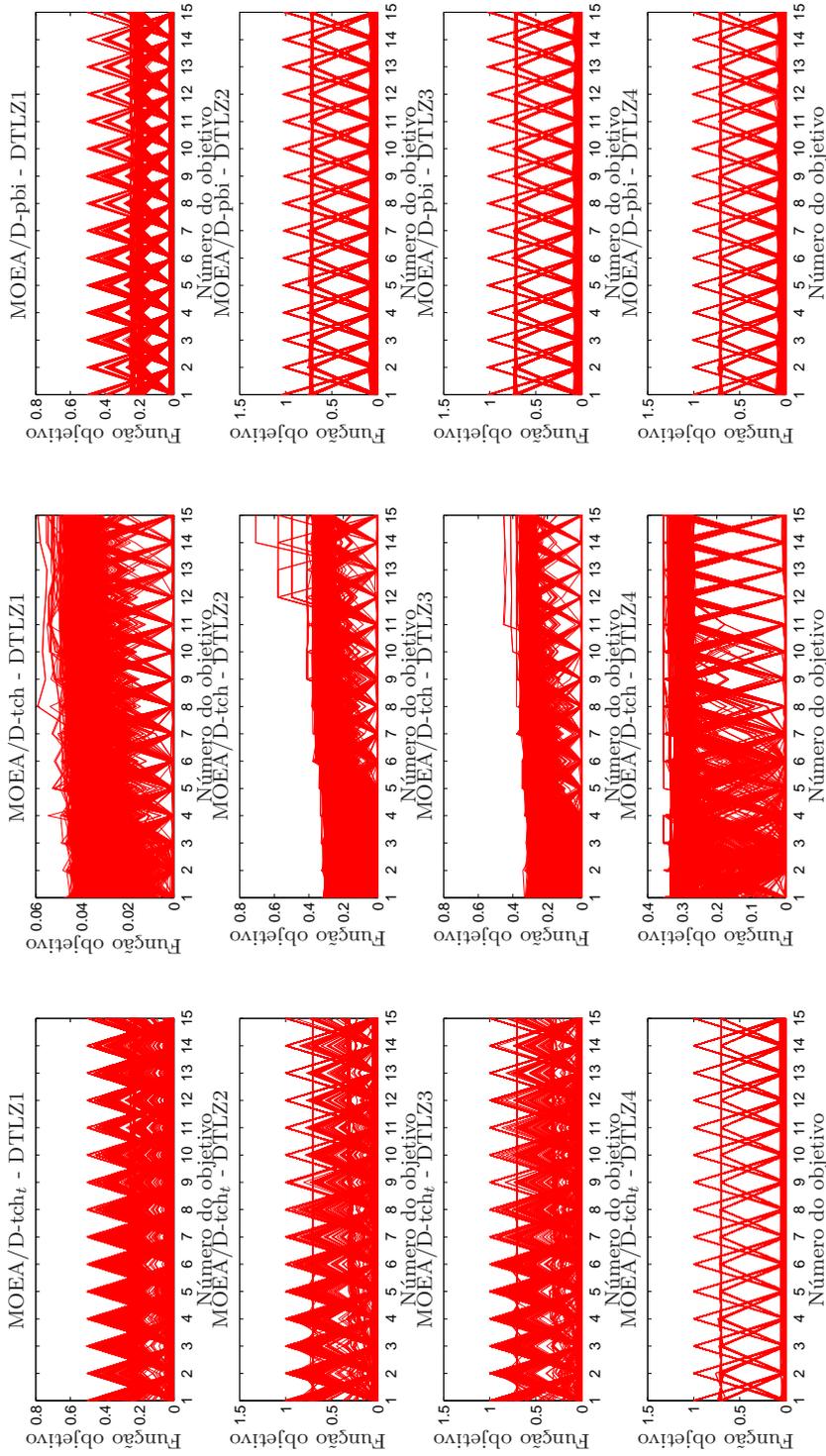


Figura 10 – Gráficos de coordenadas paralelas considerando todas as fronteiras estimadas e problemas com 15 objetivos.

4.5 Considerações Finais

Os resultados e discussões apresentados ao longo deste capítulo, embora preliminares, sugerem o potencial e competitividade de aplicação das abordagens propostas no contexto de MOEAs baseados em decomposição. O capítulo seguinte retoma as principais conclusões obtidas a partir desta análise e enumera direções para a continuidade dessa pesquisa.

