

CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS  
DEL INSTITUTO POLITÉCNICO NACIONAL

Laboratorio de Tecnologías de Información

**Optimización de problemas con más  
de tres objetivos mediante  
algoritmos evolutivos**

Tesis que presenta:

**Mario Garza Fabre**

Para obtener el grado de:

**Maestro en Ciencias  
en Computación**

Director de la Tesis:  
Dr. Gregorio Toscano Pulido



© Derechos reservados por  
Mario Garza Fabre  
2009



Esta investigación fue parcialmente financiada mediante el proyecto No. 51623 del Fondo Mixto Conacyt-Gobierno del Estado de Tamaulipas.

This research was partially funded by project number 51623 from “Fondo Mixto Conacyt-Gobierno del Estado de Tamaulipas”



La tesis presentada por Mario Garza Fabre fue aprobada por:

-----

---

Dr. Arturo Díaz Pérez

---

Dr. Carlos A. Coello Coello

---

Dr. Gregorio Toscano Pulido, Director

Cd. Victoria, Tamaulipas, México., 28 de Septiembre de 2009





*A mis padres, por ser un gran ejemplo y por su confianza y apoyo incondicional en todo momento;  
y a Gisela, mi esposa, mi principal motivación para dar lo mejor de mí en esta etapa de formación.*



# Agradecimientos

- Agradezco primeramente a Dios por la vida, la salud y todas las bendiciones con las que ha llenado cada momento de mi existencia.
- Agradezco a mis padres y hermanos por su apoyo incondicional durante toda mi vida y formación académica.
- Agradezco a mi esposa Gisela quien ha sido mi fuente de inspiración y principal motivación para dar lo mejor de mi en esta etapa tan importante.
- En especial agradezco a mi asesor, el Dr. Gregorio Toscano Pulido, por su experiencia y acertada dirección durante este proyecto de tesis.
- Agradezco muy especialmente a mis revisores, el Dr. Arturo Díaz Pérez y el Dr. Carlos A. Coello Coello, por el tiempo dedicado al evaluar este documento y por los valiosos consejos y observaciones para mejorarlo.
- Gracias a los Investigadores adscritos al Laboratorio de Tecnologías de Información (Cinvestav - Tamaulipas) por el conocimiento impartido y por el apoyo brindado durante esta etapa de formación.
- Quiero agradecer al personal administrativo y técnico del Laboratorio de Tecnologías de Información (Cinvestav - Tamaulipas) por todas las facilidades brindadas durante mi estancia en la Institución.
- Agradezco el apoyo económico otorgado por CONACyT a través de una beca escolar durante los dos años de mi estancia en el programa de maestría.
- Agradezco el apoyo recibido a través del proyecto CONACyT 90548, titulado “Uso de técnicas híbridas para resolver problemas de optimización multiobjetivo”, cuyo Investigador principal es el Dr. Gregorio Toscano Pulido.



# Índice General

<b>Índice General</b>	<b>I</b>
<b>Índice de Figuras</b>	<b>v</b>
<b>Índice de Tablas</b>	<b>vii</b>
<b>Índice de Algoritmos</b>	<b>ix</b>
<b>Publicaciones</b>	<b>xi</b>
<b>Resumen</b>	<b>xiii</b>
<b>Abstract</b>	<b>xv</b>
<b>Nomenclatura</b>	<b>xvii</b>
<b>1. Introducción</b>	<b>1</b>
1.1. Introducción . . . . .	1
1.2. Motivación . . . . .	3
1.3. Objetivos . . . . .	3
1.4. Contribuciones esperadas . . . . .	4
1.5. Organización de la tesis . . . . .	4
<b>2. Antecedentes</b>	<b>7</b>
2.1. Optimización Global . . . . .	7
2.2. Optimización Multiobjetivo . . . . .	9
2.3. Algoritmos Evolutivos . . . . .	13
2.3.1. Inspiración biológica . . . . .	13
2.3.2. Algoritmo Evolutivo Generalizado . . . . .	14
2.3.3. Algoritmos Evolutivos para Optimización Multiobjetivo . . . . .	16
2.4. Optimización de muchos objetivos . . . . .	17
2.4.1. Introducción . . . . .	17
2.4.2. Inconvenientes de la dominancia de Pareto . . . . .	18
2.4.3. Alternativas de solución . . . . .	20
2.5. Problemas de Prueba . . . . .	21
2.5.1. DTLZ1 . . . . .	22
2.5.2. DTLZ2 . . . . .	23
2.5.3. DTLZ3 . . . . .	24
2.5.4. DTLZ4 . . . . .	25
2.5.5. DTLZ5 . . . . .	26

2.5.6.	DTLZ6 . . . . .	26
2.6.	Métricas . . . . .	27
2.6.1.	Métrica de Convergencia . . . . .	27
2.6.2.	Distancia Generacional Invertida . . . . .	28
<b>3.</b>	<b>Métodos de asignación de aptitud para optimización multiobjetivo</b>	<b>31</b>
3.1.	Introducción . . . . .	31
3.2.	Enfoques tomados de la literatura . . . . .	33
3.2.1.	Métodos basados en la dominancia de Pareto . . . . .	33
3.2.2.	Composiciones jerárquicas . . . . .	34
3.2.3.	Formas relajadas de dominancia . . . . .	37
3.2.3.1.	Métodos que requieren ajuste de parámetros . . . . .	37
3.2.3.2.	Métodos que no requieren ajuste de parámetros . . . . .	40
3.3.	Enfoques propuestos . . . . .	47
3.3.1.	Composiciones jerárquicas . . . . .	47
3.3.2.	Formas relajadas de dominancia . . . . .	49
3.4.	No equivalencia entre <b>AVER</b> , <b>WSCR</b> y <b>GFLS*</b> . . . . .	55
<b>4.</b>	<b>Métodos de asignación de aptitud: Experimentación</b>	<b>57</b>
4.1.	Introducción . . . . .	57
4.2.	Distribución de jerarquías . . . . .	58
4.3.	Tiempo de cómputo . . . . .	62
4.4.	Incorporación de los métodos en un algoritmo base . . . . .	64
4.4.1.	Habilidad de convergencia . . . . .	65
4.4.2.	Distancia generacional invertida . . . . .	72
4.4.3.	Convergencia en línea . . . . .	78
4.4.4.	Síntesis de resultados . . . . .	84
4.5.	Conclusiones del capítulo . . . . .	88
<b>5.</b>	<b>Algoritmos Propuestos</b>	<b>91</b>
5.1.	Introducción . . . . .	91
5.2.	Algoritmo base . . . . .	92
5.3.	Algoritmo genético elitista basado en agrupamiento . . . . .	93
5.3.1.	Jerarquización . . . . .	94
5.3.2.	Proceso de agrupamiento . . . . .	95
5.3.3.	Estrategia de selección . . . . .	96
5.3.4.	Justificación de ajuste de parámetros y operadores . . . . .	97
5.4.	Asignación de aptitud en direcciones múltiples . . . . .	98
5.4.1.	Jerarquización <b>MDFA*</b> . . . . .	98
5.4.2.	Justificación de ajuste de parámetros y operadores . . . . .	100
5.5.	Experimentación . . . . .	101
5.5.1.	Algoritmos del estado del arte tomados para realizar la comparativa . . . . .	102
5.5.1.1.	<b>NSGA-II</b> . . . . .	102

5.5.1.2.	<b>DMO</b>	103
5.5.1.3.	<b>HypE</b>	104
5.5.1.4.	<b>MSOPS</b>	105
5.5.2.	<b>Resultados</b>	107
5.5.2.1.	Convergencia	108
5.5.2.2.	Distancia generacional invertida	111
5.6.	Conclusiones del capítulo	115
<b>6.</b>	<b>Conclusiones y trabajo futuro</b>	<b>117</b>
6.1.	Conclusiones	117
6.2.	Contribuciones	119
6.3.	Trabajo futuro	120
	<b>Bibliografía</b>	<b>123</b>





# Índice de Figuras

2.1. Dominio de los problemas de optimización multiobjetivo. . . . .	12
2.2. Proporción de soluciones no dominadas en relación con el número de objetivos. . . .	19
2.3. Aproximación al frente de Pareto para el problema DTLZ1 con 3 objetivos. . . . .	23
2.4. Aproximación al frente de Pareto para los problemas DTLZ2-DTLZ4 con 3 objetivos.	24
2.5. Aproximación al frente de Pareto para los problemas DTLZ5 y DTLZ6 con 3 objetivos.	25
3.1. Ordenamiento por no dominancia ( <i>Nondominated Sorting</i> ). . . . .	34
3.2. Punto GBEST: Mejor solución conocida. . . . .	52
3.3. Punto LWORST: Actual peor solución. . . . .	53
4.1. Entropía promedio de la distribución de jerarquías para una población de 1000 solu- ciones generadas aleatoriamente. Problemas DTLZ1-DTLZ6. . . . .	60
4.2. Tiempo de cómputo al jerarquizar 100 veces una población de 100 soluciones. . . .	63
4.3. Flujo de ejecución del MOEA implementado. . . . .	64
4.4. Interpretación de los mosaicos de datos. . . . .	79
4.5. Convergencia en línea. Problema DTLZ1 con 5, 20 y 50 objetivos. . . . .	81
4.6. Convergencia en línea. Problema DTLZ2 con 5, 20 y 50 objetivos. . . . .	82
4.7. Convergencia en línea. Problema DTLZ3 con 5, 20 y 50 objetivos. . . . .	82
4.8. Convergencia en línea. Problema DTLZ4 con 5, 20 y 50 objetivos. . . . .	83
4.9. Convergencia en línea. Problema DTLZ5 con 5, 20 y 50 objetivos. . . . .	83
4.10. Convergencia en línea. Problema DTLZ6 con 5, 20 y 50 objetivos. . . . .	84
4.11. Síntesis de resultados: promediando los seis problemas de prueba adoptados y con- siderando por separado cada cantidad de objetivos. . . . .	86
4.12. Síntesis de resultados: promediando las seis cantidades de objetivos y considerando por separado cada problema de prueba adoptado. . . . .	87
5.1. Algoritmo base: Flujo de ejecución. . . . .	93
5.2. <b>CEGA*</b> : Flujo de ejecución. . . . .	94
5.3. Agrupamiento ( <i>Clustering</i> ) jerárquico aglomerativo. . . . .	95
5.4. <b>GBST*</b> aplicado localmente para elegir el individuo representante de cada <i>cluster</i> . .	96
5.5. <b>MDFA*</b> : Flujo de ejecución. . . . .	99
5.6. Interpretación de los diagramas de cajas (boxplots). . . . .	107
5.7. Métrica de convergencia. Problema DTLZ1. . . . .	109
5.8. Métrica de convergencia. Problema DTLZ2. . . . .	110
5.9. Métrica de convergencia. Problema DTLZ3. . . . .	110
5.10. Métrica de convergencia. Problema DTLZ4. . . . .	110
5.11. Métrica de convergencia. Problema DTLZ5. . . . .	111
5.12. Métrica de convergencia. Problema DTLZ6. . . . .	111
5.13. Métrica DGI. Problema DTLZ1. . . . .	113

5.14. Métrica DGI. Problema DTLZ2. . . . .	113
5.15. Métrica DGI. Problema DTLZ3. . . . .	114
5.16. Métrica DGI. Problema DTLZ4. . . . .	114
5.17. Métrica DGI. Problema DTLZ5. . . . .	114
5.18. Métrica DGI. Problema DTLZ6. . . . .	115

# Índice de Tablas

3.1. Equivalencia <b>AVER</b> , <b>WSCR</b> y <b>GFLS*</b> . . . . .	55
3.2. Equivalencia <b>AVER</b> , <b>WSCR</b> y <b>GFLS*</b> : Contraejemplo. . . . .	56
4.1. Media y desviación estándar de la métrica de convergencia. Problema DTLZ1. . . .	66
4.2. Media y desviación estándar de la métrica de convergencia. Problema DTLZ2. . . .	67
4.3. Media y desviación estándar de la métrica de convergencia. Problema DTLZ3. . . .	68
4.4. Media y desviación estándar de la métrica de convergencia. Problema DTLZ4. . . .	69
4.5. Media y desviación estándar de la métrica de convergencia. Problema DTLZ5. . . .	70
4.6. Media y desviación estándar de la métrica de convergencia. Problema DTLZ6. . . .	71
4.7. Media y desviación estándar de la métrica DGI. Problema DTLZ1. . . . .	73
4.8. Media y desviación estándar de la métrica DGI. Problema DTLZ2. . . . .	74
4.9. Media y desviación estándar de la métrica DGI. Problema DTLZ3. . . . .	75
4.10. Media y desviación estándar de la métrica DGI. Problema DTLZ4. . . . .	76
4.11. Media y desviación estándar de la métrica DGI. Problema DTLZ5. . . . .	77
4.12. Media y desviación estándar de la métrica DGI. Problema DTLZ6. . . . .	78



# Índice de Algoritmos

1.	Método <b>SODC</b> . . . . .	46
2.	Método <b>SCOR*</b> . . . . .	48
3.	Método <b>SCOO*</b> . . . . .	49
4.	Asignación de aptitud <b>MDFA*</b> . . . . .	100
5.	<b>NSGA-II</b> : Flujo de ejecución. . . . .	102
6.	<b>DMO</b> incorporado en <b>NSGA-II</b> : Flujo de ejecución. . . . .	104
7.	<b>HypE</b> : Flujo de ejecución. . . . .	105
8.	<b>MSOPS</b> : Comparación de dos soluciones utilizando la matriz $R$ . . . . .	106



# Publicaciones

Mario Garza Fabre, Gregorio Toscano Pulido and Carlos A. Coello Coello. *Alternative Fitness Assignment Methods for Many-Objective Optimization Problems*, in 9th International Conference on Artificial Evolution (EA 2009), Springer, Strasbourg, France, October 2009.

Mario Garza Fabre, Gregorio Toscano Pulido and Carlos A. Coello Coello. *Ranking Methods for Many-Objective Optimization*, in 8th Mexican International Conference on Artificial Intelligence (MICAI 2009), Springer, Guanajuato, Gto., México, November 2009.





## **Optimización de problemas con más de tres objetivos mediante algoritmos evolutivos**

por

**Mario Garza Fabre**

Maestro en Ciencias del Laboratorio de Tecnologías de Información  
Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, 2009  
Dr. Gregorio Toscano Pulido, Director

Los algoritmos evolutivos para optimización multiobjetivo han sido exitosamente utilizados para optimizar problemas que involucran dos o tres objetivos. La mayoría de estos enfoques están basados en dominancia de Pareto, que es la relación más ampliamente utilizada para comparar soluciones en el contexto multiobjetivo. Sin embargo, diferentes trabajos han mostrado que cuando se incrementa la cantidad de objetivos, las propuestas basadas en dominancia de Pareto pierden su potencial discriminante y consecuentemente sus resultados se deterioran significativamente.

En este trabajo de tesis se aborda el problema de la disminución de capacidades convergentes que presentan los algoritmos evolutivos ante el incremento de la cantidad de objetivos. En nuestro estudio analizamos el impacto discriminante y la capacidad para guiar el proceso de búsqueda hacia la superficie compromiso que presentan distintos métodos de asignación de aptitud en un algoritmo evolutivo multiobjetivo genérico. Dicho estudio contempló diferentes alternativas de asignación de aptitud del estado del arte y algunos métodos propuestos en esta tesis. Como consecuencia del estudio realizado, se propusieron dos nuevos algoritmos con énfasis en convergencia que mantienen la diversidad mediante distintos mecanismos con la finalidad de presentar al tomador de decisiones soluciones distribuidas en el frente de Pareto. Estos nuevos enfoques se desarrollaron con la finalidad de proveer alternativas robustas (ante distintos escenarios de optimización y la dimensionalidad de su espacio de objetivos). A través de un estudio empírico se mostró que nuestros resultados mejoran el desempeño de cuatro enfoques representativos de la literatura especializada.



## **Many-objective optimization through evolutionary algorithms**

by

**Mario Garza Fabre**

Master of Science from the Information Technology Laboratory  
Research Center for Advanced Study from the National Polytechnic Institute, 2009  
Dr. Gregorio Toscano Pulido, Advisor

Multiobjective evolutionary algorithms have been successfully used in order to optimize two-objective and three-objective problems. Pareto dominance has been the most commonly adopted relation to compare solutions in a multiobjective optimization context. However, it has been shown that as the number of objectives increases, the convergence ability of approaches based on Pareto dominance decreases.

This thesis tackles the evolutionary algorithms' diminishing convergence capabilities that arise as the number of objective functions increases. In our study, we analyze the discrimination impact and the ability to guide the search process to the trade-off surface of different fitness assignment methods when they are incorporated into a generic multiobjective evolutionary algorithm. In this study, we included several state-of-the-art fitness assignment alternatives and some novel approaches of this sort. As a result, we proposed two new algorithms with emphasis on convergence which maintain diversity through different mechanisms with the aim of presenting to the decision maker a well-spread set of non dominated solutions. These new approaches were developed with the aim of providing robust alternatives (at different optimization scenarios and the number of optimization criteria). Through an empirical study we showed that our results outperformed four representative approaches of the specialized literature.



# Nomenclatura

## Acrónimos principales

AE	Algoritmo Evolutivo
ADOM	Dominancia $\alpha$ ( $\alpha$ -dominance), método de jerarquización tomado de la literatura
AVER	Jerarquía promedio ( <i>average ranking</i> ), método de jerarquización tomado de la literatura
CEDM	Dominancia por contracción/expansión ( <i>contraction/expansion dominance</i> ), método de jerarquización tomado de la literatura
CEGA*	Algoritmo genético elitista basado en agrupamiento ( <i>Clustering-based Elitist Genetic Algorithm</i> ), algoritmo propuesto como parte de este proyecto de tesis
CJ	Composiciones jerárquicas, una clasificación para los métodos de asignación de aptitud estudiados
CONC	Composición condorcet ( <i>condorcet composition</i> ), método de jerarquización tomado de la literatura
DGI	Distancia Generacional Invertida, una de las métricas utilizadas
DMO	<i>Diversity Management Operator</i> , algoritmo tomado de la literatura utilizado para comparación
DTLZ	Familia de problemas de prueba, las siglas corresponden a las iniciales de sus autores: Deb, Thiele, Laumanns y Zitzler
EMO	Optimización Evolutiva Multiobjetivo ( <i>Evolutionary Multiobjective Optimization</i> )
EP	Programación Evolutiva ( <i>Evolutionary Programming</i> )
ER	Entropía Relativa, métrica utilizada para evaluar el potencial discriminante de los métodos de jerarquización estudiados
ES	Estrategias Evolutivas ( <i>Evolution Strategies</i> )
FAVR	Relación de preferencia ( <i>favour relation</i> ), método de jerarquización tomado de la literatura
FDOM	Relación de preferencia (FAVR) sin procedimiento para identificación de ciclos
FP	Frente de Pareto, o Superficie Compromiso
FPDM	Dominancia difusa ( <i>fuzzy pareto dominance</i> ), método de jerarquización tomado de la literatura
FRD	Formas Relajadas de Dominancia, una clasificación para los métodos de asignación de aptitud estudiados
GA	Algoritmo Genético ( <i>Genetic Algorithm</i> )
GAIN*	Ganancia total ( <i>total gain</i> ), método de jerarquización propuesto como parte de este proyecto de tesis

<b>GBST*</b>	Distancia a la mejor solución conocida ( <i>distance to the best known solution</i> ), método de jerarquización propuesto como parte de este proyecto de tesis
<b>GDET*</b>	Deterioro global ( <i>global detriment</i> ), método de jerarquización propuesto como parte de este proyecto de tesis
<b>GFAC</b>	Factor de ganancia ( <i>gaining factor</i> ), método de jerarquización tomado de la literatura
<b>GFLS*</b>	Fallos globales ( <i>global failures</i> ), método de jerarquización propuesto como parte de este proyecto de tesis
<b>GIMP*</b>	Mejora global ( <i>global improvement</i> ), método de jerarquización propuesto como parte de este proyecto de tesis
<b>GWIN*</b>	Victorias globales ( <i>global wins</i> ), método de jerarquización propuesto como parte de este proyecto de tesis
<b>GWST*</b>	Distancia a la peor solución global ( <i>distance to the global worst solution</i> ), método de jerarquización propuesto como parte de este proyecto de tesis
<b>HypE</b>	<i>Hypervolume Estimation Algorithm for Multiobjective Optimization</i> , algoritmo tomado de la literatura utilizado para comparación
<b>KDOM</b>	Dominancia $k$ ( <i>k-dominance</i> ), método de jerarquización tomado de la literatura
<b>LDOM</b>	Dominancia $L$ ( <i>L-dominance</i> ), método de jerarquización tomado de la literatura
<b>LWST*</b>	Distancia a la peor solución local ( <i>distance to the local worst solution</i> ), método de jerarquización propuesto como parte de este proyecto de tesis
<b>MAXR</b>	Máxima jerarquía ( <i>maximum ranking</i> ), método de jerarquización tomado de la literatura
<b>MDFA*</b>	Asignación de aptitud en direcciones múltiples ( <i>Multi-Directional Fitness Assignment</i> ), algoritmo propuesto como parte de este proyecto de tesis
<b>MEDC</b>	Composición mediana ( <i>median composition</i> ), método de jerarquización tomado de la literatura
<b>MEPS</b>	Dominancia del $\epsilon$ mínimo ( <i>minimum <math>\epsilon</math> dominance</i> ), método de jerarquización tomado de la literatura
<b>MINR*</b>	Mínima jerarquía ( <i>minimum ranking</i> ), método de jerarquización propuesto como parte de este proyecto de tesis
<b>MOEA</b>	Algoritmo Evolutivo para Optimización Multiobjetivo ( <i>Multiobjective Evolutionary Algorithm</i> )
<b>MSOPS</b>	<i>Multiple Single Objective Pareto Sampling</i> , algoritmo tomado de la literatura utilizado para comparación
<b>NSGA-II</b>	<i>Elitist Non-Dominated Sorting Genetic Algorithm</i> , algoritmo tomado de la literatura utilizado para comparación
<b>PD</b>	Dominancia de Pareto ( <i>Pareto Dominance</i> )
<b>PORD</b>	Jerarquización por orden de eficiencia ( <i>preference order ranking</i> ), método de jerarquización tomado de la literatura
<b>PRFT*</b>	Margen de ganancia ( <i>profit</i> ), método de jerarquización propuesto como parte de este proyecto de tesis
<b>PRNK</b>	Jerarquización de Pareto ( <i>Pareto ranking</i> ), método de jerarquización tomado de la literatura

<b>PSRT</b>	Ordenamiento por no dominancia ( <i>nondominated sorting</i> ), método de jerarquización tomado de la literatura
<b>RECC</b>	Composición recíproca ( <i>reciprocal composition</i> ), método de jerarquización tomado de la literatura
<b>SCOO*</b>	Comparación sucesiva de objetivos ordenados ( <i>successive comparison of ordered objectives</i> ), método de jerarquización propuesto como parte de este proyecto de tesis
<b>SCOR*</b>	Comparación sucesiva de jerarquías ordenadas ( <i>successive comparison of ordered ranks</i> ), método de jerarquización propuesto como parte de este proyecto de tesis
<b>SODC</b>	Combinación de <b>SVEC</b> y <b>MEPS</b> ( <i>sub-objective dominance count</i> ), método de jerarquización tomado de la literatura
<b>SVEC</b>	Dominancia sub-vector ( <i>subvector dominance</i> ), método de jerarquización tomado de la literatura
<b>TOFF*</b>	Jerarquización por compromiso ( <i>trade-off ranking</i> ), método de jerarquización propuesto como parte de este proyecto de tesis
<b>VDOM</b>	Dominancia por volumen ( <i>volume dominance</i> ), método de jerarquización tomado de la literatura
<b>WSCR</b>	Record de ganancia ( <i>winning score</i> ), método de jerarquización tomado de la literatura
<b>WSUM</b>	Suma ponderada ( <i>weighted sum</i> ), método de jerarquización tomado de la literatura

## Símbolos y notación

$f(\mathbf{X})$	Función objetivo
$\mathbf{F}(\mathbf{X})$	Función vectorial, o vector de funciones objetivo
$\mathcal{F}$	Región factible
$M$	Número de objetivos
$n$	Número de variables de decisión
$N$	Tamaño de la población (número de soluciones)
$P^*$	Conjunto de puntos de referencia en el espacio de las funciones objetivo
$P$	Población de soluciones candidatas
$\mathcal{U}$	Espacio de búsqueda
$\mathbf{X}$	Vector de diseño, o vector de variables de decisión
$x$	Variable de decisión
$Z$	Conjunto aproximación: soluciones encontradas por un optimizador multiobjetivo





# 1

## Introducción

### 1.1 Introducción

*La optimización es el acto de obtener los mejores resultados bajo determinadas circunstancias* [63]. La mayoría de los problemas del mundo real requieren satisfacer simultáneamente múltiples criterios que típicamente se encuentran en conflicto entre sí, de manera que al optimizar este tipo de problemas el propósito no es encontrar una única solución óptima, sino más bien un conjunto de soluciones compromiso. Desde la década de los 1950s, la comunidad de Investigación de Operaciones ha desarrollado enfoques para resolver esta clase de problemas. Sin embargo, la mayoría de los problemas multiobjetivo tienen un espacio de búsqueda muy grande y características (por ejemplo, no continuidad o no convexidad) que limitan la utilización de tales enfoques.

La naturaleza ha sido usada como fuente de inspiración para el desarrollo de diversos algoritmos computacionales. Los Algoritmos Evolutivos (AEs) son un ejemplo, pues simulan el proceso de la selección natural al mejorar progresivamente una población de soluciones a través de la aplicación iterativa de una serie de procesos probabilísticos. Su uso de una población los hace una alternativa idónea para resolver problemas multiobjetivo, ya que son capaces de explorar distintas regiones del

espacio de búsqueda y converger hacia diferentes soluciones compromiso en una sola ejecución. Sin embargo, no fue sino hasta mediados de la década de los 1980s que se reconoció el potencial de los AEs para optimizar problemas de este tipo; desde ese entonces ha surgido una amplia variedad de algoritmos evolutivos para optimización multiobjetivo (MOEAs, por sus siglas en Inglés) que han sido exitosamente utilizados por diversos investigadores [10].

A pesar del considerable volumen de investigación sobre Optimización Evolutiva Multiobjetivo (EMO, por sus siglas en Inglés), la mayor parte de los esfuerzos realizados se han limitado a resolver problemas con dos o tres objetivos, no obstante, existen muchos problemas multiobjetivo prácticos que involucran una mayor cantidad de objetivos. Recientemente se ha despertado interés en este tema, y se han realizado distintos estudios en los que se encontró que el desempeño de la mayoría de los MOEAs reportados en la literatura se deteriora significativamente al incrementarse la cantidad de objetivos [60, 42, 37, 58, 43, 40]. Tal es la problemática que se adoptó el término *Many-Objective* para referirse a aquellos problemas que contemplan muchos (más de tres) objetivos [28].

La principal causa de esta problemática se le atribuye a la dominancia de Pareto (PD, por sus siglas en Inglés), que es la relación más ampliamente adoptada para comparar soluciones al diseñar MOEAs. Tal como sucede en la naturaleza, los AEs guían el proceso de búsqueda al otorgarles un mayor potencial reproductivo y de supervivencia a aquellos individuos mejor adaptados; v.g., soluciones con el mejor desempeño. Sin embargo, PD pierde su potencial de discriminación al incrementarse la cantidad de objetivos debido a que la mayor parte de la población se compone de soluciones no dominadas (v.g., equivalentes para PD). Bajo estas condiciones, el proceso de guiar la búsqueda se lleva a cabo prácticamente de manera aleatoria y, por lo tanto, la capacidad de convergencia de los MOEAs se deteriora.

En la literatura pueden distinguirse dos enfoques principales para hacer frente a esta problemática:

1. La reducción de dimensiones: donde se analizan las relaciones entre los diferentes objetivos con la finalidad de ignorar criterios redundantes y, de este modo, transformar el problema en uno más pequeño y manejable. Sin embargo, la reducción de dimensiones no siempre es posible y depende totalmente del escenario de optimización.

2. Uso de métodos alternativos para establecer preferencias entre las soluciones candidatas.

En este trabajo de tesis se aborda el problema de la alta dimensionalidad del espacio de objetivos siguiendo el segundo enfoque. Hemos recopilado y documentado diferentes métodos de asignación de aptitud para optimización multiobjetivo que han sido reportados en la literatura especializada. Del mismo modo, como resultado de este trabajo de tesis se proponen algunas nuevas alternativas de esta índole. Se realizó un estudio en el que los diferentes enfoques (tanto los estudiados del estado del arte, así como los propuestos) se sometieron a una serie de experimentos para evaluar su comportamiento en términos de su capacidad para guiar el proceso de búsqueda. Para estos experimentos se consideraron diferentes problemas de prueba escalables con diferentes cantidades de objetivos: 5, 10, 15, 20, 30 y 50. Como resultado de este estudio, también proponemos dos nuevos MOEAs que demostraron estadísticamente ser robustos y competitivos cuando se compararon contra cuatro enfoques representativos de la literatura.

## 1.2 Motivación

La motivación principal de este trabajo es contribuir al área de EMO al proveer un estudio que permita un mejor entendimiento del problema de convergencia cuando se presenta una alta dimensionalidad en el espacio de objetivos. También, es de particular interés desarrollar enfoques que sean capaces de lidiar efectivamente con esta problemática.

## 1.3 Objetivos

Los principales objetivos fijados para este proyecto son los siguientes:

- Estudiar a fondo la problemática de convergencia ante el incremento del número de objetivos.
- Recopilar diferentes métodos de asignación de aptitud para optimización multiobjetivo y estudiar su habilidad para guiar efectivamente el proceso de búsqueda y su escalabilidad con respecto al número de objetivos.

- Proponer al menos un nuevo método de asignación de aptitud para optimización multiobjetivo que sea efectivo y escalable con respecto al número de objetivos.
- Proponer al menos un nuevo esquema de optimización que presente una convergencia robusta ante la alta dimensionalidad del espacio de objetivos y que sea competitivo con respecto a los enfoques representativos del estado del arte.
- Validar estadísticamente el desempeño de todas las implementaciones realizadas utilizando un conjunto adecuado de métricas y funciones de prueba tomadas de la literatura especializada.

## 1.4 Contribuciones esperadas

Las principales contribuciones esperadas de este trabajo de tesis son las siguientes:

- Un estudio comparativo donde se analicen empíricamente 1) la capacidad de un determinado MOEA para guiar el proceso de búsqueda al incorporar diferentes mecanismos de asignación de aptitud y 2) su comportamiento en relación con el número de objetivos.
- Al menos un nuevo método de asignación de aptitud para optimización multiobjetivo que presente una convergencia robusta ante incrementos en el número de objetivos a optimizar.
- Al menos un nuevo MOEA capaz de resolver problemas con alta dimensionalidad del espacio de objetivos y competitivo con respecto a los enfoques representativos del estado del arte.
- Un estudio comparativo de diferentes enfoques propuestos para resolver problemas con muchos objetivos, incluyendo los enfoques resultantes de este trabajo de investigación.

## 1.5 Organización de la tesis

El Capítulo 2 presenta la terminología y los conceptos fundamentales del área, plantea la problemática y describe las funciones de prueba y métricas utilizadas para la experimentación de este

---

proyecto. El Capítulo 3 describe los diferentes mecanismos de asignación de aptitud estudiados (tanto los tomados de la literatura como las nuevas propuestas) y los resultados de los experimentos realizados se presentan en el Capítulo 4. En el Capítulo 5 se describen los dos nuevos MOEAs propuestos como parte de este proyecto y se presentan los resultados de su evaluación. Finalmente, el Capítulo 6 provee las conclusiones finales de esta tesis y se describen algunas posibles direcciones para trabajo futuro.



# 2

## Antecedentes

### 2.1 Optimización Global

El método para buscar el óptimo global (que puede no ser único) de alguna función se conoce como *Optimización Global* [11].

La optimización puede definirse como el problema de encontrar un *vector de variables de decisión* tal que maximice la satisfacción de una *función objetivo* cumpliendo con una serie de restricciones. Antes de plantear formalmente esta definición es necesario presentar algunos conceptos importantes.

**Concepto 1 (Vector de diseño o vector de variables de decisión):** Las variables de decisión son los parámetros que deben ajustarse para dar solución a un problema de optimización. Un vector de variables de decisión se puede representar de la siguiente manera:

$$\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (2.1)$$

donde  $x_i$  denota la  $i$ -ésima de  $n$  variables de decisión. Para este trabajo se utilizará por conveniencia la siguiente representación:

$$\mathbf{X} = [x_1, x_2, \dots, x_n]^T \quad (2.2)$$

**Concepto 2 (Restricciones):** Criterios o requerimientos que una solución debe satisfacer para ser considerada una solución factible. De otro modo, cuando una solución no satisface todas las restricciones del problema se considera una solución infactible. Las restricciones pueden ser de desigualdad (Ecuación (2.3)) o de igualdad (Ecuación (2.4)):

$$g_i(\mathbf{X}) \leq 0 \quad i = 1, 2, \dots, m \quad (2.3)$$

$$h_j(\mathbf{X}) = 0 \quad j = 1, 2, \dots, p \quad (2.4)$$

**Concepto 3 (Función objetivo):** Es un criterio expresado en función de las variables de decisión y se utiliza para medir la calidad de las diferentes alternativas de solución para determinado problema de optimización. Una función objetivo será representada en este trabajo de tesis como  $f(\mathbf{X})$ , donde  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  y  $\mathbf{X}$  es un vector de variables de decisión.

**Concepto 4 (Espacio de búsqueda):** El espacio de búsqueda ( $\mathcal{U}$ ) corresponde a todo el universo de soluciones candidatas al problema.

**Concepto 5 (Región factible):** La región factible ( $\mathcal{F}$ ) es el subconjunto de todas las posibles soluciones que satisfacen las restricciones de desigualdad  $g(\mathbf{X})$  y las restricciones de igualdad  $h(\mathbf{X})$  descritas en las Ecuaciones (2.3) y (2.4), respectivamente. Formalmente:

$$\mathcal{F} = \{\mathbf{X} \in \mathcal{U} : g(\mathbf{X}) \leq 0 \wedge h(\mathbf{X}) = 0\} \quad (2.5)$$

Una vez planteados los conceptos anteriores, la optimización puede definirse formalmente como el problema de encontrar el vector  $\mathbf{X}^* = [x_1^*, x_2^*, \dots, x_n^*]^T$  que satisfaga las  $m$  restricciones de



desigualdad:

$$g_i(\mathbf{X}) \leq 0 \quad i = 1, 2, \dots, m \quad (2.6)$$

las  $p$  restricciones de igualdad:

$$h_j(\mathbf{X}) = 0 \quad j = 1, 2, \dots, p \quad (2.7)$$

y que minimice<sup>1</sup> la función:

$$f(\mathbf{X}) \quad (2.8)$$

donde  $\mathbf{X}$  es un vector de variables de decisión,  $f(\mathbf{X})$  es la función objetivo.

## 2.2 Optimización Multiobjetivo

Osyczka define la optimización multiobjetivo (también conocida como optimización multicriterio u optimización vectorial) de la siguiente manera [56]:

*“El problema de encontrar un vector de variables de decisión que satisfaga las restricciones y optimice una función vectorial cuyos elementos representen las funciones objetivo. Estas funciones forman una descripción matemática de los criterios de desempeño que usualmente están en conflicto entre sí. Por lo tanto, el término “optimizar” significa encontrar una solución tal que proporcione valores para todos los objetivos que resulten aceptables para el tomador de decisiones.”*

**Concepto 6 (Función vectorial o vector de funciones objetivo):** Cuando se optimizan múltiples objetivos las  $M \geq 2$  funciones objetivos pueden representarse mediante un vector de la siguiente

---

<sup>1</sup>Sin pérdida de generalización, en este documento haremos referencia únicamente a problemas de minimización.

manera:

$$\mathbf{F}(\mathbf{X}) = \begin{bmatrix} f_1(\mathbf{X}) \\ f_2(\mathbf{X}) \\ \vdots \\ f_M(\mathbf{X}) \end{bmatrix} \quad (2.9)$$

donde  $f_m(\mathbf{X})$  denota la  $m$ -ésima función objetivo. Por conveniencia se utilizará la siguiente representación:

$$\mathbf{F}(\mathbf{X}) = [f_1(\mathbf{X}), f_2(\mathbf{X}), \dots, f_M(\mathbf{X})]^T \quad (2.10)$$

Formalmente, la optimización multiobjetivo se puede definir como el problema de encontrar el vector  $\mathbf{X}^* = [x_1^*, x_2^*, \dots, x_n^*]^T$  que satisfaga las  $m$  restricciones de desigualdad:

$$g_i(\mathbf{X}) \leq 0 \quad i = 1, 2, \dots, m \quad (2.11)$$

las  $p$  restricciones de igualdad:

$$h_j(\mathbf{X}) = 0 \quad j = 1, 2, \dots, p \quad (2.12)$$

y que optimice la función vectorial:

$$\mathbf{F}(\mathbf{X}) = [f_1(\mathbf{X}), f_2(\mathbf{X}), \dots, f_M(\mathbf{X})]^T \quad (2.13)$$

donde  $\mathbf{X} = [x_1, x_2, \dots, x_n]^T$  es el vector de  $n$  variables de decisión, y  $M$  es el número de funciones objetivo ( $M \geq 2$ ). En otras palabras, deseamos determinar de entre el conjunto  $\mathcal{F}$  de todos los valores que satisfacen (2.11) y (2.12) al conjunto  $x_1^*, x_2^*, \dots, x_n^*$  que aporte los valores óptimos para todas las funciones objetivo.

La noción de optimalidad es diferente en optimización multiobjetivo, ya que por lo general no existe una única solución que pueda ser considerada óptima. Por el contrario, se pretende encontrar los mejores compromisos posibles entre los objetivos. La noción de optimalidad generalmente adoptada

es la de *óptimo de Edgeworth-Pareto*, u *óptimo de Pareto* como es más comúnmente conocida [57]. Esta noción de optimalidad se presenta a continuación.

**Concepto 7 (Dominancia de Pareto):** Dadas dos soluciones  $\mathbf{X}_i, \mathbf{X}_j \in \mathcal{F}$ , decimos que  $\mathbf{X}_i$  domina a  $\mathbf{X}_j$  de acuerdo con la dominancia de Pareto ( $\mathbf{X}_i \prec_P \mathbf{X}_j$ ) si y sólo si:

$$\begin{aligned} \forall m \in \{1, 2, \dots, M\} : f_m(\mathbf{X}_i) &\leq f_m(\mathbf{X}_j) \quad \wedge \\ \exists m \in \{1, 2, \dots, M\} : f_m(\mathbf{X}_i) &< f_m(\mathbf{X}_j) \end{aligned} \quad (2.14)$$

De otro modo, decimos que  $\mathbf{X}_j$  es una solución no dominada con respecto a  $\mathbf{X}_i$ .

**Concepto 8 (Óptimo de Pareto):** Un vector solución es considerado *óptimo de Pareto* si no existe otro vector factible que mejore la satisfacción de alguno de los objetivos sin provocar simultáneamente un deterioro en algún otro criterio. Más formalmente, un vector de variables de decisión  $\mathbf{X}^* \in \mathcal{F}$  es un óptimo de Pareto si no existe otro vector  $\mathbf{X} \in \mathcal{F}$  tal que  $\mathbf{X}$  domine a  $\mathbf{X}^*$  de acuerdo con la dominancia de Pareto ( $\mathbf{X} \prec_P \mathbf{X}^*$ ).

**Concepto 9 (Conjunto de óptimos de Pareto):** El conjunto de óptimos de Pareto  $PS^*$  es un conjunto de puntos en el espacio de las variables de decisión que poseen las características descritas en el Concepto 8. Formalmente:

$$PS^* = \{\mathbf{X}^* \in \mathcal{F} : \neg \exists \mathbf{X} \in \mathcal{F} : \mathbf{X} \prec_P \mathbf{X}^*\} \quad (2.15)$$

**Concepto 10 (Frente de Pareto o superficie compromiso):** El conjunto de óptimos de Pareto representado en el espacio de las funciones objetivo constituye lo que se conoce como superficie compromiso o frente de Pareto (FP):

$$FP = \{\mathbf{F}(\mathbf{X}^*) : \mathbf{X}^* \in PS^*\} \quad (2.16)$$

Los conceptos y definiciones anteriores constituyen el dominio de los problemas de optimización multiobjetivo y pueden visualizarse en la figura 2.1.

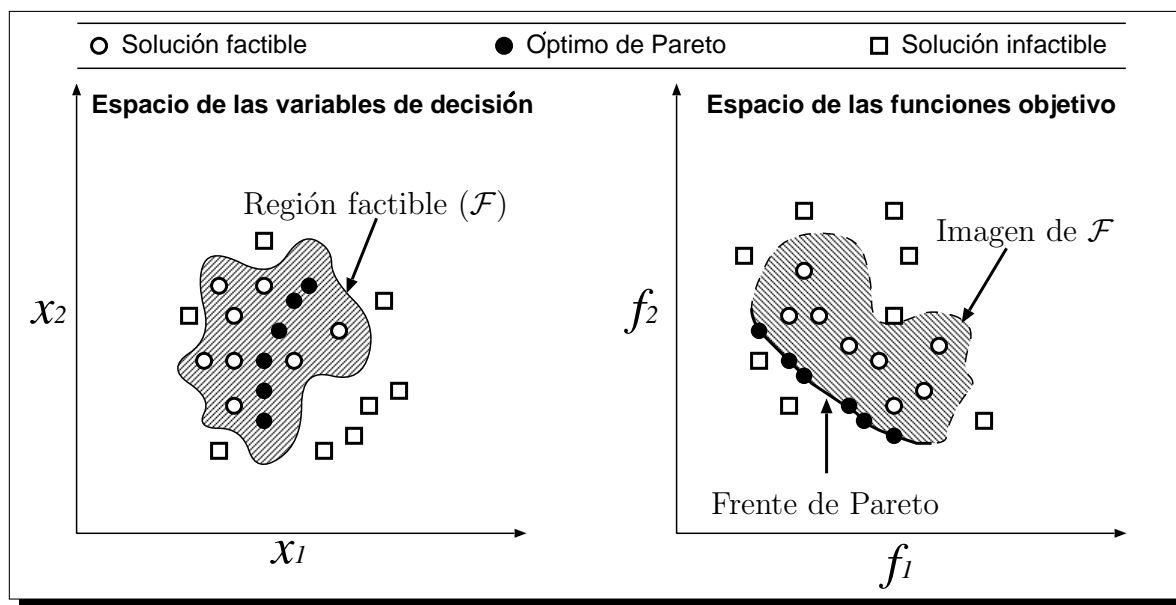


Figura 2.1: Dominio de los problemas de optimización multiobjetivo.

Existen distintos aspectos importantes al optimizar múltiples objetivos, pero antes es necesario introducir algunos conceptos.

**Concepto 11 (Conjunto aproximación):** Los optimizadores multiobjetivo por lo general intentan encontrar una aproximación al frente de Pareto (PF) para determinado problema. En este trabajo por conjunto aproximación ( $Z$ ) nos referiremos al conjunto de vectores de objetivos no dominados encontrados por un optimizador multiobjetivo.

**Concepto 12 (Tomador de decisiones):** Como ya se mencionó en la Sección 2.2, resolver un problema de optimización multiobjetivo consiste en encontrar un conjunto de soluciones compromiso, el cual puede contener un número infinito de alternativas. Sin embargo, generalmente se tendrá que elegir una de esas soluciones como curso de acción para hacer frente a dicho problema. Todas las soluciones que cumplen con las características de un óptimo de Pareto son matemáticamente equivalentes, por lo que elegir alguna de las alternativas requiere de información que no está contenida en las funciones objetivo. En este sentido, el tomador de decisiones es la persona (o conjunto de personas) que se asume tendrá(n) un amplio conocimiento del problema y que expresará(n) relaciones de preferencia entre las diferentes soluciones para tomar una decisión final [54].

De acuerdo con Purshouse [62], la tarea de un optimizador multiobjetivo es proveer una representación precisa y útil de la superficie compromiso para el tomador de decisiones. Así mismo, Purshouse destacó tres aspectos que se buscan para determinar la calidad del conjunto aproximación  $Z$  obtenido por un optimizador multiobjetivo [62]:

- **Proximidad (convergencia):** El conjunto  $Z$  debe tener soluciones cuyos vectores de objetivos correspondientes sean cercanos al frente de Pareto.
- **Diversidad:** El conjunto  $Z$  debe contener una buena distribución de las soluciones, tanto en términos de extensión como de uniformidad. Una buena diversidad es comúnmente de interés en el espacio de objetivos, pero puede requerirse también en el espacio de las variables de decisión.
- **Pertinencia:** El conjunto  $Z$  debe contener únicamente soluciones que se encuentren en la región de interés del tomador de decisiones. En la práctica, y especialmente al incrementarse la cantidad de objetivos, el tomador de decisiones estará interesado sólo en una sub-región del espacio de objetivos.

## 2.3 Algoritmos Evolutivos

### 2.3.1 Inspiración biológica

La naturaleza ha sido fuente inagotable de inspiración para el desarrollo de una amplia variedad de algoritmos computacionales. Los Algoritmos Evolutivos (AEs) engloban un conjunto de técnicas que toman inspiración y simulan el proceso natural de la evolución. Más específicamente, los AEs se inspiran en la teoría evolutiva más aceptada actualmente: el *Neo-Darwinismo*<sup>2</sup>. El Neo-Darwinismo establece que la vida puede explicarse como la ocurrencia repetida de distintos procesos estadísticos que interactúan entre sí sobre las poblaciones:

---

<sup>2</sup>El paradigma Neo-Darwiniano es el resultado de la combinación de la teoría evolutiva de Darwin, el seleccionismo de Weismann y la genética de Mendel.

- **Reproducción:** Combinación de individuos y transferencia del código genético a su descendencia.
- **Mutación:** Errores en la transferencia de información genética durante la reproducción. Es un aspecto necesario para incluir variabilidad en las especies.
- **Competencia:** Consecuencia del crecimiento de la población en un entorno con recursos limitados.
- **Selección:** Los individuos mejor adaptados tendrán mayor oportunidad de sobrevivir y reproducirse.

La evolución por sí misma puede ser vista como un proceso de optimización, donde las diferentes especies tienen que mejorar constantemente para adaptarse y competir en su entorno.

### 2.3.2 Algoritmo Evolutivo Generalizado

El término algoritmo evolutivo (AE) es generalmente utilizado para describir cualquier método de búsqueda estocástico basado en poblaciones. La idea fundamental de los AEs es evolucionar progresivamente una población de individuos (soluciones potenciales para el problema de optimización) a través de la aplicación iterativa de procesos probabilísticos de variación y selección hasta que algún criterio de terminación sea alcanzado (por lo regular un número máximo de iteraciones (generaciones)). Para simular el proceso evolutivo se requiere:

- **Representación:** una forma de codificar las soluciones candidatas para un problema dado.
- **Operadores de variación:** operadores que permitan la generación de nuevos individuos a partir de los ya existentes.
- **Función de aptitud:** una función que permita expresar la calidad de las soluciones candidatas.
- **Mecanismo de selección:** estrategia para seleccionar los individuos que podrán reproducirse de acuerdo con su aptitud.

Existen tres paradigmas principales: la Programación Evolutiva (*Evolutionary Programming*, EP) propuesta por Fogel [30], los Algoritmos Genéticos (*Genetic Algorithms*, GA) propuestos por Holland [34] y las Estrategias Evolutivas (*Evolution Strategies*, ES) propuestos por Rechenberg [64]. La principal diferencia entre estos paradigmas radica en la representación que utilizan para codificar las alternativas de solución y en los operadores de variación incorporados.

Purshouse describe la estructura general de un AE mediante la Ecuación 2.17 [62]:

$$P_{t+1} = S_s(V(S_v(P_t)), P_t) \quad (2.17)$$

donde  $P_t$  denota la población en el tiempo  $t$  (generación actual) y, por lo tanto,  $P_{t+1}$  representa la población para la siguiente generación.  $V$  es una función que representa los operadores de variación incorporados y las funciones  $S_v$  y  $S_s$  denotan los mecanismos de selección para variación (*selection-for-variation*) y selección para supervivencia (*selection-for-survival*), respectivamente. Estos tres últimos elementos serán descritos por separado a continuación:

- **Selección para variación (selection-for-variation),  $S_v$ :** Esta función se refiere a la etapa donde se seleccionan los individuos de  $P_t$  sobre los que se aplicarán los operadores de variación  $V$ . Para este proceso de selección se tomará en cuenta el valor de aptitud de las soluciones, de manera que los mejores individuos en la población generalmente tendrán una mayor oportunidad de reproducirse. Se asume que a partir de buenas soluciones es posible generar nuevas buenas soluciones.
- **Operadores de variación,  $V$ :** Los operadores de variación son los encargados de generar nuevos individuos a partir de los ya existentes. Estos operadores se clasifican de acuerdo con el número de *individuos padre* que se utilizan para generar nuevos *individuos hijo*. Cuando se utiliza un solo padre el proceso de variación se conoce usualmente como *mutación*, mientras que la combinación de múltiples padres se conoce generalmente como *cruza* o *recombinación*. Cada operador de variación tiene una probabilidad de aplicación asociada dependiendo del problema en cuestión.

- **Selección para supervivencia (selection-for-survival)**,  $S_g$ : Una vez aplicados los operadores de variación se tienen dos conjuntos de soluciones:  $P_t$  que denota la población original y  $V(S_v(P_t))$  que está formada por los individuos creados como consecuencia de los operadores de variación. Los AEs trabajan con poblaciones de tamaño finito  $N$ , por lo que en esta etapa son seleccionados los  $N$  individuos que sobrevivirán y formarán la población de padres para la siguiente generación.

### 2.3.3 Algoritmos Evolutivos para Optimización Multiobjetivo

Tal como se mencionó en las secciones anteriores, al optimizar múltiples objetivos estamos interesados en encontrar un conjunto de soluciones compromiso que aproximen de manera representativa el frente de Pareto. Dado este requerimiento, los AEs son una alternativa particularmente apropiada para resolver problemas multiobjetivo. Los AEs trabajan simultáneamente con una población de soluciones candidatas que les permite encontrar diferentes miembros del conjunto de óptimos de Pareto en una sola ejecución, en contraste con los enfoques tradicionales y algunos otras heurísticas que para lograrlo requieren llevar a cabo una serie de ejecuciones independientes. El uso de poblaciones también permite que los AEs exploren distintas regiones del espacio de búsqueda a la vez, con lo que evitan quedar atrapados en óptimos locales. Adicionalmente, la flexibilidad de los AEs los hace menos susceptibles a la forma y continuidad de la superficie compromiso, ya que pueden lidiar con frentes de Pareto desconectados y cóncavos.

En los 1960s, Rosenberg [65] estableció los primeros indicios de la utilización de algoritmos evolutivos para optimizar problemas multiobjetivo. Sin embargo, el problema con el cual trabajaba fue replanteado considerando un único objetivo. Fue hasta mediados de los 1980s cuando Schaffer realizara la primera implementación de un algoritmo evolutivo multiobjetivo (MOEA): el *Vector Evaluated Genetic Algorithm* (VEGA) [68].

En los 1990s el área de la optimización evolutiva multiobjetivo se afianzó y surgieron diversos enfoques importantes: el *Multi-Objective Genetic Algorithm* (MOGA) propuesto por Fonseca y Fleming [31]; el *Niched Pareto Genetic Algorithm* (NPGA) propuesto por Horn y Nafpliotis [35];



el *Nondominated Sorting Genetic Algorithm* (NSGA) propuesto por Srinivas y Deb [70]; y enfoques como el *Strength Pareto Evolutionary Algorithm* (SPEA) de Zitzler y Thiele [75] y el *Pareto Archived Evolution Strategy* (PEAS) de Knowles y Corne [45] que popularizaron el uso de elitismo en MOEAs.

Posteriormente, comenzó a surgir una nueva generación de MOEAs que incorporaron mecanismos más eficientes y sofisticados para mejorar la convergencia y/o preservación de diversidad. En esta nueva generación de MOEAs podemos destacar el *Elitist Non-Dominated Sorting Genetic Algorithm* (NSGA-II) propuesto por Deb *et al.* [17, 22], que se ha convertido en el MOEA más representativo de la literatura y ha sido utilizado ampliamente como punto de referencia para nuevos desarrollos; el *Pareto Envelope-based Selection Algorithm* (PESA) propuesto por Corne *et al.* [14] para incorporar ideas de SPEA y PAES, su nueva versión PESA-II [13] cuyo mecanismo de selección está basado en regiones; y SPEA2 de Zitzler *et al.* [74] que presenta un mecanismo de estimación de la densidad y un esquema mejorado y más estricto de asignación de aptitud.

Coello *et al.* [11] proporcionan una descripción más completa de cada uno de los enfoques citados.

## 2.4 Optimización de muchos objetivos

### 2.4.1 Introducción

A lo largo de la historia se han reportado diversos MOEAs (ver Sección 2.3.3), muchos de los cuales han demostrado ser competitivos al optimizar problemas con 2 o 3 objetivos manteniendo un número relativamente bajo de evaluaciones. Sin embargo, recientemente se ha reportado que el desempeño de la mayoría de estos enfoques se degrada significativamente con el aumento en el número de objetivos a optimizar [60, 42, 37, 58, 43, 40]. Tal es la problemática, que se propuso el término *Many-Objective* para referirse al subconjunto de problemas donde muchos (más de tres) objetivos requieren ser optimizados [28], término que se popularizó y se ha seguido utilizando en la literatura especializada.

La alta dimensionalidad del espacio de objetivos conlleva una amplia problemática [12, 40]:

- Proporción de soluciones no dominadas. Al ser incrementada la cantidad de objetivos, la

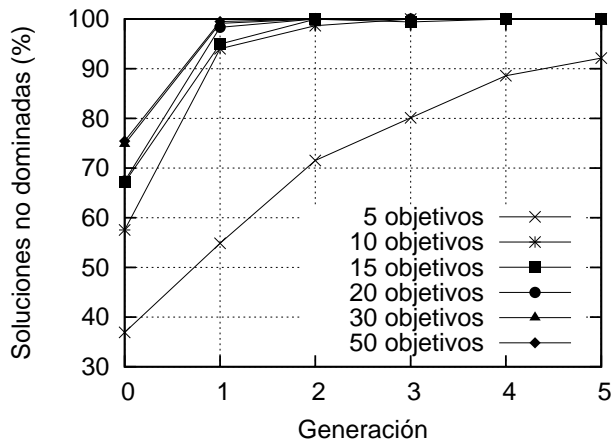
población se satura rápidamente de soluciones no dominadas, provocando el deterioro de la habilidad de búsqueda de MOEAs (ver Sección 2.4.2).

- Crecimiento exponencial en el número de soluciones requeridas para aproximar representativamente el frente de Pareto.
- Algunos MOEAs se basan en estructuras de datos y subrutinas cuyas demandas computacionales crecen exponencialmente (en espacio y tiempo, respectivamente) con el número de objetivos.
- Dificultad en la visualización de las soluciones.

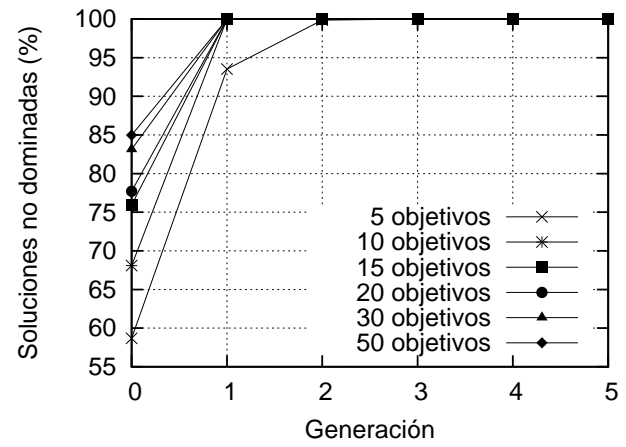
### 2.4.2 Inconvenientes de la dominancia de Pareto

El incremento en la cantidad de objetivos afecta principalmente a los enfoques que utilizan la dominancia de Pareto (PD) como criterio para discriminar entre soluciones y guiar el proceso de búsqueda (la gran mayoría de los MOEAs existentes se basan en este concepto). Esto se debe a que la proporción de soluciones no dominadas en la población se incrementa de manera drástica con el número de objetivos. Con el propósito de aclarar este punto, la figura 2.2 muestra cómo se comporta el porcentaje de soluciones no dominadas en la población de un MOEA en relación con el número de objetivos al avanzar el proceso de búsqueda. El MOEA utilizado se describe en la Sección 4.4. Para este experimento se utilizaron seis problemas de prueba escalables (DTLZ1-DTLZ6), que serán descritos en la Sección 2.5, y se utilizaron instancias con  $M = \{5, 10, 15, 20, 30, 50\}$  objetivos. Los datos graficados en la figura 2.2 corresponden a la media de 31 repeticiones independientes para cada configuración del experimento.