5 Conclusões

5.1 Considerações Finais

Nesta dissertação apresentou-se uma descrição ampliada da estrutura do MOEA/D, onde se caracterizou seus principais módulos funcionais (e.g., geradores de vetores de ponderação, funções de agregação, operadores de variação, etc.) e inúmeros trabalhos desenvolvidos no âmbito de cada um destes blocos. Essa descrição baseada em componentes do MOEA/D foi abordada na literatura pela primeira vez no artigo sob revisão (CAMPELO; BATISTA; ARANHA, 2017) e no trabalho corrente, representando portanto uma contribuição original desta pesquisa, a qual pode ser útil principalmente para guiar outros estudos relacionados a módulos específicos do MOEA/D. Além disso, investigou-se duas abordagens úteis para a otimização de problemas com muitos objetivos no contexto de MOEAs baseados em decomposição: a técnica *multiple-layer simplex-lattice design* (MLSLD) e uma função de escalarização Tchebycheff transformada (LÓPEZ; BATISTA, 2015; QI et al., 2013).

Conforme discutido no capítulo anterior, a técnica MLSLD possibilita tanto controlar, de forma mais flexível, o tamanho da população de vetores de ponderação N , por meio da definição de múltiplos tamanhos de amostras H , quanto promover uma distribuição mais uniforme dos vetores referência no domínio de objetivos. Essas características possibilitam o uso de populações com tamanhos modestos em problemas com muitos objetivos (ao contrário do que ocorre ao se usar *simplex-lattice design*) e também promovem a melhoria da diversidade das soluções estimadas, respectivamente. Os testes realizados com a aplicação da estratégia MLSLD sugerem resultados atraentes, com valores mais interessantes de HV e IGD.

A função de agregação Tchebycheff transformada aplica uma transformação sobre a função Tchebycheff padrão, possibilitando um melhor controle sobre a convergência e diversidade das soluções aproximadas. O desempenho da função de agregação proposta foi comparado ao das técnicas Tchebycheff padrão e PBI nos problemas DTLZ1–4, assumindo $n_f \in \{3,5,8,10,15\}$. Os resultados obtidos sugerem uma forte competitividade entre TCH_t e PBI, em que a primeira destaca-se por não depender de parâmetros adicionais e, principalmente, por ter apresentado valores promissores de HV e uma taxa de aproximação da fronteira Pareto-ótima de 100%, em todos os problemas testados¹. A estratégia TCH_t desempenha uma pressão seletiva *relaxada* em relação a PBI com $\theta > 1.0$,

¹ Com a expressão “taxa de aproximação da fronteira Pareto-ótima de 100%” deseja-se dizer que em nenhuma das 30 execuções o método MOEA/D-tch_t ficou estagnado em fronteiras sub-ótimas. Esse fato pode ser observado a partir das Figs. 6–10, conforme discussão apresentada no capítulo anterior.

e embora esse fato tenha comprometido a convergência das estimativas para os pontos de \mathcal{PF} , possivelmente contribuiu para uma maior promoção de diversidade entre as soluções obtidas. Além disso, a abordagem sugerida dominou a técnica Tchebycheff original em todas as situações analisadas, justificando assim sua aplicação em MOEAs baseados em decomposição.

5.2 Propostas de Continuidade

Algumas possíveis propostas para a continuidade desta pesquisa envolvem os seguintes tópicos:

1. Avaliação mais apropriada da técnica *Multiple-Layer Simplex-Lattice Design* considerando-se dimensões mais elevadas do espaço de objetivos e, especificamente, valores $H < n_f$. Investigação do efeito do parâmetro de contração τ na distribuição dos vetores de ponderação. Comparação desta abordagem com a estratégia *Uniform Design* (TAN et al., 2012), também proposta com o intuito de obter uma distribuição mais uniforme dos vetores de ponderação.
2. Estudo do comportamento da função de agregação Tchebycheff transformada considerando-se uma maior variedade de problemas teste.
3. Desenvolvimento de estratégias promissoras relacionadas aos demais blocos componentes da estrutura do MOEA/D. A literatura tem investido muito nessa linha de pesquisa, e cabem ainda muitos avanços na área de algoritmos evolutivos multiobjetivo baseados em decomposição.