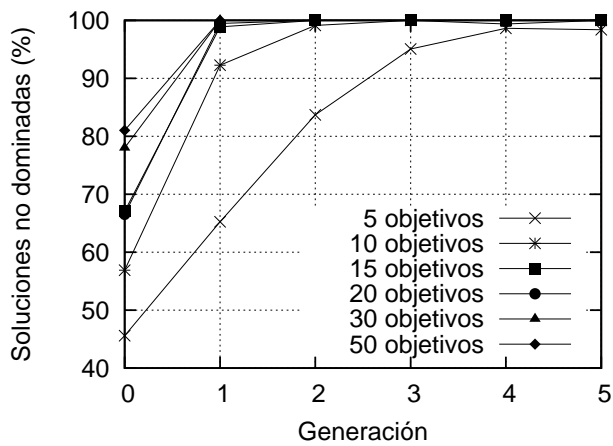
A partir de la figura 2.2 es posible observar que el aumento en el número de objetivos claramente provoca un incremento en el porcentaje de soluciones no dominadas en la población en los seis casos de prueba considerados. Este problema se acentúa rápidamente con el avance del proceso de búsqueda, de manera que en las primeras generaciones la población se satura por completo de soluciones no dominadas. Es importante notar que incluso desde la población inicial (generación



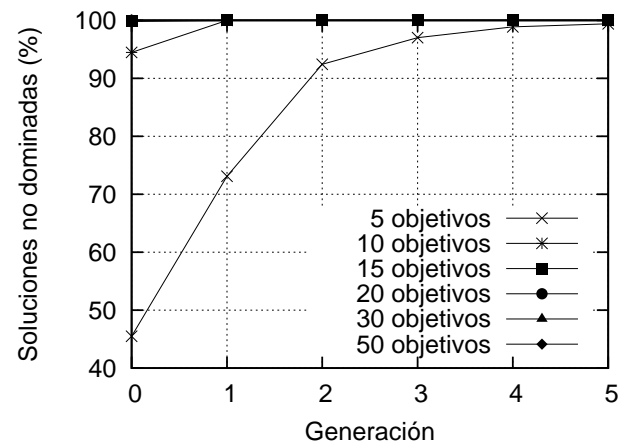
(a) DTLZ1.



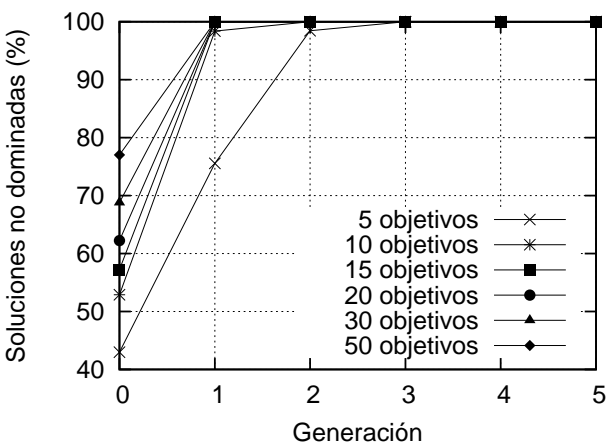
(b) DTLZ2.



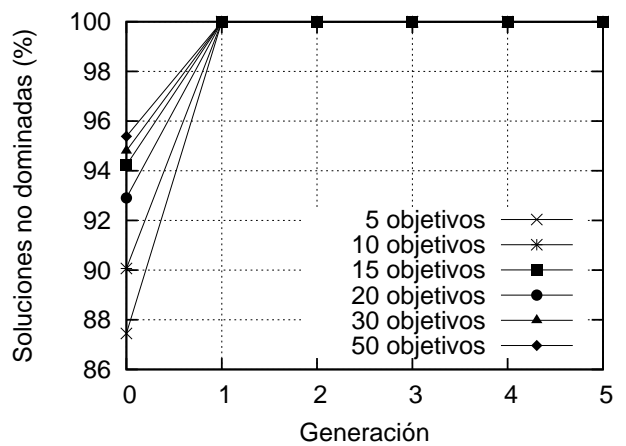
(c) DTLZ3.



(d) DTLZ4.



(e) DTLZ5.



(f) DTLZ6.

Figura 2.2: Proporción de soluciones no dominadas en relación con el número de objetivos.

0), que es generada aleatoriamente, un alto porcentaje de la población corresponde a soluciones no dominadas, especialmente para los problemas DTLZ4 y DTLZ6. Cuando todas las soluciones en la población son no dominadas (v.g., equivalentes de acuerdo con la dominancia de Pareto) no es posible discriminar entre ellas, de manera que el proceso de selección y, por lo tanto, el proceso de guiar la búsqueda se llevan a cabo prácticamente de manera aleatoria.

En realidad, PD sólo es apropiada para problemas con pocos objetivos. Supongamos que  $\mathbf{X}_i$  y  $\mathbf{X}_j$  son dos soluciones candidatas para un problema de optimización con 20 objetivos. Si  $\mathbf{X}_i$  es mejor que  $\mathbf{X}_j$  en 19 objetivos, pero  $\mathbf{X}_j$  es superior que  $\mathbf{X}_i$  en el objetivo restante, ambas soluciones serían consideradas equivalentes si se utiliza PD para compararlas. Sin embargo, en la ausencia de preferencias entre objetivos, el tomador de decisiones podría considerar que  $\mathbf{X}_i$  muestra claramente un mejor desempeño que  $\mathbf{X}_j$ . PD no considera la cantidad de objetivos en los que una solución es mejor que otra y omite también la magnitud de estas mejoras. Sin embargo, estos son aspectos muy importantes durante el proceso de toma de decisiones y pueden llevar a diferentes grados de dominancia [29, 47]; descartar esta valiosa información puede llevar a errores de discriminación.

### 2.4.3 Alternativas de solución

En la literatura se pueden distinguir dos enfoques principales para lidiar con una alta dimensionalidad en el espacio de las funciones objetivo:

- **Reducción de dimensiones.** La relación entre un par de objetivos puede ser *conflictiva* (al mejorar uno de los objetivos el otro empeora), *armoniosa* (al mejorar uno de los objetivos el otro también mejora) o *independiente* (el mejorar o empeorar de uno de los objetivos no tiene efecto sobre el otro) [59], de modo que los objetivos cuya relación es armoniosa pueden ser considerados *redundantes*. La reducción de dimensiones consiste en analizar las relaciones entre los diferentes objetivos de un problema con la finalidad de identificar redundancias y, de esta manera, reducir el problema al ignorar algunos objetivos innecesarios. Reducir las dimensiones claramente permite lidiar con las diferentes dificultades expuestas en la Sección 2.4.1; sin embargo, esta reducción no siempre es posible y depende del escenario de optimización en

cuestión. Algunos trabajos realizados al respecto son: [23, 7, 8, 6, 51, 41]

- **Métodos alternativos de discriminación.** La mayoría de los MOEAs propuestos se basan en la dominancia de Pareto (PD) para discriminar entre soluciones y guiar el proceso de búsqueda. Sin embargo, tal como se demostró en la Sección 2.4.2, al incrementarse la cantidad de objetivos PD pierde su potencial discriminante provocando que el proceso de búsqueda sea llevado a cabo prácticamente de manera aleatoria. Un enfoque de solución que pareciera natural ante esta problemática es la búsqueda de métodos alternativos para discriminar entre soluciones. La idea es utilizar métodos que permitan una discriminación más estricta (de grano más fino) y robusta que la lograda a través de PD, de manera que pueda mantenerse una presión de selección adecuada para dirigir el proceso de búsqueda. Diferentes alternativas de este tipo han sido propuestas, aunque no todas con particular interés en el aumento del número de objetivos: [4, 27, 46, 29, 26, 47, 49, 66, 76, 32]; estos y otros métodos alternativos se describen a detalle en el Capítulo 3.

El enfoque que seguiremos y estudiaremos en este proyecto de tesis es el segundo.

## 2.5 Problemas de Prueba

Para toda la experimentación involucrada en este trabajo de tesis hemos considerado seis problemas de prueba que son parte de un conjunto de funciones específicamente propuestas para estudiar la escalabilidad de MOEAs. Estos problemas de prueba son conocidos con las siglas DTLZ, que corresponden al nombre de sus autores: Deb, Thiele, Laumanns y Zitzler [24, 25]. Las razones principales que motivaron la elección de estos seis problemas de prueba son las siguientes:

1. Escalabilidad con respecto al número de objetivos y variables de decisión.
2. Se conoce de manera exacta la forma y localización del frente de Pareto (FP), así como los óptimos de Pareto correspondientes.

3. Existen ecuaciones que describen el FP para estos problemas, lo que permite evaluar analíticamente (sin utilizar un conjunto de referencia) la convergencia del conjunto aproximación obtenido por un MOEA.
4. Las dificultades de convergencia y diversidad pueden controlarse.

A continuación se describen los problemas de prueba considerados para esta tesis. Para todos los casos la notación utilizada es la siguiente: el número total de variables de decisión es  $n = M + k - 1$ , donde  $M$  es el número de objetivos y  $k$  es un parámetro de dificultad cuyo valor será especificado en la descripción de cada problema.  $\mathbf{X}$  es un vector de variables de decisión, y  $\mathbf{X}_M$  representará las últimas  $k$  variables de  $\mathbf{X}$ .  $x_i$  corresponde a la  $i$ -ésima variable del vector  $\mathbf{X}$  y está sujeta a  $0 \leq x_i \leq 1$ , para  $i = 1, 2, \dots, n$ .

### 2.5.1 DTLZ1

El problema DTLZ1 se describe formalmente de la siguiente manera:

$$\left. \begin{array}{l}
 \text{Minimizar } f_1(\mathbf{X}) = \frac{1}{2}x_1x_2 \dots x_{M-1}(1 + g(\mathbf{X}_M)), \\
 \text{Minimizar } f_2(\mathbf{X}) = \frac{1}{2}x_1x_2 \dots (1 - x_{M-1})(1 + g(\mathbf{X}_M)), \\
 \quad \quad \quad \vdots \\
 \text{Minimizar } f_{M-1}(\mathbf{X}) = \frac{1}{2}x_1(1 - x_2)(1 + g(\mathbf{X}_M)), \\
 \text{Minimizar } f_M(\mathbf{X}) = \frac{1}{2}(1 - x_1)(1 + g(\mathbf{X}_M)), \\
 \text{donde } g(\mathbf{X}_M) = 100 [|\mathbf{X}_M| + \sum_{x_i \in \mathbf{X}_M} (x_i - 0.5)^2 - \cos(20\pi(x_i - 0.5))].
 \end{array} \right\} \quad (2.18)$$

Los óptimos de Pareto para este problema corresponden a  $x_i = 0.5$  para toda  $x_i \in \mathbf{X}_M$ . El frente de Pareto (FP) para este problema corresponde al hiperplano lineal  $\sum_{m=1}^M f_m = 0.5$ . La figura 2.3 muestra una aproximación al FP para el problema DTLZ1 con tres objetivos ( $M = 3$ ). La dificultad en este problema es con respecto a la convergencia, ya que el espacio de búsqueda contiene  $(11^k - 1)$  atractores locales donde un MOEA puede estancarse antes de converger al FP. Para el caso de este problema se utilizará  $k = 5$ .

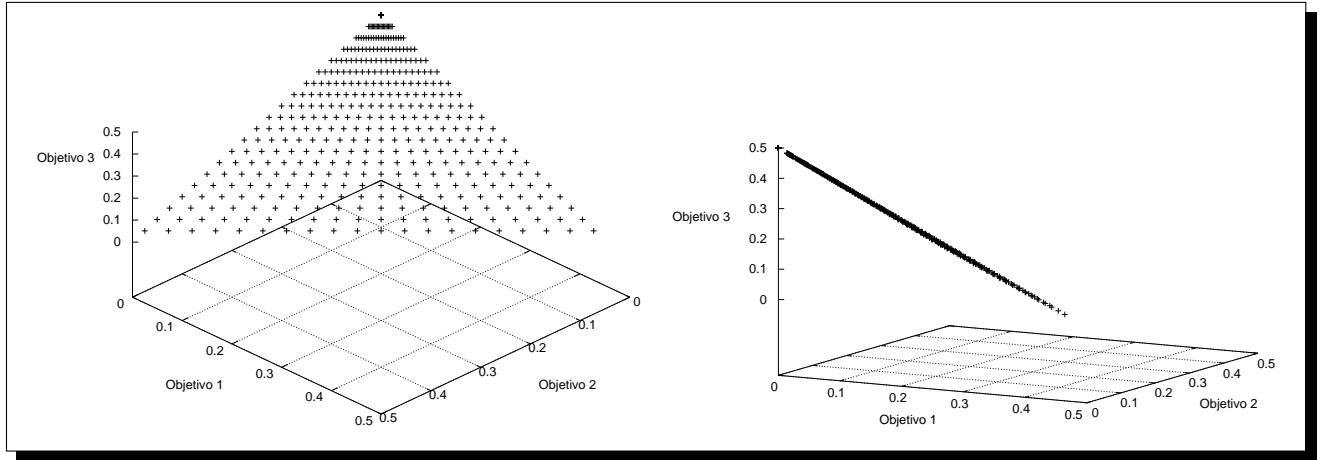


Figura 2.3: Aproximación al frente de Pareto para el problema DTLZ1 con 3 objetivos.

### 2.5.2 DTLZ2

El problema DTLZ2 tiene un FP esférico, como se muestra en la figura 2.4. Este problema se define de la siguiente manera:

$$\left. \begin{aligned}
 &\text{Minimizar } f_1(\mathbf{X}) = (1 + g(\mathbf{X}_M)) \cos\left(\frac{x_1\pi}{2}\right) \dots \cos\left(\frac{x_{M-2}\pi}{2}\right) \cos\left(\frac{x_{M-1}\pi}{2}\right), \\
 &\text{Minimizar } f_2(\mathbf{X}) = (1 + g(\mathbf{X}_M)) \cos\left(\frac{x_1\pi}{2}\right) \dots \cos\left(\frac{x_{M-2}\pi}{2}\right) \sin\left(\frac{x_{M-1}\pi}{2}\right), \\
 &\text{Minimizar } f_3(\mathbf{X}) = (1 + g(\mathbf{X}_M)) \cos\left(\frac{x_1\pi}{2}\right) \dots \sin\left(\frac{x_{M-2}\pi}{2}\right), \\
 &\quad \vdots \\
 &\text{Minimizar } f_M(\mathbf{X}) = (1 + g(\mathbf{X}_M)) \sin\left(\frac{x_1\pi}{2}\right), \\
 &\text{donde } g(\mathbf{X}_M) = \sum_{x_i \in \mathbf{X}_M} (x_i - 0.5)^2.
 \end{aligned} \right\} \quad (2.19)$$

Los óptimos de Pareto corresponden a  $x_i = 0.5$  para toda  $x_i \in \mathbf{X}_M$ . El FP para este problema es continuo, no convexo y se describe mediante la ecuación  $\sum_{m=1}^M f_m^2 = 1$ , que corresponde a una hipersfera con radio 1 y centro en el origen. El valor de  $k = 10$  será utilizado en este trabajo.

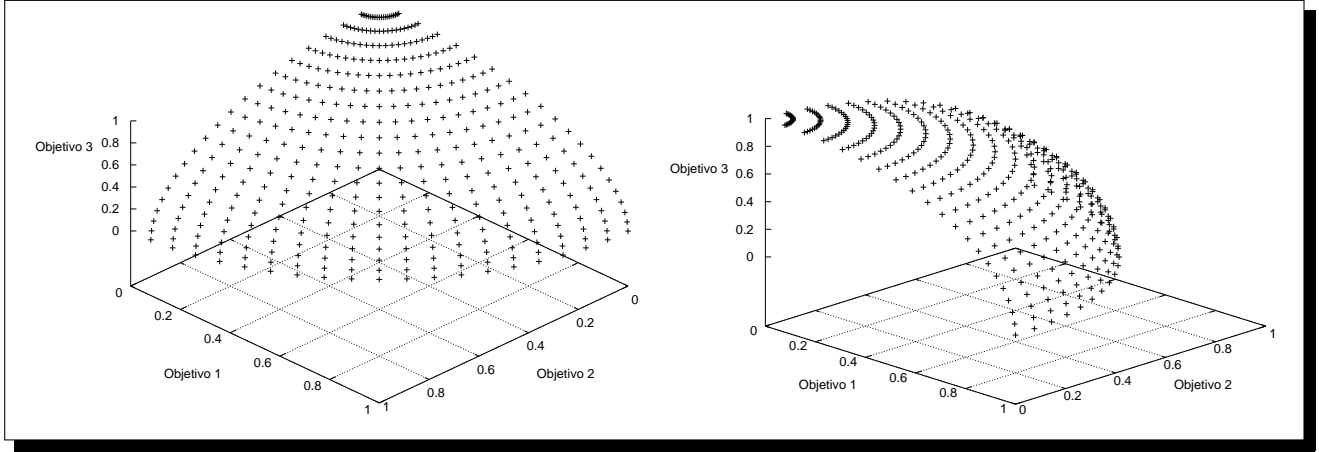


Figura 2.4: Aproximación al frente de Pareto para los problemas DTLZ2-DTLZ4 con 3 objetivos.

### 2.5.3 DTLZ3

DTLZ3 es una modificación del problema DTLZ2 descrito anteriormente. En esta versión del problema se incorporan mayores dificultades de convergencia, ya que el espacio de búsqueda incluye  $(3^k - 1)$  atractores locales que son paralelos al FP. La definición formal de este problema se plantea como sigue:

$$\left. \begin{aligned}
 &\text{Minimizar } f_1(\mathbf{X}) = (1 + g(\mathbf{X}_M)) \cos\left(\frac{x_1\pi}{2}\right) \dots \cos\left(\frac{x_{M-2}\pi}{2}\right) \cos\left(\frac{x_{M-1}\pi}{2}\right), \\
 &\text{Minimizar } f_2(\mathbf{X}) = (1 + g(\mathbf{X}_M)) \cos\left(\frac{x_1\pi}{2}\right) \dots \cos\left(\frac{x_{M-2}\pi}{2}\right) \sin\left(\frac{x_{M-1}\pi}{2}\right), \\
 &\text{Minimizar } f_3(\mathbf{X}) = (1 + g(\mathbf{X}_M)) \cos\left(\frac{x_1\pi}{2}\right) \dots \sin\left(\frac{x_{M-2}\pi}{2}\right), \\
 &\quad \vdots \\
 &\text{Minimizar } f_M(\mathbf{X}) = (1 + g(\mathbf{X}_M)) \sin\left(\frac{x_1\pi}{2}\right), \\
 &\quad \text{donde } g(\mathbf{X}_M) = 100 \left[ |\mathbf{X}_M| + \sum_{x_i \in \mathbf{X}_M} (x_i - 0.5)^2 - \cos(20\pi(x_i - 0.5)) \right].
 \end{aligned} \right\} \quad (2.20)$$

Los óptimos de Pareto corresponden a  $x_i = 0.5$  para toda  $x_i \in \mathbf{X}_M$ . Los valores óptimos para las funciones objetivo satisfacen la ecuación  $\sum_{m=1}^M f_m^2 = 1$ , que corresponde a una hipersfera con radio 1 y centro en el origen. Una aproximación al FP para esta problema con 3 objetivos puede visualizarse en la figura 2.4. El valor de  $k = 10$  será utilizado en este trabajo.



### 2.5.4 DTLZ4

El problema DTLZ2 se modificó para estudiar la habilidad de los MOEAs para mantener una buena distribución de las soluciones. DTLZ4 introduce dificultades para converger de una manera distribuida, ya que presenta un conjunto de soluciones denso en regiones específicas del espacio de búsqueda, lo que provoca la tendencia de los MOEAs a dirigir la búsqueda hacia dichas regiones. La definición formal para este problema es la siguiente:

$$\left. \begin{aligned}
 &\text{Minimizar } f_1(\mathbf{X}) = (1 + g(\mathbf{X}_M)) \cos\left(\frac{x_1^\alpha \pi}{2}\right) \dots \cos\left(\frac{x_{M-2}^\alpha \pi}{2}\right) \cos\left(\frac{x_{M-1}^\alpha \pi}{2}\right), \\
 &\text{Minimizar } f_2(\mathbf{X}) = (1 + g(\mathbf{X}_M)) \cos\left(\frac{x_1^\alpha \pi}{2}\right) \dots \cos\left(\frac{x_{M-2}^\alpha \pi}{2}\right) \sin\left(\frac{x_{M-1}^\alpha \pi}{2}\right), \\
 &\text{Minimizar } f_3(\mathbf{X}) = (1 + g(\mathbf{X}_M)) \cos\left(\frac{x_1^\alpha \pi}{2}\right) \dots \sin\left(\frac{x_{M-2}^\alpha \pi}{2}\right), \\
 &\quad \vdots \\
 &\text{Minimizar } f_M(\mathbf{X}) = (1 + g(\mathbf{X}_M)) \sin\left(\frac{x_1^\alpha \pi}{2}\right), \\
 &\text{donde } g(\mathbf{X}_M) = \sum_{x_i \in \mathbf{X}_M} (x_i - 0.5)^2.
 \end{aligned} \right\} \quad (2.21)$$

El FP para este problema es esférico y con las mismas características del FP descrito para los problemas DTLZ2 y DTLZ3 (ver figura 2.4). Los valores de los parámetros  $k = 10$  y  $\alpha = 100$  serán utilizados tal como se sugiere en [24].

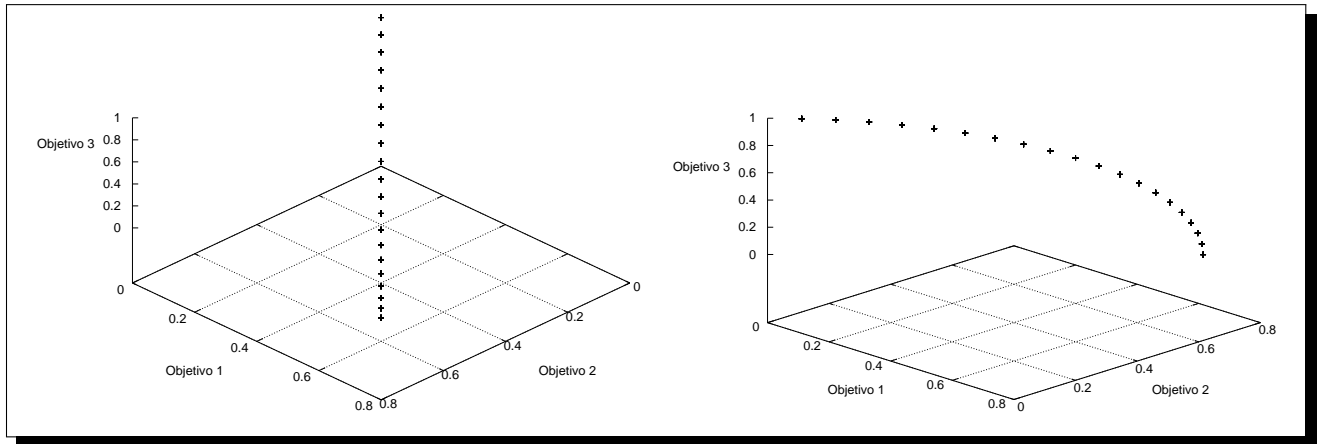


Figura 2.5: Aproximación al frente de Pareto para los problemas DTLZ5 y DTLZ6 con 3 objetivos.

### 2.5.5 DTLZ5

El problema DTLZ5 se define formalmente de la siguiente manera:

$$\left. \begin{aligned}
 &\text{Minimizar } f_1(\mathbf{X}) = (1 + g(\mathbf{X}_M)) \cos\left(\frac{\theta_1\pi}{2}\right) \dots \cos\left(\frac{\theta_{M-2}\pi}{2}\right) \cos\left(\frac{\theta_{M-1}\pi}{2}\right), \\
 &\text{Minimizar } f_2(\mathbf{X}) = (1 + g(\mathbf{X}_M)) \cos\left(\frac{\theta_1\pi}{2}\right) \dots \cos\left(\frac{\theta_{M-2}\pi}{2}\right) \sin\left(\frac{\theta_{M-1}\pi}{2}\right), \\
 &\text{Minimizar } f_3(\mathbf{X}) = (1 + g(\mathbf{X}_M)) \cos\left(\frac{\theta_1\pi}{2}\right) \dots \sin\left(\frac{\theta_{M-2}\pi}{2}\right), \\
 &\quad \vdots \\
 &\text{Minimizar } f_M(\mathbf{X}) = (1 + g(\mathbf{X}_M)) \sin\left(\frac{\theta_1\pi}{2}\right), \\
 &\quad \text{donde } \theta_i = \frac{\pi}{4(1+g(\mathbf{X}_M))} (1 + 2g(\mathbf{X}_M)x_i), \text{ para } i = 1, 2, \dots, (M-1), \\
 &\quad \text{y } g(\mathbf{X}_M) = \sum_{x_i \in \mathbf{X}_M} (x_i - 0.5)^2.
 \end{aligned} \right\} \quad (2.22)$$

Los óptimos para las funciones objetivo satisfacen la ecuación  $\sum_{m=1}^M f_m^2 = 1$ , y sus correspondientes valores en el espacio de decisión son  $x_i = 0.5$  para toda  $x_i \in \mathbf{X}_M$ . El problema DTLZ5 sirve para investigar la habilidad de MOEAs para converger a una curva. La figura 2.5 muestra una aproximación al FP para este problema usando tres objetivos. Para este trabajo se utilizará  $k = 10$ .

### 2.5.6 DTLZ6

Al igual que el problema DTLZ5, DTLZ6 está diseñado para investigar la habilidad de MOEAs para converger a una curva (ver figura 2.5). Sin embargo, DTLZ6 incorpora mayores dificultades de convergencia. Este problema se define formalmente de la siguiente manera:

$$\left. \begin{aligned}
 &\text{Minimizar } f_1(\mathbf{X}) = (1 + g(\mathbf{X}_M)) \cos\left(\frac{\theta_1\pi}{2}\right) \dots \cos\left(\frac{\theta_{M-2}\pi}{2}\right) \cos\left(\frac{\theta_{M-1}\pi}{2}\right), \\
 &\text{Minimizar } f_2(\mathbf{X}) = (1 + g(\mathbf{X}_M)) \cos\left(\frac{\theta_1\pi}{2}\right) \dots \cos\left(\frac{\theta_{M-2}\pi}{2}\right) \sin\left(\frac{\theta_{M-1}\pi}{2}\right), \\
 &\text{Minimizar } f_3(\mathbf{X}) = (1 + g(\mathbf{X}_M)) \cos\left(\frac{\theta_1\pi}{2}\right) \dots \sin\left(\frac{\theta_{M-2}\pi}{2}\right), \\
 &\quad \vdots \\
 &\text{Minimizar } f_M(\mathbf{X}) = (1 + g(\mathbf{X}_M)) \sin\left(\frac{\theta_1\pi}{2}\right), \\
 &\quad \text{donde } \theta_i = \frac{\pi}{4(1+g(\mathbf{X}_M))} (1 + 2g(\mathbf{X}_M)x_i) \text{ para } i = 1, 2, \dots, (M-1), \\
 &\quad \text{y } g(\mathbf{X}_M) = \sum_{x_i \in \mathbf{X}_M} x_i^{0.1}
 \end{aligned} \right\} \quad (2.23)$$

A diferencia de los problemas descritos anteriormente, las soluciones óptimas de Pareto para DTLZ6 corresponden a  $x_i = 0$  para toda  $x_i \in \mathbf{X}_M$ . Los valores óptimos para las funciones objetivo satisfacen la ecuación  $\sum_{m=1}^M f_m^2 = 1$ . El valor de  $k = |\mathbf{X}_M| = 10$  será utilizado en este trabajo.

## 2.6 Métricas

### 2.6.1 Métrica de Convergencia

Deb y Jain [20] propusieron en 2002 una métrica para evaluar la capacidad de convergencia de los MOEAs. Esta métrica consiste en calcular la distancia promedio entre las soluciones del conjunto aproximación  $Z$  obtenido por un MOEA y el punto más cercano a cada una de ellas en el frente de Pareto, mismo que es representado mediante un conjunto de puntos de referencia  $P^*$ . Inicialmente se calcula para cada punto  $\mathbf{X} \in Z$  la distancia Euclidiana normalizada al punto más cercano en  $P^*$ :

$$d(\mathbf{X}) = \min_{\mathbf{X}^* \in P^*} \sqrt{\sum_{m=1}^M \left( \frac{f_m(\mathbf{X}) - f_m(\mathbf{X}^*)}{f_m^{max} - f_m^{min}} \right)^2} \quad (2.24)$$

donde  $f_m^{max}$  y  $f_m^{min}$  denotan respectivamente el valor máximo y mínimo para el  $m$ -ésimo objetivo en  $P^*$ . Finalmente la métrica de convergencia se calcula promediando las distancias normalizadas de las diferentes  $\mathbf{X} \in Z$ :

$$C(Z) = \frac{\sum_{\mathbf{X} \in Z} d(\mathbf{X})}{|Z|} \quad (2.25)$$

Para esta tesis utilizaremos una variante de la métrica de convergencia de Deb y Jain [20] previamente descrita. Dado que para los seis problemas de prueba considerados se conocen ecuaciones que describen el FP, el valor de  $d(\mathbf{X})$  para cada  $\mathbf{X} \in Z$  se determinó analíticamente (sin utilizar un conjunto de referencia) [42, 72].

Por otra parte, el conjunto aproximación  $Z$  puede incluir a la vez soluciones buenas (cercanas al PF) y malas (lejanas al PF), aún cuando son no dominadas con respecto a la dominancia de Pareto

(situación especialmente común al optimizar muchos objetivos). De este modo, la distancia promedio puede estar influenciada por las malas soluciones, pudiendo opacar en algunos casos la habilidad de un determinado MOEA para encontrar soluciones que presenten una buena convergencia. Por estos motivos, en lugar de calcular la distancia promedio se utilizará como medida de convergencia la distancia entre la mejor solución encontrada (solución con mejor convergencia) y el FP de los diferentes problemas de prueba:

$$C(Z) = \min_{\mathbf{X} \in Z} d(\mathbf{X}) \quad (2.26)$$

Mientras más pequeño sea el valor para esta métrica, mejor será el desempeño del algoritmo evaluado.

### 2.6.2 Distancia Generacional Invertida

La *distancia generacional* (DG) es una métrica propuesta por Van Veldhuizen en 1999 [71]. DG consiste en calcular la distancia promedio de un conjunto de soluciones candidatas  $Z$ , con respecto a un conjunto de referencia  $P^*$  que representa al frente de Pareto (FP). Formalmente, DG se define por la Ecuación 2.29:

$$DG = \frac{\sqrt{\sum_{\mathbf{X} \in Z} d(\mathbf{X})^2}}{|Z|} \quad (2.27)$$

donde  $d(\mathbf{X})$  es la distancia Euclidiana entre la solución  $\mathbf{X}$  y el punto en  $P^*$  más cercano a ella:

$$d(\mathbf{X}) = \min_{\mathbf{X}^* \in P^*} \sqrt{\sum_{m=1}^M (f_m(\mathbf{X}) - f_m(\mathbf{X}^*))^2} \quad (2.28)$$

Aunque DG es una métrica para evaluar convergencia, si invertimos los papeles de  $Z$  y  $P^*$  en las Ecuaciones (2.29) y (2.30), lo que se conoce como *distancia generacional invertida* (DGI), es posible tomar en cuenta también la diversidad del conjunto  $Z$ . De esta manera, un valor bajo para DGI indicará tanto una buena convergencia como buena distribución de las soluciones. De manera

formal, la métrica DGI se define como sigue:

$$DGI = \frac{\sqrt{\sum_{\mathbf{X}^* \in P^*} d(\mathbf{X}^*)^2}}{|P^*|} \quad (2.29)$$

donde  $d(\mathbf{X}^*)$  es la distancia Euclidiana entre el punto de referencia  $\mathbf{X}^*$  y la solución en  $Z$  más cercana:

$$d(\mathbf{X}^*) = \min_{\mathbf{X} \in Z} \sqrt{\sum_{m=1}^M (f_m(\mathbf{X}) - f_m(\mathbf{X}^*))^2} \quad (2.30)$$

Para este trabajo el conjunto  $P^*$  estará formado únicamente por los extremos del FP y el punto conocido como la *rodilla* (*knee*) del FP [50]; la rodilla es el punto del FP que presenta el mejor compromiso entre los diferentes objetivos del problema, en el que podría asumirse que el tomador de decisiones estará interesado. Para el problema DTLZ1 los extremos del FP están dados por los  $M$  vectores  $\mathbf{F}_i = [f_1^i, f_2^i, \dots, f_M^i]$  tales que se cumpla que  $f_m^i = 0.5$  si  $m = i$  y  $f_m^i = 0$ , en caso contrario, para  $i \in \{1, 2, \dots, M\}$  (donde  $M$  es la cantidad de objetivos), mientras que la rodilla para este problema corresponde al punto  $\mathbf{r} = \frac{0.5}{M} \mathbf{1}$ . Por otra parte, para los problemas DTLZ2-DTLZ6 los extremos del FP están dados por los  $M$  vectores  $\mathbf{F}_i = [f_1^i, f_2^i, \dots, f_M^i]$  tales que se cumpla que  $f_m^i = 1.0$  si  $m = i$  y  $f_m^i = 0$  en caso contrario, para  $i \in \{1, 2, \dots, M\}$ . La rodilla en el caso de los problemas DTLZ2-DTLZ6 corresponde al punto  $\mathbf{r} = \frac{1}{\sqrt{M}} \mathbf{1}$ .

De este modo, el conjunto  $P^*$  utilizado para calcular la métrica DGI contiene en todos los casos únicamente  $M + 1$  puntos de referencia. Para dejar claro este punto consideremos el caso de los problemas DTLZ2-DTLZ6 con 5 objetivos. El conjunto de referencia  $P^*$  utilizado para calcular la métrica DGI es:

$$P^* = \{[0, 0, 0, 0, 1], [0, 0, 0, 1, 0], [0, 0, 1, 0, 0], [0, 1, 0, 0, 0], [1, 0, 0, 0, 0], [\frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}]\}$$

La ubicación de las soluciones encontradas por un MOEA puede favorecer o no el resultado de la métrica DGI. Por ejemplo, si un algoritmo A logra encontrar uno de los extremos del FP para determinado problema, mientras que otro algoritmo B converge exactamente a la rodilla, claramente

el B logrará un mejor valor para la métrica, pues la rodilla es el punto más cercano a los diferentes puntos de referencia incluidos en  $P^*$ .

# 3

## Métodos de asignación de aptitud para optimización multiobjetivo

### 3.1 Introducción

Para emular el proceso de la selección natural, los algoritmos evolutivos (AEs) requieren implementar una forma para expresar la calidad de las soluciones, v.g., una función de aptitud, de tal modo que se favorezca a aquellos individuos que compiten más eficientemente por recursos. En otras palabras, la aptitud permitirá identificar los mejores individuos para otorgarles mayor oportunidad de reproducirse y generar nuevos individuos. Del mismo modo, la aptitud guiará la selección de aquellos individuos mejor adaptados para sobrevivir de una generación a otra. La asignación de aptitud es de gran importancia; a través de una correcta discriminación entre soluciones será posible que un AE guíe el proceso de búsqueda hacia el óptimo.

En el caso de problemas mono-objetivo, usualmente dicha función de aptitud está directamente relacionada con la función objetivo que será optimizada. Sin embargo, cuando se optimizan múltiples objetivos la discriminación entre soluciones no es tan trivial; se requiere de un mecanismo adicional

para mapear el espacio multidimensional de las funciones objetivo a una sola dimensión, de modo que las soluciones puedan ser comparadas directamente y jerarquizadas; este mecanismo se conoce como *método de asignación de aptitud*<sup>1</sup> [38].

La dominancia de Pareto (PD) es la relación más comúnmente adoptada para comparar soluciones en el contexto multiobjetivo. No obstante, PD pierde su potencial de discriminación cuando se incrementa la cantidad de objetivos (ver Sección 2.4.2). Este comportamiento hace notoria la necesidad de identificar métodos alternativos que trabajen más adecuadamente ante alta dimensionalidad en el espacio de las funciones objetivo.

En este capítulo se describen 22 métodos de asignación de aptitud para optimización multiobjetivo que han sido reportados en la literatura especializada y 13 nuevas propuestas de esta índole resultantes de esta tesis. Los diferentes métodos estudiados se han agrupado en 3 categorías diferentes:

1. **Métodos basados en la dominancia de Pareto.** Aún cuando se ha demostrado que PD pierde efectividad al incrementarse la cantidad de objetivos, es preciso incluir métodos de jerarquización basados en este concepto como punto de referencia.
2. **Composiciones jerárquicas.** Las composiciones jerárquicas (CJ) son métodos que jerarquizan una población considerando un objetivo a la vez, de manera que cada individuo tendrá tantas jerarquías como objetivos tenga el problema. Las posiciones jerárquicas de un individuo  $\mathbf{X}_i$  están dadas por el vector  $\mathbf{R}(\mathbf{X}_i) = (r_1(\mathbf{X}_i), r_2(\mathbf{X}_i), \dots, r_M(\mathbf{X}_i))^T$ , donde  $r_m(\mathbf{X}_i)$  es la jerarquía de  $\mathbf{X}_i$  con respecto al  $m$ -ésimo objetivo. Una vez calculado el vector  $\mathbf{R}(\mathbf{X}_i)$  para cada  $\mathbf{X}_i \in P$ , las diferentes jerarquías de  $\mathbf{X}_i$  son procesadas de alguna forma para obtener una clasificación global. Giusti *et al.* [32] hacen referencia a estas dos etapas como los pasos de jerarquización y de composición, respectivamente. Para la descripción de las diferentes CJs estudiadas se asume una previa obtención del vector  $\mathbf{R}(\mathbf{X}_i)$  para cada  $\mathbf{X}_i \in P$ .

Estos métodos tienen como ventaja su eficiencia computacional, sencillez de implementación, e independencia del rango en que se encuentran los distintos objetivos del problema. Sin embargo,

---

<sup>1</sup>En esta tesis se utilizarán indistintamente los términos aptitud y jerarquía para hacer referencia al valor que expresa la calidad de las soluciones y que permite la discriminación entre ellas.



estos enfoques no toman en cuenta la magnitud de significancia entre soluciones.

3. **Formas relajadas de dominancia.** Las formas relajadas de dominancia (FRD) son alternativas que consideran de alguna forma el grado de dominancia entre soluciones, el cual puede relacionarse con el número de objetivos en los que una solución es mejor que otra y/o con la magnitud de estas mejoras [29, 47]. La idea es lograr una discriminación de grano más fino entre soluciones. Las FRDs pueden permitir que un individuo  $\mathbf{X}_i$  sea considerado mejor que otro individuo  $\mathbf{X}_j$  aún cuando originalmente estos individuos sean no dominados entre sí, o incluso en el caso de que  $\mathbf{X}_j$  domina a  $\mathbf{X}_i$  [49]. Generalmente en este tipo de métodos existe cierta tolerancia para aceptar deterioro en algunos objetivos cuando esto implica una mejora significativa en otros criterios.

## 3.2 Enfoques tomados de la literatura

### 3.2.1 Métodos basados en la dominancia de Pareto

**Jerarquización de Pareto (Pareto ranking, PRNK):** PRNK es un método de jerarquización propuesto por Fonseca y Fleming [31] en 1993. De acuerdo con PRNK, los individuos no dominados tendrán jerarquía 1, mientras que el resto serán penalizados por cada individuo de la población que los domina (con respecto a la dominancia de Pareto). Formalmente, la jerarquía de un individuo  $\mathbf{X}_i \in P$  está dada por:

$$\text{PRNK}(\mathbf{X}_i) = 1 + |\{\mathbf{X}_j \in P : \mathbf{X}_j \prec \mathbf{X}_i\}| \quad (3.1)$$

**Ordenamiento por no dominancia (nondominated sorting, PSRT):** Este método fue originalmente propuesto por Goldberg [33] y posteriormente implementado por Srinivas y Deb como parte de su algoritmo NSGA (Nondominated Sorting Genetic Algorithm) [69]. La clasificación de los individuos se realiza mediante un proceso iterativo. Inicialmente, los individuos no dominados son identificados y clasificados con la jerarquía más alta. Después, los individuos ya jerarquizados

son omitidos y las nuevas soluciones no dominadas son identificadas y clasificadas con la siguiente jerarquía. Este proceso se repite hasta que cada individuo en la población tenga una jerarquía, de manera que al final se obtiene una población organizada en frentes o niveles de dominancia (ver figura 3.1).

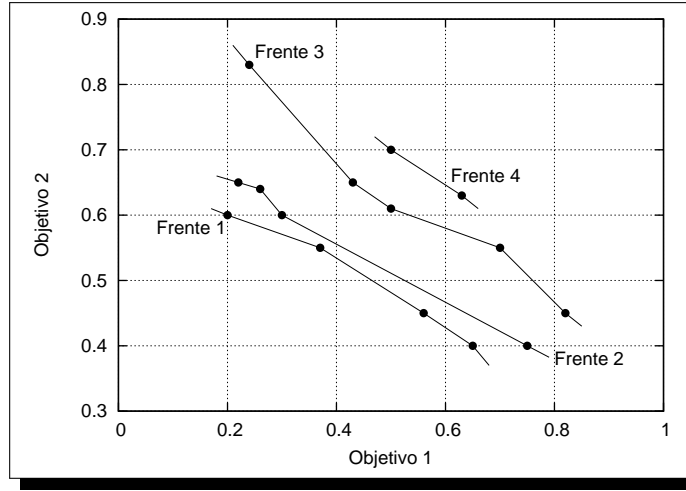


Figura 3.1: Ordenamiento por no dominancia (*Nondominated Sorting*).

### 3.2.2 Composiciones jerárquicas

En esta sección se describen diferentes composiciones jerárquicas (CJs) que han sido reportadas en la literatura especializada. Se asume la obtención previa del vector  $\mathbf{R}(\mathbf{X}_i)$  para cada  $\mathbf{X}_i \in P$  (ver Sección 3.1).

**Jerarquía promedio (average ranking, AVER):** Esta técnica fue propuesta por Bentley y Wakefield [4]. **AVER** expresa la calidad de un individuo como la sumatoria de sus diferentes posiciones jerárquicas. Formalmente, la calidad de una solución  $\mathbf{X}_i$  está dada por:

$$\text{AVER}(\mathbf{X}_i) = \sum_{m=1}^M r_m(\mathbf{X}_i) \quad (3.2)$$

**AVER** es un criterio de minimización, de modo que una solución  $\mathbf{X}_i$  domina a otra solución  $\mathbf{X}_j$ ,

con respecto a la relación  $\text{AVER}$  ( $\mathbf{X}_i \prec_{\text{AVER}} \mathbf{X}_j$ ), si y sólo si:

$$\text{AVER}(\mathbf{X}_i) < \text{AVER}(\mathbf{X}_j) \quad (3.3)$$

**Máxima jerarquía (maximum ranking, MAXR):** Bentley y Wakefield [4] propusieron esta técnica basándose en el algoritmo VEGA [67, 68]. Con **MAXR**, la clasificación para un individuo  $\mathbf{X}_i$  será la mejor jerarquía obtenida al considerar cada objetivo por separado, es decir:

$$\text{MAXR}(\mathbf{X}_i) = \min_{m=1}^M r_m(\mathbf{X}_i) \quad (3.4)$$

Dadas dos soluciones  $\mathbf{X}_i, \mathbf{X}_j \in P$ , de acuerdo con **MAXR** se dice que  $\mathbf{X}_i$  domina a  $\mathbf{X}_j$ , denotado por  $\mathbf{X}_i \prec_{\text{MAXR}} \mathbf{X}_j$ , si se cumple que:

$$\text{MAXR}(\mathbf{X}_i) < \text{MAXR}(\mathbf{X}_j) \quad (3.5)$$

Una notoria desventaja de **MAXR** es que tiende a favorecer extremos, es decir, asigna la mejor jerarquía a soluciones con el mejor valor para algún objetivo sin tomar en cuenta su desempeño para el resto de los criterios. Con la finalidad de aclarar este punto, consideremos el siguiente ejemplo: si para un problema con 20 objetivos,  $\mathbf{X}_i$  es la solución con el mejor desempeño en el primer objetivo, pero es la peor en los 19 restantes,  $\mathbf{X}_i$  será aún clasificada con la mejor jerarquía por **MAXR**. Al dar preferencia a este tipo de soluciones, **MAXR** puede rechazar aquellas con un mejor compromiso, las cuales, en la ausencia de preferencias, podría asumirse que son de mayor interés para el tomador de decisiones.

**Composición recíproca (reciprocal composition, RECC):** Esta técnica fue propuesta por Giusti *et al.* [32]. La aptitud de un individuo  $\mathbf{X}_i$  se calcula sumando la inversa de cada una de sus

posiciones jerárquicas:

$$\text{RECC}(\mathbf{X}_i) = \sum_{m=1}^M \frac{1}{r_m(\mathbf{X}_i)} \quad (3.6)$$

A diferencia de las dos CJs descritas previamente, **RECC** es un criterio de maximización. Por lo tanto, se dice que una solución  $\mathbf{X}_i$  domina a otra solución  $\mathbf{X}_j$  ( $\mathbf{X}_i \prec_{\text{RECC}} \mathbf{X}_j$ ) si y sólo si:

$$\text{RECC}(\mathbf{X}_i) > \text{RECC}(\mathbf{X}_j) \quad (3.7)$$

**Composición mediana (median composition, MEDC):** Giusti *et al.* [32] también propusieron clasificar las soluciones de la siguiente manera: ordenar el vector  $\mathbf{R}(\mathbf{X}_i)$  con las jerarquías que un determinado individuo  $\mathbf{X}_i$  obtuvo para los diferentes objetivos. La aptitud de  $\mathbf{X}_i$ ,  $\text{MEDC}(\mathbf{X}_i)$ , será igual al valor central de esta lista ordenada si  $M$  es impar o a la media de los dos valores centrales si  $M$  es par (v.g., la mediana de sus posiciones jerárquicas). De acuerdo con esta relación, se considera que una solución  $\mathbf{X}_i$  domina a otra solución  $\mathbf{X}_j$  si y sólo si se cumple que:

$$\text{MEDC}(\mathbf{X}_i) < \text{MEDC}(\mathbf{X}_j) \quad (3.8)$$

**Composición condorcet (condorcet composition, CONC):** Este método propuesto por Giusti *et al.* [32] trabaja de la siguiente manera: inicialmente, cada pareja de individuos ( $\mathbf{X}_i, \mathbf{X}_j$ ) se compara con respecto a sus diferentes posiciones jerárquicas. Si  $\mathbf{X}_i$  tiene mejor jerarquía que  $\mathbf{X}_j$  en una mayor cantidad de objetivos, entonces  $\mathbf{X}_i$  es el ganador. Si cada individuo supera al otro en igual cantidad de jerarquías, o si sus vectores de jerarquías son idénticos ( $\mathbf{R}(\mathbf{X}_i) = \mathbf{R}(\mathbf{X}_j)$ ), el resultado es un empate. Durante este proceso se mantienen contadores para las victorias y derrotas que obtiene cada individuo. Finalmente, para establecer preferencias entre dos soluciones  $\mathbf{X}_i, \mathbf{X}_j \in P$ , se dice que  $\mathbf{X}_i$  domina a  $\mathbf{X}_j$  bajo la relación **CONC** ( $\mathbf{X}_i \prec_{\text{CONC}} \mathbf{X}_j$ ) si  $\mathbf{X}_i$  tiene un mayor número de victorias que  $\mathbf{X}_j$ . En caso de que  $\mathbf{X}_i$  y  $\mathbf{X}_j$  tengan igual número de victorias, se utilizará la cantidad de derrotas como criterio de desempate.

### 3.2.3 Formas relajadas de dominancia

#### 3.2.3.1 Métodos que requieren ajuste de parámetros

La principal desventaja de los métodos que se incluyen en este apartado es que su funcionamiento depende de la correcta sintonización de algunos parámetros, lo que implica en algunos casos un conocimiento profundo del método y del problema.

**Dominancia  $\alpha$  ( $\alpha$ -dominance, ADOM):** Esta técnica fue propuesta por Ikeda *et al.* [46] como una forma para lidiar con lo que nombraron *dominance resistant solutions* (DRS), v.g., soluciones extremadamente inferiores pero difícilmente dominables. La idea es establecer cotas para la tasa de compromiso entre dos objetivos, de modo que ADOM permita que una solución  $\mathbf{X}_i$  domine a otra solución  $\mathbf{X}_j$  si  $\mathbf{X}_i$  es ligeramente inferior que  $\mathbf{X}_j$  en algún objetivo, pero significativamente superior en otro(s).

Una solución  $\mathbf{X}_i$   $\alpha$ -domina a otra solución  $\mathbf{X}_j$  ( $\mathbf{X}_i \prec_\alpha \mathbf{X}_j$ ) si y sólo si:

$$\begin{aligned} & \forall m \in \{1, 2, \dots, M\} : g_m(\mathbf{X}_i, \mathbf{X}_j) \leq 0 \\ & \wedge \exists m \in \{1, 2, \dots, M\} : g_m(\mathbf{X}_i, \mathbf{X}_j) < 0 \end{aligned} \quad (3.9)$$

donde  $g_m(\mathbf{X}_i, \mathbf{X}_j)$  es la aptitud relativa de  $\mathbf{X}_i$  con respecto a  $\mathbf{X}_j$  en el  $m$ -ésimo objetivo:

$$g_m(\mathbf{X}_i, \mathbf{X}_j) = f_m(\mathbf{X}_i) - f_m(\mathbf{X}_j) + \sum_{n \neq m} \alpha_{mn} (f_n(\mathbf{X}_i) - f_n(\mathbf{X}_j)) \quad (3.10)$$

$\alpha_{mn}$  es la tasa de compromiso entre los objetivos  $m$  y  $n$ . El usuario debe especificar este valor para cada par de objetivos. En [46], los autores usaron los valores  $\frac{1}{3}$ ,  $\frac{1}{9}$  y  $\frac{1}{100}$  para todo  $\alpha_{mn}$  tal que  $m \neq n$ . De acuerdo con sus resultados,  $\frac{1}{3}$  fue el valor con que se observó una mejor convergencia, mismo valor que se utilizará en este trabajo. ADOM es dependiente del rango de los objetivos, de tal forma que deben estar normalizados antes de su aplicación.

**Dominancia  $k$  ( $k$ -dominance, KDOM):** Farina y Amato [29] propusieron una relación de preferencia donde se considera la cantidad de objetivos en los que una solución  $\mathbf{X}_i$  es mejor, igual y peor que otra solución  $\mathbf{X}_j$ . Para esto propusieron las siguientes funciones, respectivamente:

$$n_b(\mathbf{X}_i, \mathbf{X}_j) = |\{m : f_m(\mathbf{X}_i) < f_m(\mathbf{X}_j)\}| \quad m = 1, 2, \dots, M \quad (3.11)$$

$$n_e(\mathbf{X}_i, \mathbf{X}_j) = |\{m : f_m(\mathbf{X}_i) = f_m(\mathbf{X}_j)\}| \quad m = 1, 2, \dots, M \quad (3.12)$$

$$n_w(\mathbf{X}_i, \mathbf{X}_j) = |\{m : f_m(\mathbf{X}_i) > f_m(\mathbf{X}_j)\}| \quad m = 1, 2, \dots, M \quad (3.13)$$

Por simplicidad se hará referencia a las funciones mostradas en las ecuaciones (3.11), (3.12) y (3.13) como  $n_b$ ,  $n_e$  y  $n_w$ , respectivamente. Dados dos individuos  $\mathbf{X}_i$  y  $\mathbf{X}_j$ , se dice que  $\mathbf{X}_i$   $k$ -domina<sup>2</sup> a  $\mathbf{X}_j$  ( $\mathbf{X}_i \prec_k \mathbf{X}_j$ ) si y sólo si se cumple que:

$$n_e < M \quad \wedge \quad n_b \geq \frac{M - n_e}{k + 1} \quad (3.14)$$

donde  $k$  puede tener cualquier valor en el rango  $[0, 1]$ . Si  $k = 0$ , entonces esta relación corresponde a la dominancia de Pareto. Del valor de  $k$  depende la presión que se ejerce al jerarquizar una población. En este trabajo utilizaremos  $k = 1$  para ejercer la mayor presión de selección y discriminación entre soluciones.  $k = 1$  significa que se dará preferencia a una solución  $\mathbf{X}_i$  sobre otra solución  $\mathbf{X}_j$ , si  $\mathbf{X}_i$  es mejor que  $\mathbf{X}_j$  en al menos la mitad de los objetivos en los que estas soluciones son diferentes.

**Dominancia por volumen (volume dominance, VDOM):** En 2007, Le y Landa-Silva [49] propusieron una forma relajada de dominancia que compara el volumen (del espacio de objetivos) dominado entre dos soluciones.

El volumen dominado por un individuo  $\mathbf{X}_i$  se define como la región  $R$  cuyas soluciones factibles son dominadas por  $\mathbf{X}_i$ . Un punto de referencia  $r$  debe ser definido de tal modo que sea dominado por todas las soluciones en  $R$ . El volumen dominado por una solución  $\mathbf{X}_i$ , con respecto a un punto

---

<sup>2</sup>En [29] se utiliza el término  $(1 - k)$ -domina. En este trabajo utilizaremos  $k$ -domina por simplicidad.

de referencia  $r = (r_1, r_2, \dots, r_M)$ , está dado por

$$V(\mathbf{X}_i, r) = \prod_{m=1}^M (r_m - f_m(\mathbf{X}_i)) \quad (3.15)$$

Para establecer la relación de dominancia entre dos soluciones  $\mathbf{X}_i$  y  $\mathbf{X}_j$ , es necesario calcular también el *volumen dominado compartido* (*shared dominated volume*,  $SV$ ), v.g., el volumen dominado por ambas soluciones. El volumen dominado compartido por  $\mathbf{X}_i$  y  $\mathbf{X}_j$  con respecto a un punto de referencia  $r = (r_1, r_2, \dots, r_M)$ , denotado por  $SV(\mathbf{X}_i, \mathbf{X}_j, r)$ , se define como sigue:

$$SV(\mathbf{X}_i, \mathbf{X}_j, r) = \prod_{m=1}^M (r_m - \max(f_m(\mathbf{X}_i), f_m(\mathbf{X}_j))) \quad (3.16)$$

Dado lo anterior, se dice que un individuo  $\mathbf{X}_i$   $V$ -domina (volumen-domina) a otro individuo  $\mathbf{X}_j$  ( $\mathbf{X}_i \prec_V \mathbf{X}_j$ ) si alguna de las condiciones mostradas en las ecuaciones (3.17) y (3.18) se cumple.

$$V(\mathbf{X}_j, r) = SV(\mathbf{X}_i, \mathbf{X}_j, r) \wedge V(\mathbf{X}_i, r) > SV(\mathbf{X}_i, \mathbf{X}_j, r) \quad (3.17)$$

$$V(\mathbf{X}_i, r) > V(\mathbf{X}_j, r) > SV(\mathbf{X}_i, \mathbf{X}_j, r) \wedge \frac{V(\mathbf{X}_i, r) - V(\mathbf{X}_j, r)}{SV(\mathbf{X}_i, \mathbf{X}_j, r)} > rSV \quad (3.18)$$

Un valor pequeño para  $rSV$  indica que una pequeña diferencia entre el volumen dominado por dos soluciones es suficiente para establecer preferencias entre ellas. En [49], los autores sugieren que el valor  $rSV$  esté en el rango  $[0.10, 0.20]$  para el problema de optimización sobre el cual se enfocaron en su estudio. Sin embargo, recomendaciones directas de los autores nos indican que en nuestro caso de estudio un valor en el rango  $[0.05, 0.15]$  sea razonable para  $rSV$ . Para este trabajo utilizaremos  $rSV = 0.10$ . El punto de referencia  $r$  debe ser también especificado por el usuario. Como será mencionado a detalle en el Capítulo 4, en este trabajo todos los métodos (incluyendo éste) se aplicarán sobre los objetivos normalizados al rango  $[0, 1]$ , por lo que  $r$  puede ser cualquier punto tal que  $r_m > 1$  para toda  $m \in \{1, 2, \dots, M\}$ . En este trabajo utilizaremos  $r = 1.1^M$ .

**Dominancia por contracción/expansión (contraction/expansion dominance, CEDM):**

Sato *et al.* [66] propusieron en 2007 un método para expandir/contrair el área de dominancia de las soluciones, lo que permite regular la presión en el proceso de selección. Antes de utilizar la relación de dominancia de Pareto, la aptitud que cada solución  $\mathbf{X}_i$  tiene para los diferentes objetivos se modifica de la siguiente manera:

$$f'_m(\mathbf{X}_i) = \frac{r \cdot \sin(\omega_m + S_m \cdot \pi)}{\sin(S_m \cdot \pi)} \quad \forall m \in \{1, 2, \dots, M\} \quad (3.19)$$

donde  $r$  es la norma del vector  $\mathbf{F}(\mathbf{X}_i)$  y  $\omega_m$  es el ángulo de declinación entre  $\mathbf{F}(\mathbf{X}_i)$  y  $f_m(\mathbf{X}_i)$ .  $S_m$  es un parámetro definido por el usuario cuyo valor es el que permite la expansión/contracción del área de dominancia de  $\mathbf{X}_i$  para la dimensión  $m$ . Los valores posibles para  $S_m$  están en el rango  $[0.25, 0.75]$ . Cuando  $S_m = 0.5$ , las aptitudes del individuo no son alteradas y  $f'_m(\mathbf{X}_i) = f_m(\mathbf{X}_i)$ . Por otra parte,  $S_m > 0.5$  provoca una contracción en el área de dominancia de las soluciones y una jerarquización de grano más grueso, lo que debilita el proceso de selección. De otro modo, cuando  $S_m < 0.5$  el área de dominancia se expande para lograr una jerarquización de grano más fino y una mayor presión en el proceso de selección, siendo esto de particular interés para lidiar con el incremento en la cantidad de objetivos. Para este trabajo de investigación utilizaremos  $S_m = 0.25$  para toda  $m \in \{1, 2, \dots, M\}$ , con la finalidad de lograr ejercer la máxima presión de selección y discriminación entre soluciones.