Referências

- AGUIRRE, H.; TANAKA, K. Many-objective optimization by space partitioning and adaptive ϵ -ranking on MNK-landscapes. In: *Proc. of EMO*. Nantes, France: [s.n.], 2009. (LNCS, v. 5467), p. 407–422. Citado na página 20.
- ASAFUDDOULA, M.; RAY, T.; SARKER, R. A decomposition based evolutionary algorithm for many objective optimization. *IEEE Transactions On Evolutionary Computation*, p. 1–17, 2014. Citado 3 vezes nas páginas 32, 34 e 35.
- BADER, J.; ZITZLER, E. HyPE: An algorithm for fast hypervolume-based many-objective optimization. *Evolutionary Computation*, v. 19, n. 1, p. 45–76, 2011. Citado 3 vezes nas páginas 20, 25 e 46.
- BATISTA, L. S. et al. A comparison of dominance criteria in many-objective optimization problems. In: *Proc. of CEC*. New Orleans, LA: [s.n.], 2011. p. 2359–2366. Citado 2 vezes nas páginas 20 e 24.
- CAMPELO, F.; BATISTA, L. S.; ARANHA, C. A component-wise perspective on multi-objective evolutionary algorithms based on decomposition. [*Under review*], 2017. Citado 2 vezes nas páginas 23 e 61.
- CHAN, L. Y. Optimal designs for experiments with mixtures: A survey. *Communications in Statistics: Theory and Methods*, v. 29, n. 9–10, p. 2281–2312, 2000. Citado 2 vezes nas páginas 20 e 26.
- CHIANG, T.; LAI, Y. MOEA/D-AMS: Improving MOEA/D by an adaptive mating selection mechanism. In: *IEEE Congress on Evolutionary Computation (CEC)*. [S.l.: s.n.], 2011. p. 1473–1480. Citado 2 vezes nas páginas 31 e 32.
- COELLO COELLO, C. A.; LAMONT, G. B.; van VELDHUIZEN, D. A. *Evolutionary Algorithms for Solving Multi-Objective Problems*. 2nd. ed. New York: Springer, 2007. Citado 2 vezes nas páginas 19 e 24.
- CORNELL, J. A. Experiments with mixtures: A review. *Technometrics*, v. 15, n. 3, p. 437–455, 1973. Citado na página 28.
- DAS, I.; DENNIS, J. E. Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems. *SIAM Journal of Opt.*, v. 8, n. 3, p. 631–657, 1998. Citado na página 29.
- DEB, K. *Multi-Objective Optimization Using Evolutionary Algorithms*. New York, NY, USA: John Wiley & Sons, Inc., 2001. Citado na página 24.
- DEB, K.; BEYER, H. Self-adaptive genetic algorithms with simulated binary crossover. *Evolutionary Computation*, v. 9, n. 2, p. 197–221, 2001. Citado na página 33.
- DEB, K.; JAIN, H. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, Part I: Solving problems with box constraints. *IEEE Transactions on Evolutionary Computation*, v. 18, n. 4, p. 577–601, 2014. Citado 4 vezes nas páginas 37, 41, 45 e 47.