### 3.2.3.2. Métodos que no requieren ajuste de parámetros

A diferencia de los métodos descritos en la Sección 3.2.3.1, la ventaja principal de las FRDs que serán descritas en esta sección es que su funcionamiento no depende del ajuste de ningún parámetro, lo que amplía su rango de aplicación y su facilidad de comprensión e implementación.

**Suma ponderada (weighted sum, WSUM):** La forma más natural de asignar aptitud en el contexto multiobjetivo es combinar los valores de los diferentes objetivos utilizando cualquier operador aritmético (o una combinación de ellos). La más utilizada de tales técnicas es **WSUM**, que consiste en sumar todas las funciones objetivo usando diferentes coeficientes (pesos) que denotan su



importancia relativa. Formalmente, la aptitud de una solución  $\mathbf{X}_i$  está dada por la siguiente ecuación:

$$\mathbf{WSUM}(\mathbf{X}_i) = \sum_{m=1}^M w_m f_m(\mathbf{X}_i) \quad (3.20)$$

donde  $w_m$  es el coeficiente (peso) para el  $m$ -ésimo objetivo<sup>3</sup>. **WSUM** es un criterio de minimización, por lo que se considera que una solución  $\mathbf{X}_i$  domina a otra solución  $\mathbf{X}_j$  ( $\mathbf{X}_i \prec_{\mathbf{WSUM}} \mathbf{X}_j$ ) si se cumple que  $\mathbf{WSUM}(\mathbf{X}_i) < \mathbf{WSUM}(\mathbf{X}_j)$ . Se requiere que los valores de las diferentes funciones objetivo sean normalizados al mismo rango antes de aplicar esta técnica.

**Dominancia  $L$  ( $L$ -dominance, LDOM):** Esta nueva relación de dominancia fue propuesta en 2008 por Zou *et al.* [76]. **LDOM** utiliza las funciones  $n_b$  (ecuación 3.11),  $n_e$  (ecuación 3.12) y  $n_w$  (ecuación 3.13), que calculan, respectivamente, el número de objetivos en que una solución  $\mathbf{X}_i$  es mejor, igual y peor que otra solución  $\mathbf{X}_j$ .

Dadas dos soluciones  $\mathbf{X}_i$  y  $\mathbf{X}_j$ , se dice que  $\mathbf{X}_i$   $L$ -domina a  $\mathbf{X}_j$  ( $\mathbf{X}_i \prec_L \mathbf{X}_j$ ) si y sólo si:

$$\begin{aligned} n_b - n_w &= L > 0 \\ \wedge \quad \|\mathbf{F}(\mathbf{X}_i)\|_p &< \|\mathbf{F}(\mathbf{X}_j)\|_p \text{ (para cierta } p) \end{aligned} \quad (3.21)$$

donde  $\|\mathbf{F}(\mathbf{x}_i)\|_p$  es la  $p$ -norma del vector de objetivos para una solución  $\mathbf{X}_i$ . Para este trabajo se utilizó  $p = 1$  (distancia de Manhattan). La segunda desigualdad en la ecuación (3.21) requiere que los objetivos sean normalizados al mismo rango antes de aplicar esta técnica.

**Relación de preferencia (favour relation, FAVR):** Esta alternativa de comparación fue propuesta en 2001 por Drechsler *et al.* [27]. De acuerdo con la relación **FAVR**, se dice que una

---

<sup>3</sup>Dado que en este trabajo se asume que todos los objetivos tienen la misma importancia, los coeficientes de la ecuación (3.20) son simplemente omitidos.

solución  $\mathbf{X}_i$  es superior a otra solución  $\mathbf{X}_j$  ( $\mathbf{X}_i \prec_{\text{FAVR}} \mathbf{X}_j$ ) si y sólo si:

$$|\{m : f_m(\mathbf{X}_i) < f_m(\mathbf{X}_j)\}| > |\{n : f_n(\mathbf{X}_j) < f_n(\mathbf{X}_i)\}|$$

$$m, n = 1, 2, \dots, M \quad (3.22)$$

Esta relación no es transitiva, por ejemplo, consideremos las soluciones  $\mathbf{X}_i=(8,7,1)$ ,  $\mathbf{X}_j=(1,9,6)$  y  $\mathbf{X}_k=(7,0,9)$ . Es notorio que  $\mathbf{X}_i \prec_F \mathbf{X}_j \prec_F \mathbf{X}_k \prec_F \mathbf{X}_i$ , v.g., en estas soluciones existe un ciclo de dominancia. Dada esta propiedad de la relación, los autores proponen jerarquizar una población de soluciones de la siguiente manera: utilizar una representación de grafos, donde los nodos son las soluciones y los arcos corresponden a las relaciones de preferencia entre ellas. A partir de este grafo identificar los *componentes fuertemente conectados* (*Strongly Connected Components*, SCC). Un SCC está formado por aquellos elementos no comparables entre sí (como el ciclo del ejemplo anterior). Una vez identificados los SCCs, se construye un nuevo grafo libre de ciclos donde cada SCC será un nodo y los arcos estarán dados por las relaciones de preferencia existentes entre los elementos de cada par de SCCs. De esta manera, con este nuevo grafo ahora es posible establecer un ordenamiento al asignar la misma jerarquía a los individuos agrupados en un mismo SCC.

Dada la no transitividad de la relación **FAVR**, es común la situación en que la mayor parte de los individuos en una población sean parte del mismo ciclo y, por lo tanto, existiría una distribución escasa de jerarquías (lo que debilita el proceso de selección), siendo ésta la principal desventaja de la técnica.

En este trabajo investigaremos también el uso de la relación **FAVR** sin considerar el procedimiento de jerarquización descrito previamente. En este caso, la calidad de una solución  $\mathbf{X}_i$  estará dada por la cantidad de individuos que dominan a  $\mathbf{X}_i$  (de acuerdo con la relación **FAVR**). Se hará referencia al enfoque resultante con el término **FDOM**.

**Jerarquización por orden de eficiencia (preference order ranking, PORD):** di Pierro *et al.* [26] propusieron en 2007 una estrategia de jerarquización donde se considera el orden de eficiencia de las soluciones. Se dice que un individuo  $\mathbf{X}_i$  es eficiente de orden  $k$  si es no dominado (de acuerdo

con la dominancia de Pareto) en los  $\binom{M}{k}$  sub-espacios donde se consideran sólo  $k$  objetivos a la vez. De acuerdo con esta definición, es claro que la eficiencia de orden  $M$  para un problema con exactamente  $M$  objetivos corresponde a la definición de optimalidad de Pareto.

Si  $\mathbf{X}_i$  es eficiente de orden  $k$ , entonces también es eficiente de orden  $k + 1$  [15]. Del mismo modo, si  $\mathbf{X}_i$  no es eficiente de orden  $k$ , tampoco será eficiente de orden  $k - 1$  [26]. Dadas estas propiedades, el orden de eficiencia de una solución  $\mathbf{X}_i$  es el mínimo valor de  $k$  para el que  $\mathbf{X}_i$  es eficiente. Más formalmente:

$$\text{orden\_de\_eficiencia}(\mathbf{X}_i) = \min_{k=1}^M (k : \text{isEfficient}(\mathbf{X}_i, k)) \quad (3.23)$$

donde la función  $\text{isEfficient}(\mathbf{X}_i, k)$  es verdadera si  $\mathbf{X}_i$  es eficiente de orden  $k$ . El orden de eficiencia de cada solución puede ser utilizado directamente como jerarquía. Mientras más pequeño sea el orden de eficiencia, mejor será un individuo.

Los autores proponen usar esta estrategia en combinación con el ordenamiento por no dominancia (**PSRT**) de Goldberg descrito en la Sección 3.2.1. El esquema de jerarquización resultante se describe a continuación: inicialmente se aplica **PSRT** sobre la población, lo que clasifica a los individuos en distintos frentes de dominancia. Después, las soluciones que forman cada uno de los frentes son jerarquizadas de acuerdo a su orden de eficiencia, de una forma tal que las soluciones del segundo frente sean penalizadas con la peor jerarquía obtenida en el primer frente. Del mismo modo, las soluciones del tercer frente de dominancia son penalizadas con la peor jerarquía obtenida en el segundo, y así sucesivamente.

La principal desventaja de este enfoque es que su costo computacional se incrementa notablemente con el número de objetivos, de manera que su aplicación está limitada a problemas con una cantidad relativamente baja de criterios de optimización.

**Factor de ganancia (gaining factor, GFAC):** En 2006, Burke y Landa-Silva [9] propusieron esta relación de comparación que considera la ganancia de un individuo con respecto a otro. Dado un problema con 2 objetivos, si un individuo  $\mathbf{X}_i$  es mejor que otro individuo  $\mathbf{X}_j$  en el primer objetivo

( $f_1(\mathbf{X}_i) < f_1(\mathbf{X}_j)$ ), entonces la ganancia obtenida se calcula como  $g = (f_1(\mathbf{X}_j) - f_1(\mathbf{X}_i))/f_1(\mathbf{X}_j)$ . Se considera que  $\mathbf{X}_i$  es mejor que  $\mathbf{X}_j$  si en el segundo objetivo se cumple que  $f_2(\mathbf{X}_i) < f_2(\mathbf{X}_j) \cdot (1+g)$ , lo que de acuerdo con los autores equivale a la siguiente relación:

$$\frac{f_1(\mathbf{X}_i)}{f_1(\mathbf{X}_j)} + \frac{f_2(\mathbf{X}_i)}{f_2(\mathbf{X}_j)} < 2 \quad (3.24)$$

Para un problema con  $M$  objetivos, la relación **GFAC** indica que  $\mathbf{X}_i$  domina a  $\mathbf{X}_j$  ( $\mathbf{X}_i \prec_{\text{GFAC}} \mathbf{X}_j$ ) si y sólo si [49]:

$$\frac{f_1(\mathbf{X}_i)}{f_1(\mathbf{X}_j)} + \frac{f_2(\mathbf{X}_i)}{f_2(\mathbf{X}_j)} + \dots + \frac{f_M(\mathbf{X}_i)}{f_M(\mathbf{X}_j)} = \sum_{m=1}^M \frac{f_m(\mathbf{X}_i)}{f_m(\mathbf{X}_j)} < M \quad (3.25)$$

Los valores de los objetivos deben ser normalizados al mismo rango antes de que esta técnica sea aplicada.

**Dominancia sub-vector (subvector dominance, SVEC):** Köppen y Yoshida [47] propusieron expresar la calidad de un individuo en términos del máximo número de objetivos en los que es superado por algún otro individuo de la población. Formalmente:

$$\text{SVEC}(\mathbf{X}_i) = \max_{\mathbf{X}_j \neq \mathbf{X}_i} |\{m : f_m(\mathbf{X}_j) < f_m(\mathbf{X}_i)\}| \quad m = 1, 2, \dots, M \quad (3.26)$$

Se dará preferencia a aquellas soluciones que minimicen su valor para **SVEC**. Por lo tanto, se dice que una solución  $\mathbf{X}_i$  domina a otra solución  $\mathbf{X}_j$  con respecto a la relación **SVEC**, denotado por  $\mathbf{X}_i \prec_{\text{SVEC}} \mathbf{X}_j$ , si se cumple la siguiente condición:

$$\text{SVEC}(\mathbf{X}_i) < \text{SVEC}(\mathbf{X}_j) \quad (3.27)$$

**Dominancia del  $\epsilon$  mínimo (minimum  $\epsilon$  dominance, MEPS):** Köppen y Yoshida [47] propusieron también un método de jerarquización que corresponde a la forma aditiva de  $\epsilon$ -dominance [48]. De acuerdo con este enfoque, la aptitud de una solución  $\mathbf{X}_i$  está dada por el mínimo valor

$\epsilon$  (MEPS) tal que, al ser añadido a cada uno de sus objetivos,  $\mathbf{X}_i$  sea dominada por algún otro individuo de la población. El valor MEPS para un individuo  $\mathbf{X}_i$  puede calcularse de la siguiente manera:

$$\text{MEPS}(\mathbf{X}_i) = \min_{\mathbf{X}_j \neq \mathbf{X}_i} \left( \max_{m=1}^M (f_m(\mathbf{X}_j) - f_m(\mathbf{X}_i), 0) \right) \quad (3.28)$$

Dadas dos soluciones candidatas  $\mathbf{X}_i, \mathbf{X}_j \in P$ , se puede decir que  $\mathbf{X}_i$  domina a  $\mathbf{X}_j$  de acuerdo con este método ( $\mathbf{X}_i \prec_{\text{MEPS}} \mathbf{X}_j$ ) si se cumple que:

$$\text{MEPS}(\mathbf{X}_i) > \text{MEPS}(\mathbf{X}_j) \quad (3.29)$$

Este método requiere la normalización de los objetivos al mismo rango antes de aplicarse.

**Dominancia difusa (fuzzy pareto dominance, FPDM):** Köppen y Yoshida [47] propusieron un método que admite cierta incertidumbre en la magnitud de la diferencia que existe entre dos individuos en cada objetivo, de modo que evita basar la comparación de soluciones con respecto a un solo objetivo. La calidad de un individuo  $\mathbf{X}_i$  es entonces expresada en términos de la máxima *diferencia difusa* cuando se compara contra el resto de la población. Formalmente:

$$\text{FPDM}(\mathbf{X}_i) = \max_{\mathbf{X}_j \neq \mathbf{X}_i} \prod_{m=1}^M \left[ \frac{f_m(\mathbf{X}_i)}{f_m(\mathbf{X}_j)} \right] \quad (3.30)$$

$$\text{donde} \quad \left[ \frac{a}{b} \right] = \begin{cases} a/b & a < b \\ 1 & \text{de otro modo} \end{cases}$$

Las mejores soluciones serán aquellas con el valor más pequeño para la función FPDM. De modo que se considera que una solución  $\mathbf{X}_i$  domina a otra solución  $\mathbf{X}_j$  si y sólo si:

$$\text{FPDM}(\mathbf{X}_i) < \text{FPDM}(\mathbf{X}_j) \quad (3.31)$$

Se requiere que los valores de los diferentes objetivos sean normalizados al mismo rango antes

de la aplicación de esta técnica.

**Combinación de SVEC y MEPS (sub-objective dominance count, SODC):** Köppen y Yoshida [47] propusieron un método que considera tanto la cantidad de objetivos en los que una solución es mejor que otra, como la magnitud de dichas mejoras. SODC combina los métodos SVEC y MEPS previamente descritos, de modo que se consideren simultáneamente múltiples criterios de jerarquización. El enfoque resultante se describe mediante el Algoritmo 1.

---

**Algoritmo 1** Método SODC.

---

**DEFINE** subobjective\_dominance\_count()

```

1: for all  $\mathbf{X}_i \in P$  do
2:    $S_{\mathbf{X}_i} \leftarrow \{(M - \text{svec}(\mathbf{X}_i, \mathbf{X}_j), \text{meps}(\mathbf{X}_i, \mathbf{X}_j)) : \mathbf{X}_j \neq \mathbf{X}_i\}$ 
3:    $POS_{\mathbf{X}_i} \leftarrow \text{pareto\_set}(S_{\mathbf{X}_i})$ 
4: for all  $\mathbf{X}_i \in P$  do
5:    $\text{SODC}(\mathbf{X}_i) = |\{\mathbf{X}_j : \mathbf{X}_i \in POS_{\mathbf{X}_j} \wedge \mathbf{X}_j \neq \mathbf{X}_i\}|$ 

```

**END**

donde:

$$\text{svec}(\mathbf{X}_i, \mathbf{X}_j) = |\{m : f_m(\mathbf{X}_j) < f_m(\mathbf{X}_i)\}| \quad m = 1, 2, \dots, M$$

$$\text{meps}(\mathbf{X}_i, \mathbf{X}_j) = \max_{m=1}^M (f_m(\mathbf{X}_j) - f_m(\mathbf{X}_i), 0)$$


---

Como puede apreciarse en el Algoritmo 1, para cada solución  $\mathbf{X}_i$  se forma un conjunto  $S_{\mathbf{X}_i}$  con todas las parejas formadas por los valores obtenidos para ambas métricas, al comparar  $\mathbf{X}_i$  con el resto de las soluciones. Posteriormente, el conjunto de Pareto de  $S_{\mathbf{X}_i}$  (denotado por  $POS_{\mathbf{X}_i}$ ) es identificado.  $POS_{\mathbf{X}_i}$  se forma con las soluciones que mostraron el mejor desempeño en ambas métricas al compararse contra  $\mathbf{X}_i$ . Finalmente,  $\text{SODC}(\mathbf{X}_i)$  expresará la calidad de una solución como el número de ocasiones en que aparece en los diferentes  $POS_{\mathbf{X}_j}$  para cada  $\mathbf{X}_j \neq \mathbf{X}_i$ .

Se dice que un individuo  $\mathbf{X}_i$  domina a otro individuo  $\mathbf{X}_j$  (con respecto al método SODC) si se cumple que:

$$\text{SODC}(\mathbf{X}_i) > \text{SODC}(\mathbf{X}_j) \quad (3.32)$$

**Record de ganancia (winning score, WSCR):** El método WSCR fue propuesto en 2006

por Maneeratana *et al.* [53] como parte de su algoritmo COGA (*Compressed Objective Genetic Algorithm*). **WSCR** es un criterio de maximización que contabiliza el número de objetivos en los que un individuo es superior e inferior a otro. Formalmente, la calidad de un individuo  $\mathbf{X}_i$  está dada por:

$$\text{WSCR}(\mathbf{X}_i) = \sum_{\mathbf{X}_j \neq \mathbf{X}_i} |\{m : f_m(\mathbf{X}_i) < f_m(\mathbf{X}_j)\}| - |\{n : f_n(\mathbf{X}_i) > f_n(\mathbf{X}_j)\}| \quad (3.33)$$

$$m = 1, 2, \dots, M \quad (3.34)$$

Se dice que una solución  $\mathbf{X}_i$  domina a otra solución  $\mathbf{X}_j$  con respecto a **WSCR** si y sólo si:

$$\text{WSCR}(\mathbf{X}_i) > \text{WSCR}(\mathbf{X}_j) \quad (3.35)$$

### 3.3 Enfoques propuestos

#### 3.3.1 Composiciones jerárquicas

En esta sección se proponen diferentes composiciones jerárquicas (CJs) resultantes de este proyecto de investigación. Se asume la obtención previa del vector  $\mathbf{R}(\mathbf{X}_i)$  para cada  $\mathbf{X}_i \in P$  (ver Sección 3.1).

**Mínima jerarquía (minimum ranking, MINR<sup>\*4</sup>):** En la Sección 3.2.2 se mencionaron algunos inconvenientes del método **MAXR** propuesto por Bentley y Wakefield [4]. **MAXR** clasifica las soluciones de acuerdo con la mejor jerarquía que obtuvieron para los diferentes objetivos, de modo que asigna una alta jerarquía a individuos con buen desempeño para algún objetivo aún cuando tienen un mal desempeño global (favorece extremos). Basándonos en estas observaciones y de manera opuesta a **MAXR**, nosotros proponemos utilizar la peor jerarquía de un individuo como su valor de

---

<sup>4</sup>Usamos un \* para distinguir los métodos propuestos como resultado de este proyecto.

aptitud. La calidad de una solución  $X_i$  se calcula entonces de la siguiente manera:

$$\text{MINR}^*(X_i) = \max_{m=1}^M r_m(X_i) \quad (3.36)$$

$\text{MINR}^*$  es un criterio de minimización, por lo que se dice que una solución  $X_i$  domina a otra solución  $X_j$ , con respecto a la relación  $\text{MINR}^*$  ( $X_i \prec_{\text{MINR}^*} X_j$ ), si y sólo si:

$$\text{MINR}^*(X_i) < \text{MINR}^*(X_j) \quad (3.37)$$

A diferencia de  $\text{MAXR}$ ,  $\text{MINR}^*$  es capaz de filtrar aquellas soluciones con mal desempeño en cualquier objetivo, favoreciendo individuos con un mejor compromiso.

**Comparación sucesiva de jerarquías ordenadas (successive comparison of ordered ranks,  $\text{SCOR}^*$ ):** Este nuevo método de jerarquización se describe mediante el Algoritmo 2.

---

**Algoritmo 2** Método  $\text{SCOR}^*$ .

---

**DEFINE**  $\text{SCOR\_Ranking}()$

- 1:  $R'(X_i) \leftarrow \text{ordenar\_descendentemente}(R(X_i)) \quad \forall X_i \in P$
- 2:  $\text{SCOR}^*(X_i) \leftarrow 0 \quad \forall X_i \in P$
- 3: **for all**  $(X_i, X_j) : X_i, X_j \in P \wedge X_i \neq X_j$  **do**
- 4:   **for**  $m = 1$  to  $M$  **do**
- 5:     **if**  $r'_m(X_i) < r'_m(X_j)$  **then**
- 6:        $\text{SCOR}^*(X_j) \leftarrow \text{SCOR}^*(X_j) + 1$
- 7:        $m \leftarrow M$  (terminar ciclo de la línea 4)
- 8:     **else if**  $r'_m(X_j) < r'_m(X_i)$  **then**
- 9:        $\text{SCOR}^*(X_i) \leftarrow \text{SCOR}^*(X_i) + 1$
- 10:       $m \leftarrow M$  (terminar ciclo de la línea 4)

**END**

---

Inicialmente, el vector de jerarquías  $R(X_i)$  se ordena descendientemente para cada  $X_i \in P$ . Después, cada pareja de soluciones se compara elemento a elemento con respecto a su vector ordenado de jerarquías, de tal modo que la primera jerarquía en la que una de las soluciones sea mejor que la otra determinará la relación de preferencia entre ellas. Una vez terminado este proceso,  $\text{SCOR}^*(X_i)$  expresará la calidad de una solución  $X_i$  en términos del número de individuos que de acuerdo con



esta comparativa fueron mejores que  $X_i$ . De esta manera, se considera que  $X_i$  domina a  $X_j$  si se cumple que  $SCOR^*(X_i) < SCOR^*(X_j)$ .

### 3.3.2 Formas relajadas de dominancia

**Comparación sucesiva de objetivos ordenados (successive comparison of ordered objectives, SCOO\*):** SCOO\* utiliza el procedimiento de jerarquización propuesto para el método SCOR\* descrito en la Sección 3.4, pero usando los valores de los objetivos en lugar de las posiciones jerárquicas de los individuos. Inicialmente, se ordena descendentemente el vector de objetivos  $F(X_i)$  para cada solución  $X_i$  en la población. Después, cada pareja de soluciones se compara elemento a elemento con respecto a su vector ordenado de objetivos, de tal modo que la primera comparación en la que una solución sea mejor que la otra permitirá establecer preferencias entre ellas. Durante este proceso se llevará un contador  $SCOO^*(X_i)$  (para cada  $X_i \in P$ ) que será incrementado por cada individuo que de acuerdo con esta comparación es superior a  $X_i$ . Este procedimiento se detalla en el Algoritmo 3.

---

#### Algoritmo 3 Método SCOO\*.

---

**DEFINE** SCOO\_Ranking()

```

1:  $F'(X_i) \leftarrow \text{ordenar\_descendentemente}(F(X_i)) \quad \forall X_i \in P$ 
2:  $SCOO^*(X_i) \leftarrow 0 \quad \forall X_i \in P$ 
3: for all  $(X_i, X_j) : X_i, X_j \in P \wedge X_i \neq X_j$  do
4:   for  $m = 1$  to  $M$  do
5:     if  $f'_m(X_i) < f'_m(X_j)$  then
6:        $SCOO^*(X_j) \leftarrow SCOO^*(X_j) + 1$ 
7:        $m \leftarrow M$  (terminar ciclo de la línea 4)
8:     else if  $f'_m(X_j) < f'_m(X_i)$  then
9:        $SCOO^*(X_i) \leftarrow SCOO^*(X_i) + 1$ 
10:       $m \leftarrow M$  (terminar ciclo de la línea 4)
```

**END**

---

De acuerdo con el método SCOO\*, se dice que una solución  $X_i$  domina a otra solución  $X_j$ , denotado por  $X_i \prec_{SCOO^*} X_j$ , si se cumple que  $SCOO^*(X_i) < SCOO^*(X_j)$ . Para aplicar este método los objetivos deben estar normalizados en el mismo rango.

**Fallos globales (global failures, GFLS\*):** GFLS\* penaliza a los individuos por cada ocasión en que son superados al compararse, objetivo por objetivo, con respecto al resto de la población. Más formalmente, la cantidad global de fallos de un individuo  $\mathbf{X}_i$  está dada por:

$$\text{GFLS}^*(\mathbf{X}_i) = \sum_{\mathbf{X}_j \neq \mathbf{X}_i} |\{m : f_m(\mathbf{X}_j) < f_m(\mathbf{X}_i)\}| \quad m = 1, 2, \dots, M \quad (3.38)$$

De acuerdo con GFLS\*, una solución  $\mathbf{X}_i$  se considera mejor que otra solución  $\mathbf{X}_j$ , si y sólo si:

$$\text{GFLS}^*(\mathbf{X}_i) < \text{GFLS}^*(\mathbf{X}_j) \quad (3.39)$$

**Victorias globales (global wins, GWIN\*):** De manera opuesta al método GFLS\* descrito previamente, GWIN\* es un criterio de maximización que acumula la cantidad de objetivos en los que un individuo  $\mathbf{X}_i$  es superior a cada otro individuo de la población. Formalmente:

$$\text{GWIN}^*(\mathbf{X}_i) = \sum_{\mathbf{X}_j \neq \mathbf{X}_i} |\{m : f_m(\mathbf{X}_i) < f_m(\mathbf{X}_j)\}| \quad m = 1, 2, \dots, M \quad (3.40)$$

Dadas dos soluciones  $\mathbf{X}_i, \mathbf{X}_j \in P$ , se dice que  $\mathbf{X}_i$  domina a  $\mathbf{X}_j$  si se cumple que:

$$\text{GWIN}^*(\mathbf{X}_i) < \text{GWIN}^*(\mathbf{X}_j) \quad (3.41)$$

**Deterioro global (global detriment, GDET\*):** El método GFLS\* descrito previamente únicamente toma en cuenta la cantidad de objetivos en los que una solución es mejor que otra; sin embargo, un aspecto de gran importancia es medir qué tan significativa es la diferencia entre ellas. GDET\* penaliza a cada individuo por la magnitud de la diferencia con que es superado cuando se compara objetivo por objetivo con respecto al resto de la población. Formalmente, el deterioro global de un individuo  $\mathbf{X}_i$  está dado por:

$$\text{GDET}^*(\mathbf{X}_i) = \sum_{\mathbf{X}_j \neq \mathbf{X}_i} \sum_{m=1}^M \max(f_m(\mathbf{X}_i) - f_m(\mathbf{X}_j), 0) \quad (3.42)$$

Se considera que una solución  $\mathbf{X}_i$  domina a otra solución  $\mathbf{X}_j$  con respecto a este enfoque si y sólo si  $\mathbf{GDET}^*(\mathbf{X}_i) < \mathbf{GDET}^*(\mathbf{X}_j)$ .  $\mathbf{GDET}^*$  es susceptible al rango de los objetivos, por lo que deben ser normalizados antes de su aplicación.

**Mejora global (global improvement, GIMP\*):** De manera similar a la técnica  $\mathbf{GDET}^*$  previamente descrita,  $\mathbf{GIMP}^*$  es un método que toma en cuenta qué tan significativamente es mejor una solución que otra. Sin embargo, de manera opuesta a  $\mathbf{GDET}^*$ ,  $\mathbf{GIMP}^*$  es un criterio de maximización que consiste en acumular la magnitud de la diferencia con la que un individuo es superior en cada uno de sus objetivos, frente a cada otro individuo en la población. De manera formal,  $\mathbf{GIMP}^*$  expresa la calidad de una solución  $\mathbf{X}_i$  de la siguiente manera:

$$\mathbf{GIMP}^*(\mathbf{X}_i) = \sum_{\mathbf{X}_j \neq \mathbf{X}_i} \sum_{m=1}^M \max(f_m(\mathbf{X}_j) - f_m(\mathbf{X}_i), 0) \quad (3.43)$$

Sean  $\mathbf{X}_i, \mathbf{X}_j \in P$  dos soluciones candidatas, se dice que  $\mathbf{X}_i$  domina a  $\mathbf{X}_j$  de acuerdo con  $\mathbf{GIMP}^*$  si se cumple que  $\mathbf{GIMP}^*(\mathbf{X}_i) > \mathbf{GIMP}^*(\mathbf{X}_j)$ . Los diferentes objetivos deben ser normalizados al mismo rango antes de aplicar este método.

**Margen de ganancia (profit, PRFT\*):**  $\mathbf{PRFT}^*$  expresa la calidad de las soluciones en términos del margen entre su máxima ganancia y su máxima pérdida. Para clarificar lo anterior, el margen de ganancia para un individuo  $\mathbf{X}_i$  es expresado formalmente por la ecuación (3.44):

$$\mathbf{PRFT}^*(\mathbf{X}_i) = \max_{\mathbf{X}_j \neq \mathbf{X}_i} (\text{ganancia}(\mathbf{X}_i, \mathbf{X}_j)) - \max_{\mathbf{X}_j \neq \mathbf{X}_i} (\text{ganancia}(\mathbf{X}_j, \mathbf{X}_i)) \quad (3.44)$$

donde  $\text{ganancia}(\mathbf{X}_i, \mathbf{X}_j)$  es la ganancia de  $\mathbf{X}_i$  con respecto a  $\mathbf{X}_j$ :

$$\text{ganancia}(\mathbf{X}_i, \mathbf{X}_j) = \sum_{m=1}^M \max(f_m(\mathbf{X}_j) - f_m(\mathbf{X}_i), 0) \quad (3.45)$$

Se dará preferencia a las soluciones que maximicen su valor para  $\mathbf{PRFT}^*$ . De este modo, se dice que

una solución  $\mathbf{X}_i$  domina a otra solución  $\mathbf{X}_j$  si y sólo si  $\mathbf{PRFT}^*(\mathbf{X}_i) > \mathbf{PRFT}^*(\mathbf{X}_j)$ . Este método debe aplicarse sobre los objetivos normalizados al mismo rango.

**Distancia a la mejor solución conocida (distance to the best known solution, GBST\*):**

La mejor solución conocida, a la que haremos referencia como punto GBEST, es aquella formada por el mejor valor conocido para los diferentes objetivos; v.g. GBEST domina a todas las soluciones conocidas desde el inicio de la búsqueda. La figura 3.2 ilustra un ejemplo del cálculo del punto GBEST sobre un conjunto de soluciones conocidas para un problema con dos objetivos.

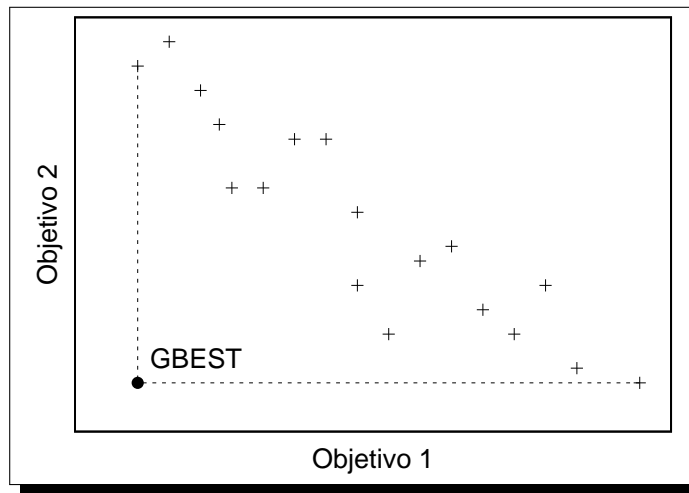


Figura 3.2: Punto GBEST: Mejor solución conocida.

El punto GBEST no debe confundirse con el *vector de objetivos ideal* (*ideal objective vector*) cuyos elementos corresponden al valor óptimo para cada dimensión considerando todas las soluciones dentro de la región factible [16].

Este método consiste en calcular la cercanía de cada individuo al punto GBEST (este trabajo usa la distancia Euclidiana), de modo que mientras más reducida sea esta distancia, mejor será un individuo. Se requiere que los objetivos estén normalizados antes de aplicar esta técnica.

**Distancia a la peor solución local (distance to the local worst solution, LWST\*):** La peor solución local, o punto LWORST, es aquella que es dominada por todas las soluciones en la

población actual. El punto LWORST está formado por el peor valor para cada objetivo, considerando solamente la población actual. La figura 3.3 ilustra un ejemplo del cálculo del punto LWORST para un problema con dos objetivos.

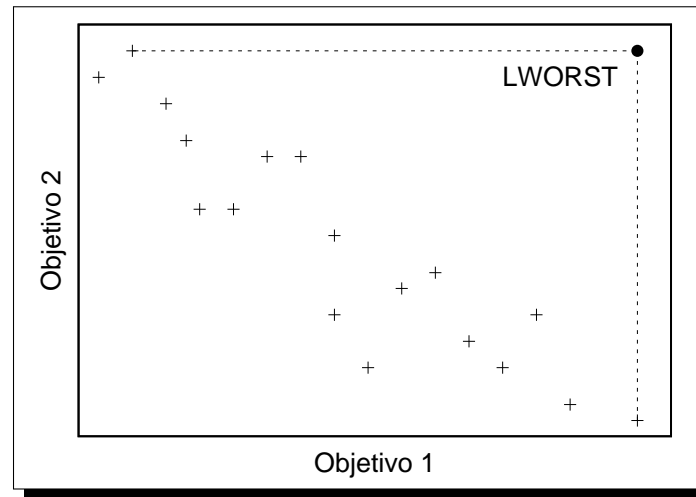


Figura 3.3: Punto LWORST: Actual peor solución.

El punto LWORST no debe confundirse con el punto conocido como *vector de objetivos nadir* (*nadir objective vector*) cuyos elementos corresponden al peor valor para cada objetivo considerando únicamente los puntos del conjunto de óptimos de Pareto [16].

Mientras más grande sea la distancia de una solución con respecto al punto LWORST, mejor será la calidad de dicha solución. Se requiere que los objetivos se encuentren normalizados en el mismo rango antes de aplicar esta técnica.

**Distancia a la peor solución global (distance to the global worst solution, GWST\*):** De manera similar al método **LWST\*** descrito anteriormente, este método calcula la distancia a la peor solución. Sin embargo, en este caso haremos referencia a dicha solución como el punto GWORST, ya que estará formada por el peor valor conocido globalmente para los diferentes objetivos (desde el inicio de la búsqueda). Se dará preferencia a aquellas soluciones que maximicen su distancia al punto GWORST (este trabajo usa la distancia Euclidiana).

El punto GWORST no debe confundirse con el punto conocido como *vector de objetivos nadir* (*nadir objective vector*) cuyos elementos corresponden al peor valor para cada objetivo considerando

únicamente los puntos del conjunto de óptimos de Pareto [16].

**Jerarquización por compromiso (trade-off ranking, TOFF\*):** En la ausencia de preferencias puede asumirse que el tomador de decisiones estará interesado en soluciones con el mejor compromiso entre objetivos. Tomando esto en cuenta, **TOFF\*** es un método de jerarquización que está basado en medir el nivel de compromiso de cada solución. Para estos propósitos, **TOFF\*** calcula la aptitud de un determinado individuo como el acumulado de las diferencias entre el valor de cada pareja de sus objetivos. De manera formal, el grado de compromiso de una solución  $\mathbf{X}_i$  está dado por:

$$\text{TOFF}^*(\mathbf{X}_i) = \sum_{m=1}^{M-1} \sum_{n=m+1}^M |f_m(\mathbf{X}_i) - f_n(\mathbf{X}_i)| \quad (3.46)$$

De acuerdo con este enfoque, se dice que una solución  $\mathbf{X}_i$  es superior a otra solución  $\mathbf{X}_j$  si se cumple que  $\text{TOFF}^*(\mathbf{X}_i) < \text{TOFF}^*(\mathbf{X}_j)$ . Está claro que para el caso de problemas de minimización, favorecer valores pequeños para  $\text{TOFF}^*(\mathbf{X}_i)$  permitirá también dar preferencia a soluciones con la mejor convergencia, ya que al aproximarse al frente de Pareto las diferencias entre el valor de los diferentes objetivos se vuelven cada vez menos significativas. De esta manera, proponemos a **TOFF\*** como una alternativa que permite lograr una buena convergencia y al mismo tiempo dirigir la búsqueda hacia regiones que se asumen son de mayor interés, v.g. el mejor compromiso.

**Ganancia total (total gain, GAIN\*):** Basándonos en la relación **GFAC** de Burke y Landa-Silva [9] descrita en la Sección 3.2.3, nosotros proponemos una forma diferente para establecer preferencias entre soluciones. Empíricamente, con pequeñas instancias hemos observado que esta alternativa permite una jerarquización de grano más fino que la relación original, por lo que optamos contemplarla en este trabajo. Sean  $\mathbf{X}_i$  y  $\mathbf{X}_j$  dos individuos, la *ganancia total* de  $\mathbf{X}_i$  con respecto a  $\mathbf{X}_j$  está dada por:

$$\text{GAIN}^*(\mathbf{X}_i, \mathbf{X}_j) = \sum \frac{f_m(\mathbf{X}_j) - f_m(\mathbf{X}_i)}{f_m(\mathbf{X}_j)} \quad \forall m \in \{1, 2, \dots, M\} : f_m(\mathbf{X}_i) < f_m(\mathbf{X}_j) \quad (3.47)$$

Se considera que  $X_i$  domina a  $X_j$  ( $X_i \prec_{\text{GAIN}^*} X_j$ ) si la ganancia total de  $X_i$  con respecto a  $X_j$  es mayor a la ganancia total de  $X_j$  con respecto a  $X_i$ :

$$X_i \prec_{\text{GAIN}^*} X_j \Leftrightarrow \text{GAIN}^*(X_i, X_j) > \text{GAIN}^*(X_j, X_i) \quad (3.48)$$

### 3.4 No equivalencia entre AVER, WSCR y GFLS\*

Corne y Knowles demostraron en 2007 [12] la equivalencia de los métodos AVER [4] y WSCR [53]. Del mismo modo, nuestro enfoque GFLS\* presentado en la Sección 3.3.2 proporciona resultados similares a los de ambos métodos. Consideremos como ejemplo las 6 soluciones para un problema con 3 objetivos mostradas en la tabla 3.1. La tabla 3.1 desglosa los resultados de aplicar cada uno de los tres métodos sobre este conjunto de soluciones<sup>5</sup>. Así mismo, se muestra la jerarquía (ordenamiento) final donde se transformó el resultado obtenido con cada método al rango  $[1, N]$ ; 1 y  $N$  para la mejor y peor solución respectivamente, considerando que AVER y GFLS\* son criterios de minimización mientras que, por el contrario, WSCR es un criterio de maximización.

Tabla 3.1: Equivalencia AVER, WSCR y GFLS\*.

Solución ( $f_1, f_2, f_3$ )	Jerarquía $f_1$	Jerarquía $f_2$	Jerarquía $f_3$	Resultado método → Jerarquía final		
				AVER (-)	WSCR (+)	GFLS* (-)
(1, 5, 6)	1	4	6	11 → 3	-1 → 3	8 → 3
(7, 1, 2)	6	1	2	9 → 2	3 → 2	6 → 2
(2, 4, 4)	2	3	4	9 → 2	3 → 2	6 → 2
(5, 3, 1)	4	2	1	7 → 1	7 → 1	4 → 1
(6, 8, 5)	5	5	5	15 → 5	-9 → 5	12 → 5
(3, 9, 3)	3	6	3	12 → 4	-3 → 4	9 → 4

Tal como se observa en la tabla 3.1, la jerarquía final lograda con los tres métodos es la misma. La equivalencia de nuestro enfoque GFLS\* con respecto al método AVER puede plantearse mediante la ecuación (3.49):

$$\text{GFLS}^*(X_i) = \text{AVER}(X_i) - M \quad (3.49)$$

<sup>5</sup>Para los detalles del cálculo de los resultados mostrados en la tabla 3.1, referirse a la sección correspondiente a cada uno de los métodos considerados.

Aunque de manera general los resultados de estos tres métodos son similares, existen casos específicos en los que cada enfoque discrimina de forma diferente. Como contraejemplo de esta equivalencia consideremos ahora las soluciones en la tabla 3.2. Este contraejemplo es una modificación del ejemplo presentado en la tabla 3.1; el valor del primer objetivo ( $f_1$ ) para la primera y última solución han sido modificados. En este tipo de escenarios, donde existen valores repetidos para alguno(s) de los objetivos, el orden de preferencias que cada uno de estos tres enfoques induce puede ser diferente. En este caso, el orden de preferencia para la segunda y tercer solución es diferente para cada enfoque; mientras que para **AVER** la segunda solución es mejor que la tercera, **WSCR** las considera equivalentes y nuestro método **GFLS\*** establece una preferencia opuesta a **AVER** (las diferencias han sido resaltadas con un \* en la tabla 3.2).

Tabla 3.2: Equivalencia **AVER**, **WSCR** y **GFLS\***: Contraejemplo.

Solución ( $f_1, f_2, f_3$ )	Jerarquía $f_1$	Jerarquía $f_2$	Jerarquía $f_3$	Resultado método → Jerarquía final		
				<b>AVER</b> (-)	<b>WSCR</b> (+)	<b>GFLS*</b> (-)
(2, 5, 6)	1	4	6	11 → 5	-3 → 4	8 → 5
(7, 1, 2)	4	1	2	7 → 2*	3 → 2*	6 → 3*
(2, 4, 4)	1	3	4	8 → 3*	3 → 2*	5 → 2*
(5, 3, 1)	2	2	1	5 → 1	7 → 1	4 → 1
(6, 8, 5)	3	5	5	13 → 6	-9 → 5	12 → 6
(2, 9, 3)	1	6	3	10 → 4	-1 → 3	7 → 4

Con este contraejemplo se demuestra que los tres enfoques mencionados no son equivalentes. Sin embargo, los resultados que estos métodos le permitan obtener a un determinado MOEA serán, por lo general, similares.



# 4

## Métodos de asignación de aptitud: Experimentación

### 4.1 Introducción

Con la finalidad de investigar las propiedades de cada uno de los enfoques descritos en el Capítulo 3, en este capítulo se presentan los resultados de una serie de experimentos comparativos. Estos experimentos se realizaron considerando los problemas DTLZ1-DTLZ6 descritos en la Sección 2.5, ya que estos pueden ser escalados a cualquier número de objetivos y variables de decisión. El número total de variables de decisión para estos problemas es  $n = M + k - 1$ , donde  $M$  es la cantidad de objetivos. Se utilizó el valor de  $k = 5$  para el problema DTLZ1, y  $k = 10$  para el resto de los problemas. Dado que el interés particular de este trabajo de tesis radica en el incremento en la cantidad de objetivos, los experimentos se realizaron considerando instancias con  $M = \{5, 10, 15, 20, 30, 50\}$  objetivos para los diferentes problemas de prueba. Sin embargo, ya que el método **PORD** descrito en la Sección 3.2.3.2 se vuelve computacionalmente costoso con el aumento en el número de objetivos (ver Sección 4.3), este enfoque sólo se aplicó para instancias con un máximo de 20 objetivos.

## 4.2 Distribución de jerarquías

Corne y Knowles [12] propusieron calcular la efectividad de un método de asignación de aptitud mediante la medición de la *entropía relativa* (*relative entropy*,  $ER$ ) de la distribución de jerarquías que se induce sobre un conjunto de soluciones. Dada una población con  $N$  soluciones jerarquizadas, donde existen como máximo  $N$  jerarquías (cada individuo tiene una jerarquía diferente) y como mínimo 1 (todos los individuos comparten la misma jerarquía), la  $ER$  de la distribución de jerarquías  $D$  está dada por la ecuación (4.1):

$$ER(D) = \frac{\sum_r \frac{D_r}{N} \log\left(\frac{D_r}{N}\right)}{\log(1/N)} \quad (4.1)$$

donde  $D_r$  denota la cantidad de soluciones con jerarquía  $r$ .  $ER(D)$  tiende a 1 al aproximarse a la situación ideal en la que cada solución tiene una jerarquía diferente (ordenamiento completo de la población). Por otro lado, el peor caso sería aquel en el que todos los individuos comparten la misma clasificación, situación en la que  $ER(D)$  es igual a cero. De acuerdo con los autores, se puede decir que un método que provea un ordenamiento de grano fino favorecerá el proceso de selección y, por lo tanto, permitirá un mejor desempeño para el esquema de optimización que lo utilice.

López Jaimes *et al.* [52] realizaron un estudio similar donde analizaron los histogramas de la distribución de jerarquías lograda por distintos enfoques. De acuerdo con los autores, una forma de medir la escalabilidad de un método es analizar si la forma y rango de su distribución de jerarquías se mantiene al incrementarse la cantidad de objetivos. En acuerdo con lo expuesto por Corne and Knowles [12], López Jaimes *et al.* argumentan que un método de jerarquización favorecerá el proceso de selección si es capaz de generar un amplio rango de jerarquías.

La cantidad de métodos, problemas de prueba y cantidades de objetivos considerados para este estudio hacen prohibitivo analizar los histogramas con la distribución de jerarquías para cada configuración. En este trabajo nos concentraremos en la  $ER$  de Corne y Knowles [12] para analizar el potencial de discriminación entre soluciones que provee cada uno de los enfoques estudiados.

La figura 4.1 muestra la *ER* de la distribución de jerarquías lograda por cada una de las técnicas que se estudian. Para este experimento se generaron aleatoriamente 1,000 vectores de decisión y se evaluaron en los problemas DTLZ1-DTLZ6. Para cada experimento se realizaron 31 ejecuciones independientes. Dada la similitud en los resultados para los diferentes problemas de prueba, y con la finalidad de optimizar el espacio, los datos de la figura 4.1 corresponden al promedio de los resultados para los seis problemas adoptados, de manera que se analiza el comportamiento global de los enfoques.

A partir de la figura 4.1 es posible identificar comportamientos particulares de algunos de los enfoques estudiados. Los métodos basados en la dominancia de Pareto (**PRNK** y **PSRT**) presentan una entropía baja incluso desde las instancias más pequeñas. Sin embargo, es notable que cuando se incrementó la cantidad de objetivos, los niveles de entropía logrados por **PRNK** y **PSRT** se deterioraron significativamente, lo que confirma que la dominancia de Pareto no es capaz de proveer una discriminación de soluciones adecuada en espacios de objetivos con grandes dimensiones. **PSRT** es en todos los casos ligeramente mejor que **PRNK** al jerarquizar una población aleatoria; sin embargo, tal como fue expuesto en la Sección 2.4.2, al progresar la búsqueda de un MOEA la población se satura de soluciones no dominadas (de acuerdo con la dominancia de Pareto), por lo que la jerarquización de ambos enfoques normalmente será similar.

Aún cuando **PORD** solo fue aplicado para instancias con un máximo de 20 objetivos, a partir de sus resultados en instancias pequeñas es posible señalar que debido a que discrimina entre soluciones con una misma clasificación, sus niveles de entropía serán por lo general ligeramente mejores a los de **PSRT**.

**FAVR** es el método que mostró los niveles más bajos de entropía en todas las gráficas de la figura 4.1. Este comportamiento de **FAVR** se atribuye a su mecanismo propuesto de jerarquización, en el que, dado que no es una relación transitiva, se identifican ciclos en el grafo de preferencias resultante (ver detalles en Sección 3.2.3.2). Es común la situación en que gran parte de los individuos de la población formen parte de un mismo ciclo y, de este modo, compartan la misma jerarquía. Por otra parte, **FDOM**, que utiliza la misma relación de preferencia pero no el método propuesto de jerarquización, mostró niveles altos de entropía y consistencia cuando se incrementó la cantidad de

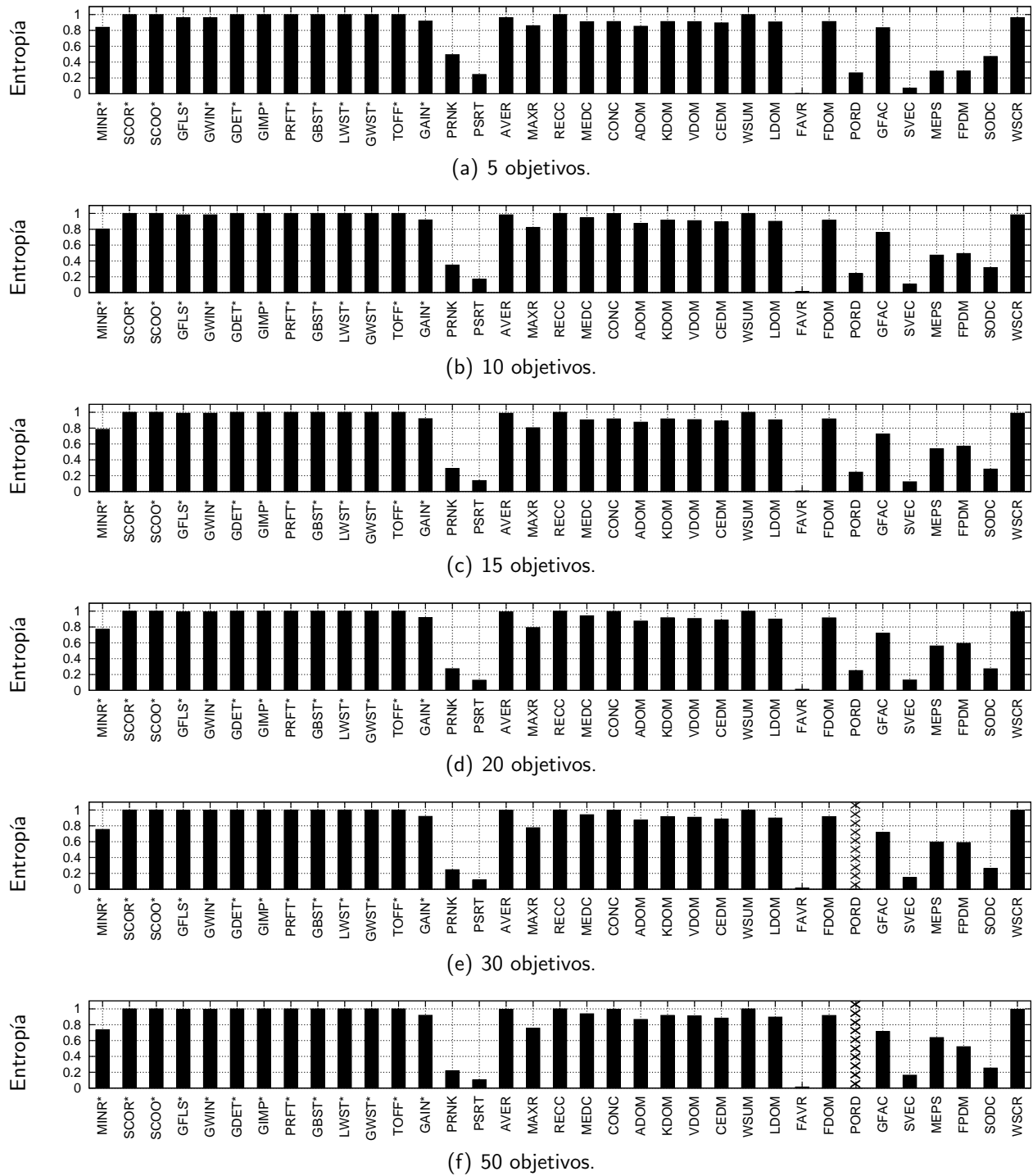


Figura 4.1: Entropía promedio de la distribución de jerarquías para una población de 1000 soluciones generadas aleatoriamente. Problemas DTLZ1-DTLZ6.

objetivos.

SVEC es otra técnica que mostró niveles bajos de entropía en todas las instancias de este experimento. Este comportamiento está dado por la naturaleza misma del método, ya que SVEC asigna como aptitud de un individuo el mayor número de criterios en los que es superado por algún otro individuo, de modo que habrá como máximo  $M$  jerarquías diferentes (donde  $M$  es la cantidad de objetivos). De este modo, la figura 4.1 muestra que al aumentarse la cantidad de objetivos también aumentó el desempeño de este método, no obstante, sus niveles de entropía nunca llegaron a ser competitivos.

De los métodos propuestos como parte de este proyecto de investigación, MINR\* es el método que mostró el nivel más bajo de entropía, mismo que se deterioró ligeramente de forma gradual con el incremento en la cantidad de objetivos. Sin embargo, sus niveles de entropía fueron en todos los casos competitivos con respecto al resto de los enfoques.

Fuera de estas observaciones particulares, se puede decir que los distintos enfoques propuestos, al igual que algunos de los enfoques del estado del arte, mostraron buenos niveles de entropía y escalabilidad con respecto al número de objetivos.

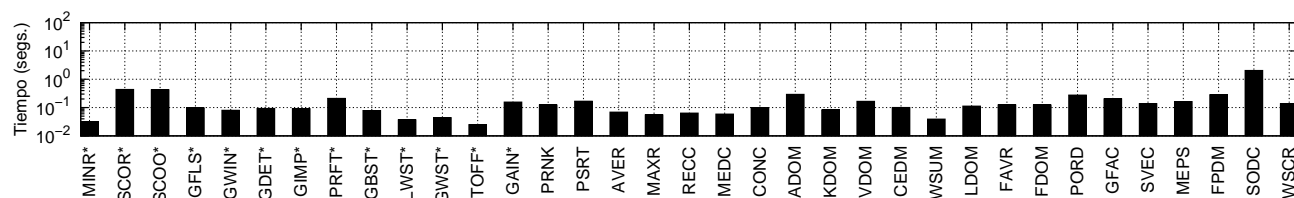
Los resultados de este experimento pueden ser vistos como una guía sobre el comportamiento relativo que tendría un MOEA al incorporar los distintos esquemas de jerarquización. Sin embargo, el hecho de que un método provea una alta discriminación entre soluciones no garantiza que brindará efectividad al guiar el proceso de búsqueda, ya que esta discriminación podría ser errónea y dar preferencia a soluciones no necesariamente con el mejor desempeño. Por lo tanto, en la siguiente sección se estudiará de manera empírica la capacidad que tiene cada alternativa para guiar la búsqueda hacia el óptimo para este mismo conjunto de funciones de prueba.

## 4.3 Tiempo de cómputo

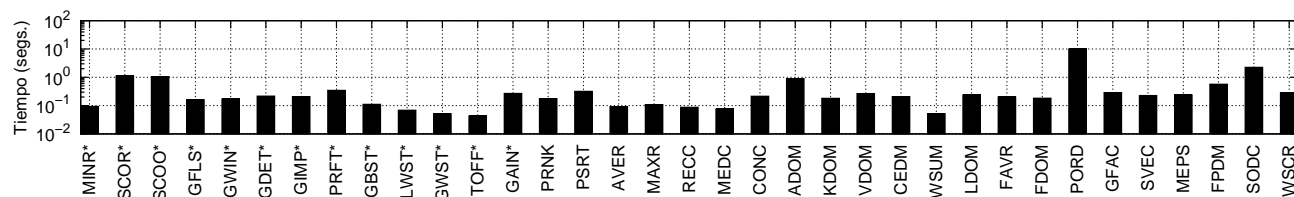
Un aspecto sin duda importante es el tiempo de cómputo que requiere cada una de las alternativas de jerarquización estudiadas en relación con el incremento del número de objetivos. En este experimento se midió el tiempo que le tomó a cada enfoque para jerarquizar 100 veces una población de 100 soluciones generadas aleatoriamente en el espacio de las funciones objetivo. Se consideraron 6 cantidades de objetivos:  $M = \{5, 10, 15, 20, 30, 50\}$ . Los resultados de este experimento se muestran en la figura 4.2.

A partir de la figura 4.2 es posible apreciar que la mayor parte de los enfoques estudiados demandan una cantidad razonable de tiempo de cómputo. PORD es el método más costoso (computacionalmente hablando), ya que sus tiempos de ejecución aumentaron de manera considerable con el incremento en el número de objetivos. Esto limita su aplicación y justifica que en nuestra experimentación sólo se consideró para problemas de optimización con hasta 20 objetivos.

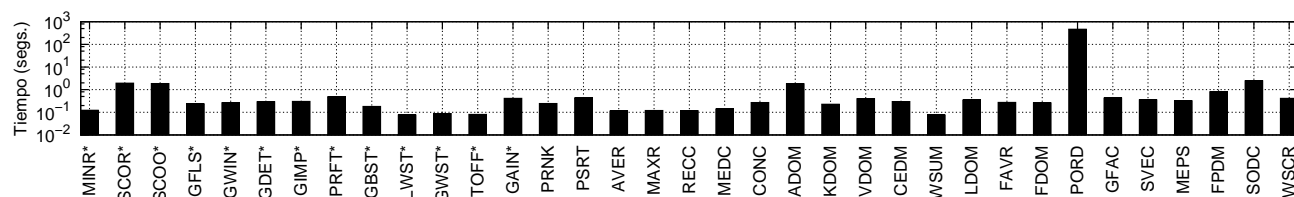
Por otra parte, nuestras propuestas **SCOR\*** y **SCOO\***, al igual que el método **ADOM** tomado de la literatura, mostraron un incremento importante en su tiempo de cómputo cuando se aumentó la cantidad de objetivos. Sin embargo, consideramos que las demandas computacionales de estos enfoques son aún razonables, por lo que no resultó prohibitiva su aplicación en los diferentes experimentos realizados.