- DEB, K. et al. Scalable test problems for evolutionary multiobjective optimization. In: ABRAHAM, A.; JAIN, L.; GOLDBERG, R. (Ed.). *Evolutionary Multiobjective Optimization*. [S.l.]: Springer London, 2005, (Advanced Information and Knowledge Processing). p. 105–145. Citado 3 vezes nas páginas 20, 45 e 46.
- DEB, K. D. D. Analysing mutation schemes for real-parameter genetic algorithms. *International Journal of Artificial Intelligence and Soft Computing*, v. 4, n. 1, p. 1–28, 2014. Citado na página 33.
- di Pierro, F.; KHU, S.-T.; SAVIC, D. A. An investigation on preference order ranking scheme for multiobjective evolutionary optimization. *IEEE Transactions on Evolutionary Computation*, v. 11, n. 1, p. 17–45, 2007. Citado na página 25.
- DRECHSLER, N.; DRECHSLER, R.; BECKER, B. Multi-objective optimisation based on relation favour. In: *Proceedings of the 1st International Conference on Evolutionary Multi-Criterion Optimization (EMO)*. Zurich, Switzerland: [s.n.], 2001. (Lecture Notes in Computer Science, v. 1993), p. 154–166. Citado na página 24.
- FANG, K. T.; LIN, D. K. J. *Uniform designs and their application in industry*. [S.l.]: Elsevier, 2003. v. 22. (Handbook of Statistics, v. 22). 131–170. Citado na página 27.
- FANG, K. T.; WANG, Y. *Number-Theoretic Methods in Statistics*. London: Chapman and Hall, 1993. v. 51. (Monographs on Statistics and Applied Probability, v. 51). Citado na página 28.
- FARINA, M.; AMATO, P. On the optimal solution definition for many-criteria optimization problems. In: *Proceedings of the North American Fuzzy Information Processing Society (NAFIPS)*. [S.l.: s.n.], 2002. p. 233–238. Citado 2 vezes nas páginas 20 e 24.
- FLEMING, P. J.; PURSHOUSE, R. C.; LYGOE, R. J. Many-objective optimization: An engineering design perspective. In: *Proc. of EMO*. Guanajuato, Mexico: Springer, 2005. (LNCS), p. 14–32. Citado na página 19.
- HUGHES, E. J. Msops-ii: A general-purpose many-objective optimiser. In: *IEEE Congress on Evolutionary Computation, 2007 (CEC)*. Singapore: [s.n.], 2007. p. Sept. Citado 2 vezes nas páginas 20 e 25.
- ISHIBUCHI, H.; AKEDO, N.; NOJIMA, Y. Relation between neighborhood size and MOEA/D performance on many-objective problems. In: *Evolutionary Multi-Criterion Optimization (EMO)*. [S.l.]: Springer, 2013. (Lecture Notes in Computer Science, v. 7811), p. 459–474. Citado na página 31.
- ISHIBUCHI, H. et al. Effects of using two neighborhood structures on the performance of cellular evolutionary algorithms for many-objective optimization. In: *IEEE Congress on Evolutionary Computation*. Trondheim, Norway: [s.n.], 2009. p. 2508–2515. Citado na página 31.
- ISHIBUCHI, H.; TSUKAMOTO, N.; NOJIMA, Y. Evolutionary many-objective optimization: A short review. In: *Proc. of CEC*. Hong Kong: [s.n.], 2008. p. 2419–2426. Citado na página 19.

- KRUISSELBRINK, J. W. et al. Combining aggregation with Pareto optimization: A case study in evolutionary molecular design. In: *Proceedings of the 5th International Conference on Evolutionary Multi-Criterion Optimization (EMO)*. Nantes, France: [s.n.], 2009. (Lecture Notes in Computer Science, v. 5467), p. 453–467. Citado na página 25.
- LI, H.; ZHANG, Q. Multiobjective optimization problems with complicated Pareto sets, MOEA/D and NSGA-II. *IEEE Transactions on Evolutionary Computation*, v. 13, n. 2, p. 284–302, 2009. Citado 7 vezes nas páginas 25, 26, 31, 32, 34, 35 e 47.
- LI, K. et al. An evolutionary many-objective optimization algorithm based on dominance and decomposition. *IEEE Transactions On Evolutionary Computation*, p. 1–23, 2014. Citado 3 vezes nas páginas 32, 42 e 46.
- LI, Y.; ZHOU, A.; ZHANG, G. An MOEA/D with multiple differential evolution mutation operators. In: *IEEE Congress on Evolutionary Computation (CEC)*. Beijing, China: [s.n.], 2014. p. 397–404. Citado na página 34.
- LÓPEZ, L. S.; BATISTA, L. S. Uma estratégia de decomposição Tchebycheff transformada para otimização evolucionária multiobjetivo. In: *XII Congresso Brasileiro de Inteligência Computacional (CBIC)*. [S.l.: s.n.], 2015. v. 1, p. 1–6. Citado na página 61.
- LUCKEN, C. von; BARÁN, B.; BRIZUELA, C. A survey on multi-objective evolutionary algorithms for many-objective problems. *Computational Optimization and Applications*, v. 58, n. 3, p. 707–756, 2014. Citado 2 vezes nas páginas 19 e 24.
- MIETTINEN, K. *Nonlinear Multiobjective Optimization*. Boston: Kluwer Academic Publishers, 1999. Citado 5 vezes nas páginas 23, 24, 28, 29 e 36.
- MURATA, T.; TAKI, A. Many-objective optimization for knapsack problems using correlation-based weighted sum approach. In: *Proc. of EMO*. Nantes, France: [s.n.], 2009. (LNCS, v. 5467), p. 468–480. Citado 2 vezes nas páginas 20 e 25.
- PRICE, K. V.; STORN, R. M.; LAMPINEN, J. A. *Differential Evolution: A Practical Approach to Global Optimization*. 1st. ed. [S.l.]: Springer, 2005. (Natural Computing Series). Citado na página 34.
- PURSHOUSE, R. C.; FLEMING, P. J. On the evolutionary optimization of many conflicting objectives. *IEEE Transactions On Evolutionary Computation*, v. 11, n. 6, p. 770–784, 2007. Citado na página 19.
- QI, Y. et al. MOEA/D with adaptive weight adjustment. *Evolutionary Computation*, v. 22, n. 2, p. 231–264, 2013. Citado 2 vezes nas páginas 44 e 61.
- SATO, H. Inverted PBI in MOEA/D and its impact on the search performance on multi and many-objective optimization. In: *Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation (GECCO)*. Vancouver, BC, Canada: [s.n.], 2014. p. 645–652. Citado na página 30.
- SATO, H.; AGUIRRE, H. E.; TANAKA, K. Pareto partial dominance MOEA and hybrid archiving strategy included CDAS in many-objective optimization. In: *Proc. of CEC*. Barcelona: [s.n.], 2010. p. 1–8. Citado 2 vezes nas páginas 20 e 24.