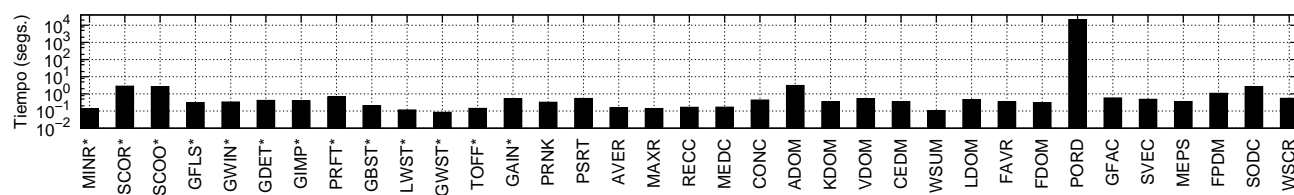
(a) 5 objetivos.



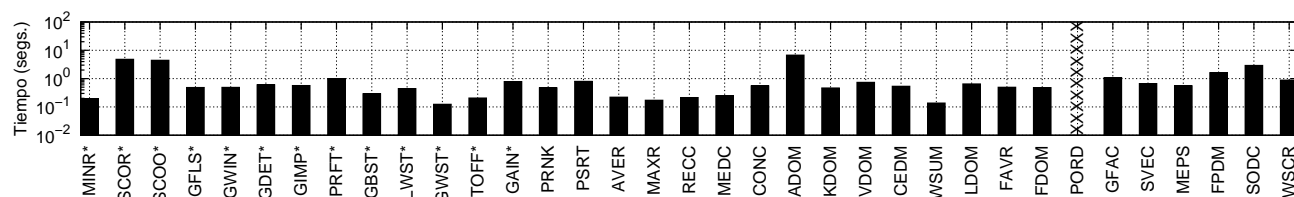
(b) 10 objetivos.



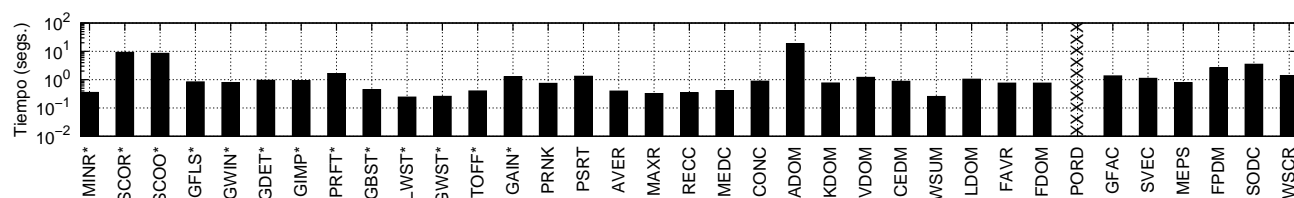
(c) 15 objetivos.



(d) 20 objetivos.



(e) 30 objetivos.



(f) 50 objetivos.

Figura 4.2: Tiempo de cómputo al jerarquizar 100 veces una población de 100 soluciones.

## 4.4 Incorporación de los métodos en un algoritmo base

Los distintos métodos de asignación de aptitud estudiados fueron incorporados en un MOEA con la finalidad de investigar su habilidad para guiar el proceso de búsqueda y su escalabilidad con respecto al número de objetivos. El flujo de ejecución del MOEA implementado se describe mediante la figura 4.3. En un inicio se genera aleatoriamente una población de  $N$  padres. Esta

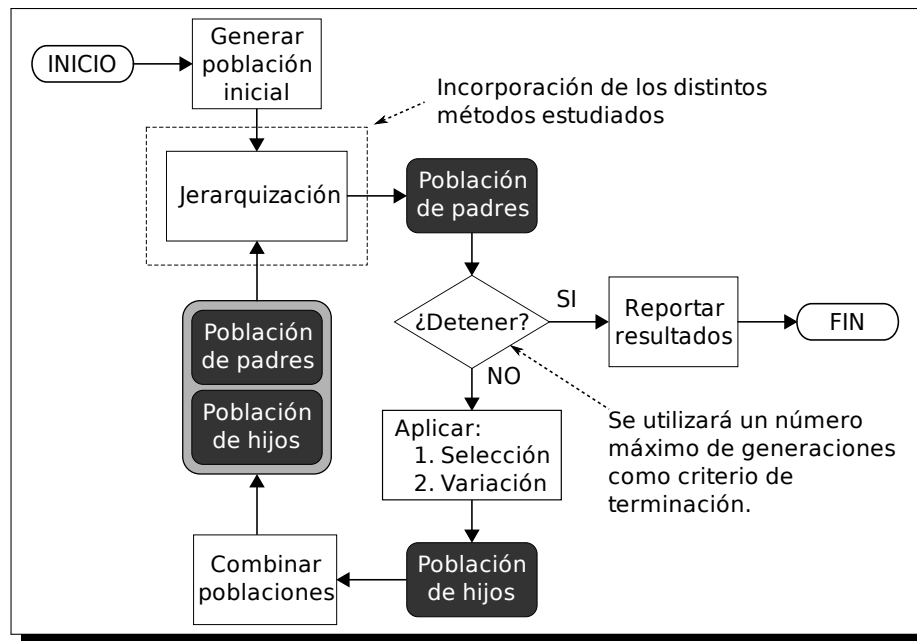


Figura 4.3: Flujo de ejecución del MOEA implementado.

población se jerarquiza (de acuerdo con el método incorporado) para que a través del proceso de selección se determine qué individuos podrán reproducirse. Sobre el conjunto de padres seleccionados se aplican los operadores de variación para generar una población de  $N$  hijos. Una vez aplicados los operadores de variación, la población de padres y la población de hijos se combinan, dando lugar a una población de tamaño  $2N$ . Posteriormente, esta población combinada se jerarquiza para que los mejores  $N$  individuos sobrevivan y formen la población de padres para la siguiente generación (MOEA elitista [18]). Iterativamente se seleccionarán los individuos con la mejor jerarquía, y si durante esta selección se da el caso en que existen más individuos compartiendo la mejor jerarquía que la capacidad disponible, entonces los elementos requeridos serán elegidos de manera aleatoria.



El MOEA implementado utiliza selección mediante torneo binario (binary tournament selection) basada en la jerarquía (aptitud) de las soluciones. Como operadores de variación, este MOEA incorpora la cruce binaria simulada (simulated binary crossover, SBX) ( $\eta_c = 15$ ) [21] y mutación polinomial (polynomial mutation) ( $\eta_m = 20$ ) [19], con probabilidades de 1.0 y  $1/n$  respectivamente, donde  $n$  es el número de variables de decisión. Para todos los experimentos reportados en esta sección se utilizó una población de  $N = 100$  individuos y 300 generaciones para la búsqueda. Con el propósito de evitar alteraciones en el comportamiento de los métodos estudiados, el MOEA implementado no incorpora ningún mecanismo adicional para preservar diversidad en la población.

El desempeño de algunos enfoques depende fuertemente del rango en que se encuentren los valores de los objetivos, de tal forma, todos los métodos (sin excepción) fueron aplicados sobre los objetivos normalizados mediante la ecuación (4.2):

$$f'_m(\mathbf{X}_i) = \frac{f_m(\mathbf{X}_i) - GMIN_m}{GMAX_m - GMIN_m} \quad \forall m \in \{1, 2, \dots, M\} \quad (4.2)$$

donde  $GMAX_m$  y  $GMIN_m$  denotan respectivamente los valores máximo y mínimo conocidos (desde el inicio de la búsqueda) para el  $m$ -ésimo objetivo. Dado que para los problemas de prueba adoptados se conoce a priori que  $GMIN_m = 0$ , la ecuación (4.2) se reduce a:

$$f'_m(\mathbf{X}_i) = \frac{f_m(\mathbf{X}_i)}{GMAX_m} \quad \forall m \in \{1, 2, \dots, M\} \quad (4.3)$$

#### 4.4.1 Habilidad de convergencia

Para este experimento se utilizó la métrica de convergencia tal como fue descrita en la Sección 2.6.1. Las tablas 4.1 a 4.6 muestran los resultados obtenidos por el MOEA al utilizar cada uno de los enfoques de jerarquización estudiados para optimizar los problemas de prueba de nuestro interés. Los datos en estas tablas corresponden a la media y desviación estándar de la métrica de convergencia aplicada sobre la población final del MOEA para 31 ejecuciones independientes de cada experimento. Los mejores valores logrados para cada instancia de los problemas han sido resaltados con negritas para facilitar la interpretación de estas tablas.

Tabla 4.1: Media y desviación estándar de la métrica de convergencia. Problema DTLZ1.

Método	5 Obj.	10 Obj.	15 Obj.	20 Obj.	30 Obj.	50 Obj.
MINR*	0.016 ± 0.079	0.140 ± 0.306	0.086 ± 0.090	0.226 ± 0.313	0.294 ± 0.333	0.570 ± 0.400
SCOR*	0.017 ± 0.056	0.131 ± 0.269	0.213 ± 0.322	0.302 ± 0.467	0.246 ± 0.205	0.687 ± 0.610
SCOO*	0.000 ± 0.001	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	0.001 ± 0.001	0.017 ± 0.039
GFLS*	<b>0.000 ± 0.000</b>	0.011 ± 0.039	0.005 ± 0.023	0.040 ± 0.092	0.111 ± 0.181	0.390 ± 0.476
GWIN*	<b>0.000 ± 0.000</b>	0.016 ± 0.047	0.009 ± 0.032	0.015 ± 0.047	0.105 ± 0.110	0.341 ± 0.228
GDET*	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	0.006 ± 0.023	0.018 ± 0.030
GIMP*	0.000 ± 0.001	0.000 ± 0.001	0.001 ± 0.001	0.001 ± 0.001	0.001 ± 0.001	0.037 ± 0.047
PRFT*	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	0.000 ± 0.001	0.006 ± 0.022	<b>0.011 ± 0.026</b>
GBST*	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	0.030 ± 0.043
LWST*	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	0.004 ± 0.023	<b>0.000 ± 0.000</b>	0.003 ± 0.016	0.023 ± 0.038
GWST*	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	0.003 ± 0.016	0.030 ± 0.056
TOFF*	0.001 ± 0.001	0.002 ± 0.001	<b>0.000 ± 0.000</b>	0.004 ± 0.020	0.007 ± 0.023	0.012 ± 0.032
GAIN*	<b>0.000 ± 0.000</b>	0.006 ± 0.028	0.010 ± 0.032	0.055 ± 0.133	0.231 ± 0.357	0.571 ± 0.635
PRNK	0.644 ± 0.502	3.808 ± 3.101	1.976 ± 1.669	2.042 ± 1.492	2.256 ± 1.632	1.612 ± 0.846
PSRT	0.637 ± 0.555	3.313 ± 2.545	2.357 ± 1.489	2.147 ± 1.160	1.130 ± 0.789	1.458 ± 0.925
AVER	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	0.005 ± 0.023	0.019 ± 0.041	0.107 ± 0.117	0.353 ± 0.259
MAXR	37.85 ± 14.64	27.10 ± 7.857	22.81 ± 6.854	21.91 ± 7.283	14.28 ± 7.615	10.77 ± 5.471
RECC	<b>0.000 ± 0.000</b>	0.021 ± 0.066	0.031 ± 0.140	0.044 ± 0.068	0.108 ± 0.134	0.347 ± 0.333
MEDC	37.80 ± 12.28	30.02 ± 10.80	24.98 ± 9.257	19.42 ± 6.324	19.53 ± 7.093	12.57 ± 5.016
CONC	<b>0.000 ± 0.000</b>	0.006 ± 0.028	0.001 ± 0.001	0.056 ± 0.102	0.089 ± 0.092	0.443 ± 0.419
ADOM	0.002 ± 0.001	0.001 ± 0.001	0.002 ± 0.002	0.001 ± 0.002	0.028 ± 0.053	0.040 ± 0.047
KDOM	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	0.021 ± 0.048	0.033 ± 0.058	0.130 ± 0.107	0.477 ± 0.329
VDOM	0.435 ± 0.277	0.463 ± 0.397	0.424 ± 0.405	0.499 ± 0.432	0.416 ± 0.280	0.393 ± 0.338
CEDM	0.002 ± 0.001	0.002 ± 0.001	0.002 ± 0.001	0.005 ± 0.020	0.002 ± 0.002	0.029 ± 0.035
WSUM	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	0.004 ± 0.023	<b>0.000 ± 0.000</b>	0.003 ± 0.016	0.014 ± 0.052
LDOM	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	0.006 ± 0.022	0.023 ± 0.041
FAVR	0.001 ± 0.001	0.016 ± 0.047	0.102 ± 0.145	0.198 ± 0.256	0.634 ± 0.710	2.543 ± 2.713
FDOM	0.001 ± 0.001	0.011 ± 0.039	0.032 ± 0.103	0.033 ± 0.051	0.143 ± 0.199	0.609 ± 0.529
PORD	0.496 ± 0.331	0.824 ± 0.731	0.841 ± 0.755	0.657 ± 0.628	-	-
GFAC	33.22 ± 10.83	16.47 ± 11.83	19.61 ± 8.689	17.90 ± 6.820	15.51 ± 6.519	10.89 ± 4.512
SVEC	0.586 ± 0.516	0.753 ± 0.942	0.472 ± 0.597	0.690 ± 0.665	0.576 ± 0.472	0.377 ± 0.335
MEPS	0.004 ± 0.004	0.019 ± 0.048	0.021 ± 0.043	0.047 ± 0.076	0.122 ± 0.188	0.071 ± 0.072
FPDM	0.752 ± 0.467	0.936 ± 0.754	1.020 ± 0.660	1.208 ± 0.894	1.135 ± 0.662	0.877 ± 0.750
SODC	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	0.006 ± 0.024	0.012 ± 0.033	0.033 ± 0.064	0.083 ± 0.090
WSCR	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	0.009 ± 0.032	0.044 ± 0.061	0.114 ± 0.180	0.372 ± 0.241

A partir de la tabla 4.1 es posible apreciar que la mayoría de los enfoques propuestos, al igual que algunos de los métodos del estado del arte, lograron una buena convergencia para las diferentes instancias del problema DTLZ1. Se puede decir que **GBST\*** es el método que mostró el mejor desempeño para este problema, ya que obtuvo el mejor resultado en 5 de 6 instancias. Sin embargo, algunas otras propuestas como **GDET\***, **PRFT\***, **GWST\***, **LWST\*** y **SCOO\***, y métodos tomados de la literatura como **LDOM** y **WSUM**, lograron también los mejores resultados en parte de las instancias y cercanía al mejor resultado en el resto de los casos. Por otra parte, aunque los métodos **TOFF\***, **GIMP\*** y **CEDM** no alcanzaron la mejor convergencia en las diferentes instancias del experimento, sus resultados fueron muy cercanos a los mejores y su desviación estándar nos indica que posiblemente algunas malas corridas influyeron en su resultado.

El peor desempeño mostrado para DTLZ1 fue el de los métodos MEDC, MAXR y GFAC, ya que sus resultados fueron aún peores que los obtenidos por los enfoques basados en la dominancia de Pareto. Otros métodos que no lograron una buena convergencia para este problema fueron FPDm, PRNK, PSRT, SVEC, PORD y VDOM, cuyo desempeño fue pobre incluso para las instancias pequeñas.

Tabla 4.2: Media y desviación estándar de la métrica de convergencia. Problema DTLZ2.

Método	5 Obj.	10 Obj.	15 Obj.	20 Obj.	30 Obj.	50 Obj.
MINR*	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
SCOR*	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
SCOO*	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
GFLS*	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.018 ± 0.033
GWIN*	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.008 ± 0.010
GDET*	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
GIMP*	0.000 ± 0.001	0.001 ± 0.001	0.000 ± 0.001	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
PRFT*	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
GBST*	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
LWST*	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
GWST*	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
TOFF*	0.000 ± 0.000	0.000 ± 0.002	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
GAIN*	0.000 ± 0.000	0.000 ± 0.000	0.011 ± 0.010	0.052 ± 0.038	0.263 ± 0.148	0.482 ± 0.183
PRNK	0.009 ± 0.004	0.033 ± 0.020	0.064 ± 0.047	0.079 ± 0.047	0.074 ± 0.042	0.099 ± 0.090
PSRT	0.009 ± 0.003	0.033 ± 0.016	0.059 ± 0.039	0.066 ± 0.066	0.064 ± 0.034	0.086 ± 0.055
AVER	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.001	0.015 ± 0.036
MAXR	0.140 ± 0.100	0.065 ± 0.083	0.041 ± 0.098	0.011 ± 0.024	0.003 ± 0.004	0.004 ± 0.007
RECC	0.002 ± 0.004	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.002 ± 0.002	0.132 ± 0.113
MEDC	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.004 ± 0.005	0.280 ± 0.200
CONC	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.001 ± 0.003	0.069 ± 0.068
ADOM	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
KDOM	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.001 ± 0.001	0.078 ± 0.089
VDOM	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
CEDM	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
WSUM	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
LDOM	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
FAVR	0.001 ± 0.004	0.002 ± 0.010	0.189 ± 0.103	0.185 ± 0.136	0.400 ± 0.176	0.532 ± 0.203
FDOM	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.001 ± 0.001	0.053 ± 0.047
PORD	0.002 ± 0.001	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	-	-
GFAC	0.122 ± 0.075	0.133 ± 0.102	0.131 ± 0.058	0.185 ± 0.078	0.289 ± 0.143	0.392 ± 0.199
SVEC	0.002 ± 0.001	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
MEPS	0.007 ± 0.002	0.012 ± 0.005	0.014 ± 0.006	0.018 ± 0.006	0.018 ± 0.009	0.016 ± 0.008
FPDM	0.033 ± 0.011	0.096 ± 0.040	0.124 ± 0.048	0.145 ± 0.071	0.137 ± 0.078	0.199 ± 0.096
SODC	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
WSCR	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.017 ± 0.023

De acuerdo con la tabla 4.2, la mayoría de los enfoques no encontraron dificultades serias para converger en las diferentes instancias del problema DTLZ2. Estos resultados dificultan resaltar qué enfoques fueron los mejores; sin embargo, es posible destacar cuáles mostraron un mal desempeño. GFAC y FPDm fueron los métodos que obtuvieron los peores resultados para este experimento, ya que para todas las instancias lograron una convergencia pobre y se deterioraron con el aumento

en la cantidad de objetivos. Por otra parte, **FAVR** y **GAIN\*** obtuvieron buenos resultados para las instancias pequeñas, pero el incremento en el número de objetivos provocó que sus resultados para el resto de las instancias se deterioraran significativamente.

Tabla 4.3: Media y desviación estándar de la métrica de convergencia. Problema DTLZ3.

Método	5 Obj.	10 Obj.	15 Obj.	20 Obj.	30 Obj.	50 Obj.
<b>MINR*</b>	5.716 ± 5.351	23.26 ± 15.82	27.55 ± 14.02	33.54 ± 17.84	42.65 ± 16.48	48.79 ± 17.00
<b>SCOR*</b>	7.239 ± 4.691	19.74 ± 9.966	29.11 ± 13.81	38.78 ± 17.76	38.24 ± 17.37	45.90 ± 20.77
<b>SCOO*</b>	0.209 ± 0.470	<b>0.024 ± 0.036</b>	0.208 ± 0.396	0.311 ± 0.450	0.601 ± 0.710	1.965 ± 2.246
<b>GFLS*</b>	0.855 ± 1.093	3.853 ± 4.181	7.914 ± 4.512	18.84 ± 10.31	42.06 ± 14.81	70.64 ± 26.11
<b>GWIN*</b>	0.772 ± 0.649	3.309 ± 2.556	8.821 ± 5.631	21.96 ± 10.99	48.33 ± 17.64	65.49 ± 29.49
<b>GDET*</b>	0.171 ± 0.369	0.205 ± 0.471	0.309 ± 0.535	0.336 ± 0.530	0.695 ± 0.686	1.695 ± 1.206
<b>GIMP*</b>	0.386 ± 0.707	0.287 ± 0.488	0.416 ± 0.552	0.770 ± 1.201	1.401 ± 1.071	2.490 ± 2.342
<b>PRFT*</b>	0.094 ± 0.261	0.043 ± 0.176	0.173 ± 0.371	0.236 ± 0.489	<b>0.340 ± 0.529</b>	1.388 ± 1.086
<b>GBST*</b>	0.110 ± 0.295	0.041 ± 0.176	0.098 ± 0.265	<b>0.151 ± 0.334</b>	0.693 ± 0.891	1.419 ± 1.300
<b>LBST*</b>	0.108 ± 0.294	0.142 ± 0.336	0.174 ± 0.365	0.304 ± 0.521	0.792 ± 1.035	<b>1.319 ± 1.706</b>
<b>GWST*</b>	0.074 ± 0.247	0.076 ± 0.354	0.195 ± 0.365	0.303 ± 0.923	0.632 ± 0.701	1.802 ± 2.238
<b>TOFF*</b>	0.820 ± 0.808	0.159 ± 0.426	0.173 ± 0.444	0.176 ± 0.365	0.703 ± 0.993	2.158 ± 1.847
<b>GAIN*</b>	2.145 ± 1.820	13.23 ± 9.236	40.30 ± 21.15	61.33 ± 24.59	89.34 ± 29.35	93.35 ± 36.98
<b>PRNK</b>	61.38 ± 21.65	227.2 ± 75.90	218.2 ± 63.13	195.0 ± 73.63	186.5 ± 62.86	148.4 ± 76.61
<b>PSRT</b>	59.71 ± 19.96	260.6 ± 82.82	244.3 ± 103.2	204.9 ± 86.09	191.7 ± 103.2	164.9 ± 81.59
<b>AVER</b>	0.920 ± 1.199	3.470 ± 2.344	9.891 ± 6.485	18.21 ± 8.368	50.08 ± 15.78	61.35 ± 22.84
<b>MAXR</b>	496.1 ± 145.3	290.8 ± 243.5	175.2 ± 189.4	88.15 ± 111.0	45.24 ± 41.57	38.24 ± 11.55
<b>RECC</b>	39.65 ± 20.71	17.96 ± 9.694	25.01 ± 10.72	39.66 ± 19.06	45.57 ± 16.98	87.05 ± 35.02
<b>MEDC</b>	0.890 ± 0.880	2.881 ± 3.072	6.861 ± 4.084	16.98 ± 9.435	56.03 ± 26.50	121.0 ± 69.04
<b>CONC</b>	0.571 ± 0.883	2.424 ± 1.859	5.255 ± 3.219	15.76 ± 8.973	47.19 ± 14.72	69.87 ± 29.17
<b>ADOM</b>	0.051 ± 0.177	0.150 ± 0.333	<b>0.074 ± 0.308</b>	0.407 ± 0.553	0.643 ± 0.905	1.546 ± 1.340
<b>KDOM</b>	0.740 ± 0.928	2.174 ± 2.166	5.721 ± 3.481	14.75 ± 7.477	46.36 ± 13.60	76.43 ± 19.59
<b>VDOM</b>	2.877 ± 4.292	3.820 ± 4.183	4.054 ± 3.883	3.243 ± 2.570	7.173 ± 5.991	8.942 ± 7.351
<b>CEDM</b>	<b>0.009 ± 0.007</b>	0.107 ± 0.296	0.174 ± 0.369	0.268 ± 0.506	0.801 ± 0.868	1.509 ± 1.241
<b>WSUM</b>	0.078 ± 0.245	0.140 ± 0.333	0.144 ± 0.332	0.239 ± 0.489	0.823 ± 1.277	1.507 ± 1.010
<b>LDOM</b>	0.141 ± 0.343	0.438 ± 0.706	0.592 ± 1.359	0.886 ± 0.863	1.276 ± 1.159	2.683 ± 2.219
<b>FAVR</b>	14.11 ± 34.73	25.27 ± 21.41	179.9 ± 85.30	179.6 ± 88.61	328.3 ± 129.2	416.3 ± 133.7
<b>FDOM</b>	0.767 ± 0.874	3.310 ± 2.398	5.101 ± 2.448	14.87 ± 7.534	50.09 ± 21.36	67.11 ± 21.54
<b>PORD</b>	35.77 ± 15.75	33.46 ± 18.47	24.74 ± 10.84	22.50 ± 9.913	-	-
<b>GFAC</b>	543.0 ± 96.84	339.1 ± 203.6	296.4 ± 126.7	242.4 ± 188.2	178.0 ± 127.9	238.9 ± 117.8
<b>SVEC</b>	33.30 ± 11.19	36.83 ± 17.45	31.64 ± 11.50	20.40 ± 12.52	17.03 ± 10.34	20.74 ± 9.260
<b>MEPS</b>	0.461 ± 0.600	2.509 ± 1.631	4.351 ± 3.308	6.427 ± 7.118	7.871 ± 5.271	7.824 ± 4.456
<b>FPDM</b>	41.81 ± 15.55	78.75 ± 30.34	121.7 ± 49.37	118.0 ± 39.84	97.95 ± 43.65	89.22 ± 39.42
<b>SODC</b>	4.482 ± 3.724	14.38 ± 9.012	19.44 ± 8.422	17.69 ± 9.147	15.29 ± 6.637	15.30 ± 7.649
<b>WSCR</b>	0.596 ± 0.750	3.757 ± 2.732	7.315 ± 4.961	18.95 ± 8.948	48.91 ± 16.69	67.09 ± 27.52

Las instancias del problema DTLZ3 (tabla 4.3) fueron difíciles de optimizar para los diferentes enfoques; sin embargo, es posible puntualizar qué métodos lograron el mejor desempeño. Aún cuando para cada instancia el mejor resultado fue obtenido por un método diferente, nosotros consideramos que **PRFT\*** y **GBST\*** son las mejores alternativas para este caso de prueba, ya que sus resultados fueron próximos a los mejores y esta característica se preservó para las diferentes instancias del problema. **GFAC** obtuvo los peores resultados para las primeras 4 instancias, mientras que **FAVR** fue el método con el peor desempeño para las últimas 2. Algunos otros métodos que presentaron

consistentemente un mal desempeño son MAXR, FPDM, RECC, nuestros métodos MINR\*, SCOR\* y GAIN\*, y nuevamente los enfoques basados en la dominancia de Pareto. Un comportamiento que vale la pena discutir es el de MAXR, que comenzó con los peores resultados pero mejoró en instancias más grandes del problema (característica presentada también para las funciones DTLZ1 y DTLZ2). En nuestra opinión, este comportamiento se debe a que cuando se incrementa la dimensionalidad del espacio de objetivos, también aumentan las posibilidades de encontrar soluciones que puedan superar malos ejemplares retenidos por el método a lo largo de la búsqueda (mientras más objetivos tenga el problema, más extremos compiten con la jerarquía más alta).

Tabla 4.4: Media y desviación estándar de la métrica de convergencia. Problema DTLZ4.

Método	5 Obj.	10 Obj.	15 Obj.	20 Obj.	30 Obj.	50 Obj.
MINR*	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.001 ± 0.001
SCOR*	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.001 ± 0.000
SCOO*	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
GFLS*	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
GWIN*	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
GDET*	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
GIMP*	0.000 ± 0.000	0.003 ± 0.005	0.004 ± 0.004	0.005 ± 0.005	0.009 ± 0.006	0.013 ± 0.007
PRFT*	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
GBST*	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
LWST*	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
GWST*	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
TOFF*	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
GAIN*	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
PRNK	0.004 ± 0.002	0.064 ± 0.044	0.346 ± 0.139	0.475 ± 0.113	0.517 ± 0.171	0.548 ± 0.200
PSRT	0.004 ± 0.002	0.066 ± 0.043	0.346 ± 0.139	0.475 ± 0.113	0.517 ± 0.171	0.548 ± 0.200
AVER	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
MAXR	0.476 ± 0.165	0.499 ± 0.166	0.490 ± 0.157	0.481 ± 0.180	0.499 ± 0.172	0.549 ± 0.165
RECC	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.031 ± 0.056
MEDC	0.225 ± 0.124	0.196 ± 0.108	0.165 ± 0.077	0.136 ± 0.089	0.092 ± 0.053	0.018 ± 0.026
CONC	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
ADOM	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
KDOM	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
VDOM	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.001 ± 0.003	0.017 ± 0.017
CEDM	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
WSUM	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
LDOM	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
FAVR	0.272 ± 0.133	0.184 ± 0.092	0.507 ± 0.202	0.342 ± 0.140	0.409 ± 0.164	0.432 ± 0.192
FDOM	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
PORD	0.002 ± 0.001	0.001 ± 0.001	0.001 ± 0.001	0.000 ± 0.000	-	-
GFAC	0.459 ± 0.162	0.452 ± 0.151	0.431 ± 0.158	0.484 ± 0.141	0.504 ± 0.181	0.565 ± 0.170
SVEC	0.008 ± 0.019	0.003 ± 0.003	0.002 ± 0.002	0.002 ± 0.002	0.000 ± 0.000	0.000 ± 0.000
MEPS	0.005 ± 0.002	0.008 ± 0.003	0.010 ± 0.004	0.011 ± 0.005	0.013 ± 0.005	0.019 ± 0.008
FPDM	0.015 ± 0.009	0.000 ± 0.001	0.001 ± 0.004	0.000 ± 0.000	0.516 ± 0.158	0.525 ± 0.211
SODC	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
WSCR	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000

La tabla 4.4 muestra que la mayoría de los enfoques fueron capaces de optimizar las diferentes instancias del problema DTLZ4. El peor desempeño se observó al utilizar principalmente los métodos

GFAC, MAXR, FAVR, MEDC, PRNK y PSRT.