- SAXENA, D. K. et al. Objective reduction in many-objective optimization: Linear and nonlinear algorithms. *IEEE Trans. Evol. Comput.*, v. 17, n. 1, p. 77–99, 2013. Citado 2 vezes nas páginas 20 e 25.
- SCHEFFÉ, H. Experiments with mixtures. *Journal of Royal Statistical Society, Series B*, v. 20, n. 2, p. 344–360, 1958. Citado 3 vezes nas páginas 20, 26 e 28.
- SCHEFFÉ, H. Simplex-centroid designs for experiments with mixtures. *Journal of Royal Statistical Society, Series B*, v. 25, n. 2, p. 235–263, 1963. Citado na página 28.
- SINGH, H. K.; ISAACS, A.; RAY, T. A Pareto corner search evolutionary algorithm and dimensionality reduction in many-objective optimization problems. *IEEE Transactions on Evolutionary Computation*, v. 15, n. 4, p. 539–556, 2011. Citado na página 36.
- SULFLOW, A.; DRECHSLER, N.; DRECHSLER, R. Robust multi-objective optimization in high dimensional spaces. In: *Proceedings of the 4th International Conference on Evolutionary Multi-Criterion Optimization (EMO)*. Matsushima, Japan: [s.n.], 2007. (Lecture Notes in Computer Science, v. 4403), p. 715–726. Citado na página 24.
- TAN, Y. et al. A modification to MOEA/D-DE for multiobjective optimization problems with complicated Pareto sets. *Information Sciences*, v. 213, p. 14–38, 2012. Citado 4 vezes nas páginas 27, 32, 34 e 62.
- TAN, Y. et al. MOEA/D + uniform design: A new version of MOEA/D for optimization problems with many objectives. *Computers & Operations Research*, v. 40, n. 6, p. 1648–1660, 2013. Citado 2 vezes nas páginas 27 e 45.
- WANG, Z. et al. A replacement strategy for balancing convergence and diversity in MOEA/D. In: *IEEE Congress on Evolutionary Computation (CEC)*. Beijing, China: [s.n.], 2014. p. 2132–2139. Citado 2 vezes nas páginas 34 e 35.
- ZHANG, Q.; LI, H. MOEA/D: A multiobjective evolutionary algorithm based on decomposition. *IEEE Transactions On Evolutionary Computation*, v. 11, n. 6, p. 712–731, 2007. Citado 9 vezes nas páginas 20, 25, 26, 29, 31, 32, 34, 36 e 47.
- ZITZLER, E.; KUNZLI, S. Indicator-based selection in multiobjective search. In: *Proceedings of the 8th International Conference on Parallel Problem Solving from Nature (PPSN VIII)*. Birmingham, UK: Springer, 2004. (Lecture Notes in Computer Science, v. 3242), p. 832–842. Citado na página 25.
- ZITZLER, E.; THIELE, L. Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach. *IEEE Transactions on Evolutionary Computation*, v. 3, n. 4, p. 257–271, 1999. Citado 2 vezes nas páginas 20 e 46.
- ZITZLER, E. et al. Performance assessment of multiobjective optimizers: An analysis and review. *IEEE Trans. Evol. Comput.*, v. 7, n. 2, p. 117–132, 2003. Citado 2 vezes nas páginas 20 e 45.