Para el problema DTLZ5 (tabla 4.5) podemos apreciar resultados muy polarizados. Por un lado, nuestros métodos **GDET\***, **GIMP\***, **PRFT\***, **GBST\***, **LWST\*** y **GWST\***, al igual que los métodos **VDOM**, **CEDM** y **WSUM**, obtuvieron la mejor convergencia en todas las instancias de este problema, mientras que el desempeño del resto de los enfoques, con algunas excepciones, quedó lejos de ser aceptable en la mayoría de los casos.

Tabla 4.5: Media y desviación estándar de la métrica de convergencia. Problema DTLZ5.

Método	5 Obj.	10 Obj.	15 Obj.	20 Obj.	30 Obj.	50 Obj.
<b>MINR*</b>	0.001 ± 0.001	0.006 ± 0.004	0.008 ± 0.004	0.006 ± 0.003	0.010 ± 0.012	0.104 ± 0.077
<b>SCOR*</b>	0.002 ± 0.002	0.006 ± 0.002	0.007 ± 0.003	0.006 ± 0.002	0.016 ± 0.023	0.122 ± 0.066
<b>SCOO*</b>	0.003 ± 0.009	0.004 ± 0.009	0.007 ± 0.018	0.001 ± 0.004	0.002 ± 0.005	0.005 ± 0.022
<b>GFLS*</b>	<b>0.000 ± 0.000</b>	2.499 ± 0.003	2.500 ± 0.000	2.500 ± 0.000	2.500 ± 0.000	2.500 ± 0.000
<b>GWIN*</b>	<b>0.000 ± 0.000</b>	2.499 ± 0.003	2.500 ± 0.000	2.500 ± 0.000	2.500 ± 0.000	2.500 ± 0.000
<b>GDET*</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>
<b>GIMP*</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>
<b>PRFT*</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>
<b>GBST*</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>
<b>LWST*</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>
<b>GWST*</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>
<b>TOFF*</b>	0.435 ± 0.212	0.374 ± 0.118	0.199 ± 0.186	0.007 ± 0.039	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>
<b>GAIN*</b>	<b>0.000 ± 0.000</b>	2.500 ± 0.000	2.500 ± 0.000	2.500 ± 0.000	2.500 ± 0.000	2.499 ± 0.002
<b>PRNK</b>	0.133 ± 0.059	0.226 ± 0.084	0.307 ± 0.129	0.264 ± 0.118	0.325 ± 0.139	0.363 ± 0.142
<b>PSRT</b>	0.157 ± 0.078	0.270 ± 0.114	0.253 ± 0.105	0.298 ± 0.123	0.321 ± 0.136	0.374 ± 0.133
<b>AVER</b>	<b>0.000 ± 0.000</b>	2.499 ± 0.003	2.500 ± 0.000	2.500 ± 0.000	2.500 ± 0.000	2.500 ± 0.000
<b>MAXR</b>	0.172 ± 0.176	0.346 ± 0.361	0.400 ± 0.441	0.410 ± 0.347	0.512 ± 0.263	0.764 ± 0.321
<b>RECC</b>	<b>0.000 ± 0.000</b>	0.009 ± 0.017	0.260 ± 0.202	0.508 ± 0.336	1.072 ± 0.419	1.480 ± 0.252
<b>MEDC</b>	<b>0.000 ± 0.000</b>	2.498 ± 0.003	2.500 ± 0.000	2.500 ± 0.000	2.500 ± 0.000	2.500 ± 0.000
<b>CONC</b>	<b>0.000 ± 0.000</b>	2.500 ± 0.001	2.500 ± 0.000	2.500 ± 0.000	2.500 ± 0.000	2.500 ± 0.000
<b>ADOM</b>	0.044 ± 0.040	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>
<b>KDOM</b>	<b>0.000 ± 0.000</b>	2.495 ± 0.026	2.500 ± 0.000	2.500 ± 0.000	2.500 ± 0.000	2.500 ± 0.000
<b>VDOM</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>
<b>CEDM</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>
<b>WSUM</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>
<b>LDOM</b>	<b>0.000 ± 0.000</b>	0.036 ± 0.016	0.016 ± 0.011	0.012 ± 0.007	0.008 ± 0.005	0.007 ± 0.005
<b>FAVR</b>	<b>0.000 ± 0.000</b>	2.207 ± 0.102	2.181 ± 0.078	2.318 ± 0.045	2.356 ± 0.032	2.401 ± 0.022
<b>FDOM</b>	<b>0.000 ± 0.000</b>	2.499 ± 0.003	2.500 ± 0.000	2.500 ± 0.000	2.500 ± 0.000	2.500 ± 0.000
<b>PORD</b>	0.001 ± 0.002	2.232 ± 0.417	2.443 ± 0.166	2.498 ± 0.003	-	-
<b>GFAC</b>	0.135 ± 0.162	1.827 ± 0.625	1.890 ± 0.526	1.822 ± 0.490	2.074 ± 0.428	2.245 ± 0.282
<b>SVEC</b>	0.001 ± 0.001	2.274 ± 0.428	2.475 ± 0.042	2.498 ± 0.004	2.500 ± 0.001	2.500 ± 0.000
<b>MEPS</b>	0.062 ± 0.022	0.128 ± 0.042	0.164 ± 0.068	0.201 ± 0.103	0.190 ± 0.095	0.173 ± 0.056
<b>FPDM</b>	0.078 ± 0.036	2.071 ± 0.100	2.294 ± 0.044	2.360 ± 0.034	2.416 ± 0.038	2.473 ± 0.022
<b>SODC</b>	<b>0.000 ± 0.000</b>	<b>0.000 ± 0.000</b>	0.003 ± 0.001	0.009 ± 0.003	0.014 ± 0.010	0.021 ± 0.023
<b>WSCR</b>	<b>0.000 ± 0.000</b>	2.499 ± 0.002	2.500 ± 0.000	2.500 ± 0.000	2.500 ± 0.000	2.500 ± 0.000

Finalmente, los resultados en la tabla 4.6 indican que el FP del problema DTLZ6 fue difícil de aproximar. El mejor resultado en 5 de 6 instancias fue logrado por nuestras propuestas. De acuerdo con estos resultados, consideramos que los métodos **PRFT\*** y **CEDM** fueron los mejores para este experimento; sin embargo, **GBST\***, **GWST\***, **WSUM**, **GDET\*** y **LWST\*** presentaron un

Tabla 4.6: Media y desviación estándar de la métrica de convergencia. Problema DTLZ6.

Método	5 Obj.	10 Obj.	15 Obj.	20 Obj.	30 Obj.	50 Obj.
MINR*	2.887 ± 0.556	5.041 ± 1.023	6.231 ± 1.441	6.385 ± 1.118	6.855 ± 0.898	7.343 ± 0.726
SCOR*	2.891 ± 0.524	4.961 ± 1.039	5.990 ± 1.150	6.491 ± 1.050	6.861 ± 0.895	7.375 ± 0.635
SCOO*	0.085 ± 0.023	0.087 ± 0.032	0.087 ± 0.028	0.092 ± 0.022	0.101 ± 0.026	0.180 ± 0.050
GFLS*	0.075 ± 0.022	10.00 ± 0.000	10.00 ± 0.000	10.00 ± 0.000	10.00 ± 0.000	10.00 ± 0.000
GWIN*	0.070 ± 0.030	10.00 ± 0.003	10.00 ± 0.000	10.00 ± 0.000	10.00 ± 0.000	10.00 ± 0.000
GDET*	0.085 ± 0.027	0.082 ± 0.028	0.089 ± 0.027	0.085 ± 0.029	<b>0.081 ± 0.033</b>	0.096 ± 0.027
GIMP*	0.167 ± 0.193	0.176 ± 0.319	0.157 ± 0.169	0.121 ± 0.151	0.232 ± 0.313	0.253 ± 0.454
PRFT*	0.080 ± 0.029	<b>0.076 ± 0.030</b>	0.079 ± 0.029	0.081 ± 0.033	0.089 ± 0.036	<b>0.089 ± 0.033</b>
GBST*	0.086 ± 0.035	0.085 ± 0.027	0.084 ± 0.034	<b>0.078 ± 0.028</b>	0.088 ± 0.034	0.090 ± 0.022
LWST*	0.076 ± 0.020	0.086 ± 0.030	0.083 ± 0.031	0.095 ± 0.031	0.089 ± 0.026	0.098 ± 0.031
GWST*	<b>0.068 ± 0.030</b>	0.077 ± 0.028	0.082 ± 0.035	0.089 ± 0.031	0.092 ± 0.032	0.105 ± 0.031
TOFF*	8.993 ± 0.515	4.445 ± 3.912	0.089 ± 0.031	0.100 ± 0.033	0.091 ± 0.036	0.093 ± 0.039
GAIN*	10.00 ± 0.000	10.00 ± 0.000	10.00 ± 0.000	10.00 ± 0.000	9.999 ± 0.004	9.996 ± 0.004
PRNK	5.782 ± 0.434	7.978 ± 0.439	8.085 ± 0.450	8.170 ± 0.471	8.386 ± 0.548	8.500 ± 0.393
PSRT	5.787 ± 0.503	8.011 ± 0.404	8.036 ± 0.472	8.154 ± 0.483	8.466 ± 0.410	8.529 ± 0.323
AVER	0.081 ± 0.032	10.00 ± 0.000	10.00 ± 0.000	10.00 ± 0.000	10.00 ± 0.000	10.00 ± 0.000
MAXR	6.505 ± 2.207	8.481 ± 0.859	8.494 ± 0.837	8.558 ± 0.832	8.289 ± 0.745	8.074 ± 0.586
RECC	5.087 ± 1.099	7.242 ± 0.641	7.990 ± 0.632	7.984 ± 0.745	8.762 ± 0.606	9.332 ± 0.588
MEDC	0.146 ± 0.056	9.999 ± 0.004	10.00 ± 0.000	10.00 ± 0.000	10.00 ± 0.000	10.00 ± 0.000
CONC	0.073 ± 0.024	10.00 ± 0.000	10.00 ± 0.000	10.00 ± 0.000	10.00 ± 0.000	10.00 ± 0.000
ADOM	0.069 ± 0.023	0.460 ± 0.086	0.546 ± 0.070	0.560 ± 0.069	0.596 ± 0.088	0.611 ± 0.077
KDOM	0.077 ± 0.027	10.00 ± 0.001	10.00 ± 0.001	10.00 ± 0.000	10.00 ± 0.000	10.00 ± 0.000
VDOM	0.071 ± 0.027	0.088 ± 0.039	0.128 ± 0.068	0.130 ± 0.063	0.295 ± 0.236	0.503 ± 0.350
CEDM	0.072 ± 0.027	0.076 ± 0.036	<b>0.075 ± 0.031</b>	0.084 ± 0.033	0.096 ± 0.028	0.101 ± 0.036
WSUM	0.080 ± 0.030	0.081 ± 0.025	0.089 ± 0.029	0.085 ± 0.031	0.087 ± 0.034	0.091 ± 0.029
LDOM	0.072 ± 0.028	6.712 ± 0.303	6.388 ± 0.376	6.193 ± 0.330	5.981 ± 0.437	5.428 ± 0.297
FAVR	5.347 ± 2.238	9.821 ± 0.195	9.380 ± 0.245	9.821 ± 0.084	9.693 ± 0.108	9.573 ± 0.137
FDOM	0.090 ± 0.025	10.00 ± 0.000	10.00 ± 0.000	10.00 ± 0.000	10.00 ± 0.000	10.00 ± 0.000
PORD	4.152 ± 0.631	9.773 ± 0.154	9.939 ± 0.036	9.976 ± 0.018	-	-
GFAC	7.737 ± 1.249	9.291 ± 0.245	9.629 ± 0.268	9.712 ± 0.235	9.847 ± 0.166	9.896 ± 0.139
SVEC	4.236 ± 0.632	9.764 ± 0.139	9.928 ± 0.055	9.976 ± 0.023	9.995 ± 0.005	9.999 ± 0.001
MEPS	1.259 ± 0.195	5.094 ± 0.375	6.019 ± 0.582	6.377 ± 0.605	6.528 ± 0.678	6.558 ± 0.656
FPDM	3.918 ± 0.409	9.072 ± 0.216	9.545 ± 0.104	9.624 ± 0.100	9.746 ± 0.073	9.839 ± 0.059
SODC	0.290 ± 0.056	0.685 ± 0.157	1.439 ± 0.357	2.215 ± 0.407	3.786 ± 0.505	5.584 ± 0.438
WSCR	0.077 ± 0.035	10.00 ± 0.001	10.00 ± 0.000	10.00 ± 0.000	10.00 ± 0.000	10.00 ± 0.000

comportamiento similar.

Como observaciones generales, es posible destacar que la mayoría de los métodos propuestos fueron capaces de guiar el proceso de búsqueda para las diferentes instancias de este experimento. En nuestra opinión, de manera general podemos resaltar que el mejor desempeño en estos experimentos fue principalmente el de nuestras propuestas **PRFT\***, **GBST\*** y **GDET\***, así como los métodos de la literatura **WSUM** y **CEDM**. Un aspecto destacable de estas alternativas es la robustez que mostraron ante la variedad de problemas de prueba y el incremento en la cantidad de objetivos. Uno de nuestros hallazgos más importantes se relaciona con los resultados obtenidos por **WSUM**, ya que el desempeño de esta simple pero eficiente alternativa fue mejor que el de muchos métodos más elaborados; la convergencia lograda al utilizar **WSUM** fue muy cercana a la mejor aproximación

alcanzada en la mayoría de los casos.

Se puede decir que **GFAC**, **MAXR**, **FAVR**, **MEDC**, **PRNK** y **PSRT** fueron las alternativas que mostraron el peor desempeño global, ya que sus resultados para los diferentes problemas de prueba indican que no son capaces de guiar el proceso de búsqueda en espacios de objetivos con alta dimensionalidad. El desempeño de **PRNK** y **PSRT** confirman que la relación de dominancia de Pareto pierde efectividad al incrementarse la cantidad de objetivos, de manera que con estos resultados podemos respaldar la importancia y necesidad de contar con métodos alternativos para discriminar entre soluciones candidatas para este tipo de problemas.

#### 4.4.2 Distancia generacional invertida

En la Sección 4.4.1 se analizó la capacidad de convergencia que provee cada una de las estrategias de jerarquización estudiadas. Sin embargo, aunque la convergencia debe ser prioridad en cualquier tarea de optimización, un aspecto sin duda importante al optimizar múltiples objetivos es converger de una manera distribuida para proveer al tomador de decisiones una una aproximación suficientemente representativa de la superficie compromiso. Los resultados del MOEA al utilizar los diferentes mecanismos de jerarquización han sido analizados utilizando la distancia generacional invertida (DGI) descrita en la Sección 2.6.2. DGI es una métrica que permite medir convergencia al mismo tiempo que considera la distribución de las soluciones en la superficie compromiso.

Las tablas 4.7 a 4.12 muestran los resultados de la métrica DGI para los problemas DTLZ1 a DTLZ6, respectivamente. Estos resultados corresponden a la media y desviación estándar de 31 corridas independientes para cada instancia del experimento. Los mejores resultados de estas tablas han sido resaltados en negritas.

**MEPS** fue el método que logró los mejores resultados para la métrica DGI al optimizar los problemas DTLZ1, DTLZ2 y DTLZ4 (tablas 4.7, 4.8 y 4.10, respectivamente). Aunque **MEPS** no fue identificado en la Sección 4.4.1 por proveer la mejor capacidad de convergencia para estos tres problemas de prueba, sus resultados no fueron tan despreciables en la mayoría de los casos; v.g., los niveles de convergencia logrados por este método fueron aproximados a los mejores niveles



Tabla 4.7: Media y desviación estándar de la métrica DGI. Problema DTLZ1.

Método	5 Obj.	10 Obj.	15 Obj.	20 Obj.	30 Obj.	50 Obj.
MINR*	0.259 ± 0.063	0.318 ± 0.277	0.235 ± 0.077	0.348 ± 0.297	0.389 ± 0.320	0.632 ± 0.396
SCOR*	0.259 ± 0.041	0.306 ± 0.244	0.356 ± 0.299	0.416 ± 0.447	0.342 ± 0.196	0.746 ± 0.600
SCOO*	0.171 ± 0.001	0.146 ± 0.001	0.127 ± 0.002	0.114 ± 0.001	<b>0.096 ± 0.001</b>	<b>0.081 ± 0.011</b>
GFLS*	0.246 ± 0.004	0.206 ± 0.030	0.171 ± 0.018	0.182 ± 0.081	0.217 ± 0.155	0.462 ± 0.462
GWIN*	0.244 ± 0.008	0.207 ± 0.032	0.173 ± 0.026	0.158 ± 0.034	0.211 ± 0.097	0.411 ± 0.223
GDET*	0.243 ± 0.009	0.195 ± 0.004	0.165 ± 0.005	0.147 ± 0.003	0.126 ± 0.014	0.110 ± 0.023
GIMP*	0.080 ± 0.030	0.143 ± 0.013	0.130 ± 0.008	0.120 ± 0.007	0.102 ± 0.006	0.098 ± 0.020
PRFT*	0.231 ± 0.008	0.188 ± 0.006	0.162 ± 0.003	0.144 ± 0.003	0.125 ± 0.018	0.106 ± 0.021
GBST*	0.199 ± 0.011	0.166 ± 0.008	0.140 ± 0.006	0.128 ± 0.005	0.107 ± 0.005	0.101 ± 0.024
LWST*	0.172 ± 0.005	0.151 ± 0.003	0.132 ± 0.007	0.117 ± 0.002	0.099 ± 0.004	0.087 ± 0.015
GWST*	0.247 ± 0.000	0.198 ± 0.000	0.168 ± 0.000	0.149 ± 0.000	0.126 ± 0.013	0.123 ± 0.050
TOFF*	0.171 ± 0.001	0.144 ± 0.001	0.150 ± 0.019	0.151 ± 0.016	0.127 ± 0.020	0.107 ± 0.027
GAIN*	0.247 ± 0.000	0.202 ± 0.022	0.176 ± 0.025	0.197 ± 0.119	0.332 ± 0.345	0.642 ± 0.625
PRNK	0.439 ± 0.299	2.899 ± 2.401	1.561 ± 1.313	1.662 ± 1.166	1.755 ± 1.211	1.254 ± 0.718
PSRT	0.448 ± 0.362	2.542 ± 1.845	1.942 ± 1.227	1.654 ± 0.913	0.892 ± 0.602	1.145 ± 0.720
AVER	0.243 ± 0.010	0.196 ± 0.004	0.171 ± 0.018	0.161 ± 0.033	0.213 ± 0.106	0.420 ± 0.250
MAXR	32.39 ± 11.93	24.94 ± 7.451	21.83 ± 6.574	20.56 ± 6.714	13.65 ± 7.143	10.28 ± 5.047
RECC	0.247 ± 0.004	0.213 ± 0.054	0.195 ± 0.127	0.180 ± 0.052	0.217 ± 0.122	0.415 ± 0.323
MEDC	33.54 ± 11.84	26.82 ± 10.69	23.18 ± 9.022	18.23 ± 6.094	18.97 ± 6.974	11.81 ± 4.810
CONC	0.247 ± 0.000	0.202 ± 0.022	0.168 ± 0.000	0.196 ± 0.089	0.200 ± 0.081	0.515 ± 0.411
ADOM	0.153 ± 0.020	0.150 ± 0.005	0.132 ± 0.006	0.119 ± 0.003	0.111 ± 0.025	0.099 ± 0.023
KDOM	0.247 ± 0.000	0.198 ± 0.000	0.185 ± 0.038	0.175 ± 0.048	0.235 ± 0.099	0.547 ± 0.322
VDOM	0.296 ± 0.111	0.339 ± 0.175	0.345 ± 0.220	0.359 ± 0.221	0.313 ± 0.159	0.294 ± 0.195
CEDM	0.174 ± 0.010	0.153 ± 0.003	0.132 ± 0.003	0.121 ± 0.008	0.100 ± 0.002	0.092 ± 0.014
WSUM	0.247 ± 0.000	0.198 ± 0.000	0.172 ± 0.018	0.149 ± 0.000	0.126 ± 0.013	0.110 ± 0.047
LDOM	0.233 ± 0.024	0.184 ± 0.016	0.161 ± 0.011	0.141 ± 0.010	0.124 ± 0.015	0.113 ± 0.032
FAVR	0.239 ± 0.021	0.198 ± 0.037	0.235 ± 0.101	0.304 ± 0.219	0.656 ± 0.595	2.519 ± 2.701
FDOM	0.229 ± 0.029	0.192 ± 0.020	0.184 ± 0.086	0.161 ± 0.034	0.223 ± 0.146	0.613 ± 0.447
PORD	0.363 ± 0.193	0.606 ± 0.414	0.649 ± 0.450	0.525 ± 0.382	-	-
GFAC	23.84 ± 6.686	13.14 ± 9.421	15.47 ± 5.988	15.23 ± 5.980	13.67 ± 5.711	10.11 ± 4.305
SVEC	0.535 ± 0.448	0.784 ± 0.881	0.571 ± 0.581	0.772 ± 0.649	0.630 ± 0.476	0.409 ± 0.338
MEPS	<b>0.075 ± 0.011</b>	<b>0.084 ± 0.005</b>	<b>0.087 ± 0.007</b>	<b>0.098 ± 0.013</b>	0.107 ± 0.046	0.086 ± 0.019
FPDM	0.809 ± 0.437	1.019 ± 0.721	1.101 ± 0.634	1.267 ± 0.868	1.193 ± 0.645	0.923 ± 0.724
SODC	0.247 ± 0.000	0.198 ± 0.000	0.173 ± 0.019	0.158 ± 0.026	0.151 ± 0.055	0.169 ± 0.082
WSCR	0.245 ± 0.005	0.197 ± 0.003	0.173 ± 0.022	0.182 ± 0.048	0.224 ± 0.168	0.439 ± 0.229

alcanzados. Los resultados de este experimento nos indican que a pesar de que la habilidad de convergencia de **MEPS** no fue tan destacable como la de otros enfoques, este método provee mejor preservación de diversidad durante el proceso de búsqueda, lo que le permitió lograr poblaciones finales suficientemente distribuidas como para inclinar a su favor los resultados para la métrica DGI. El método **MEPS** esta basado en  $\epsilon$ -dominance [48], que es una técnica propuesta específicamente para promover simultáneamente distribución y convergencia. Este comportamiento de **MEPS** sólo se observó para estos tres problemas de prueba. En el resto de los casos de prueba (problemas DTLZ3, DTLZ5 y DTLZ6) los niveles de convergencia alcanzados por **MEPS** fueron significativamente inferiores a los mejores logrados, de modo que la buena distribución que **MEPS** mantuvo a lo largo de la búsqueda no fue suficiente para favorecer sus resultados para la métrica en estas instancias del

Tabla 4.8: Media y desviación estándar de la métrica DGI. Problema DTLZ2.

Método	5 Obj.	10 Obj.	15 Obj.	20 Obj.	30 Obj.	50 Obj.
MINR*	0.502 ± 0.004	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
SCOR*	0.502 ± 0.005	0.400 ± 0.000	0.339 ± 0.001	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
SCOO*	0.395 ± 0.002	0.353 ± 0.004	0.316 ± 0.003	0.285 ± 0.001	0.241 ± 0.001	0.193 ± 0.000
GFLS*	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.198 ± 0.003
GWIN*	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.197 ± 0.001
GDET*	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
GIMP*	0.308 ± 0.068	0.339 ± 0.033	0.314 ± 0.015	0.285 ± 0.008	0.242 ± 0.005	0.193 ± 0.002
PRFT*	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
GBST*	0.503 ± 0.000	0.400 ± 0.002	0.339 ± 0.000	0.299 ± 0.001	0.249 ± 0.000	0.196 ± 0.000
LWST*	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
GWST*	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
TOFF*	0.395 ± 0.002	0.386 ± 0.023	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
GAIN*	0.503 ± 0.000	0.400 ± 0.000	0.341 ± 0.002	0.307 ± 0.006	0.284 ± 0.021	0.248 ± 0.021
PRNK	0.233 ± 0.031	0.304 ± 0.022	0.301 ± 0.012	0.282 ± 0.008	0.244 ± 0.006	0.200 ± 0.009
PSRT	0.242 ± 0.040	0.296 ± 0.021	0.300 ± 0.010	0.279 ± 0.012	0.242 ± 0.005	0.199 ± 0.005
AVER	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.197 ± 0.004
MAXR	0.471 ± 0.074	0.393 ± 0.026	0.334 ± 0.023	0.290 ± 0.008	0.244 ± 0.002	0.194 ± 0.001
RECC	0.361 ± 0.053	0.377 ± 0.006	0.327 ± 0.001	0.292 ± 0.000	0.245 ± 0.000	0.209 ± 0.012
MEDC	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.250 ± 0.001	0.226 ± 0.022
CONC	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.203 ± 0.007
ADOM	0.386 ± 0.055	0.377 ± 0.010	0.324 ± 0.011	0.289 ± 0.007	0.243 ± 0.004	0.194 ± 0.001
KDOM	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.204 ± 0.009
VDOM	0.473 ± 0.005	0.392 ± 0.001	0.335 ± 0.001	0.297 ± 0.001	0.248 ± 0.000	0.195 ± 0.000
CEDM	0.503 ± 0.000	0.400 ± 0.003	0.339 ± 0.002	0.299 ± 0.001	0.249 ± 0.001	0.196 ± 0.000
WSUM	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
LDOM	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
FAVR	0.503 ± 0.001	0.401 ± 0.002	0.373 ± 0.019	0.329 ± 0.023	0.304 ± 0.026	0.254 ± 0.024
FDOM	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.201 ± 0.005
PORD	0.225 ± 0.070	0.357 ± 0.011	0.331 ± 0.007	0.299 ± 0.001	-	-
GFAC	0.408 ± 0.050	0.361 ± 0.060	0.334 ± 0.022	0.319 ± 0.019	0.280 ± 0.022	0.236 ± 0.024
SVEC	0.226 ± 0.060	0.360 ± 0.013	0.330 ± 0.009	0.298 ± 0.002	0.249 ± 0.000	0.196 ± 0.000
MEPS	<b>0.144 ± 0.020</b>	<b>0.156 ± 0.017</b>	<b>0.173 ± 0.017</b>	<b>0.185 ± 0.010</b>	<b>0.178 ± 0.006</b>	<b>0.167 ± 0.005</b>
FPDM	0.178 ± 0.033	0.307 ± 0.017	0.303 ± 0.012	0.288 ± 0.012	0.251 ± 0.010	0.210 ± 0.010
SODC	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.001	0.299 ± 0.001	0.249 ± 0.001	0.196 ± 0.000
WSCR	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.197 ± 0.002

experimento.

La tabla 4.9 confirma que el problema DTLZ3 fue difícil de optimizar para la mayoría de los métodos estudiados. En este caso el mejor desempeño para la métrica DGI no puede atribuirse a un único método, ya que **ADOM**, **SCOO\***, **GBST\***, **PRFT\*** y **LWST\*** obtuvieron cada uno el mejor resultado en diferentes instancias.

Para los problemas DTLZ5 (tabla 4.11) y DTLZ6 (tabla 4.12), nuestro método **SCOO\*** fue la alternativa que mostró los mejores resultados en la mayoría de las instancias. Este método, que también obtuvo un buen desempeño para el resto de los problemas, provee buena habilidad de convergencia de acuerdo con los resultados reportados en la Sección 4.4.1. No obstante, los resultados de **SCOO\*** para este experimento nos indican que este método también fue capaz de lograr una

Tabla 4.9: Media y desviación estándar de la métrica DGI. Problema DTLZ3.

Método	5 Obj.	10 Obj.	15 Obj.	20 Obj.	30 Obj.	50 Obj.
MINR*	2.690 ± 2.163	7.290 ± 4.767	7.125 ± 3.503	7.529 ± 3.893	7.837 ± 2.960	6.971 ± 2.381
SCOR*	3.297 ± 1.902	6.229 ± 2.999	7.514 ± 3.452	8.673 ± 3.876	7.045 ± 3.118	6.566 ± 2.908
SCOO*	0.460 ± 0.150	<b>0.362 ± 0.007</b>	0.359 ± 0.079	0.340 ± 0.079	0.332 ± 0.109	0.441 ± 0.300
GFLS*	0.794 ± 0.399	1.474 ± 1.242	2.230 ± 1.113	4.326 ± 2.247	7.730 ± 2.660	10.03 ± 3.656
GWIN*	0.754 ± 0.221	1.310 ± 0.750	2.453 ± 1.400	5.005 ± 2.397	8.856 ± 3.168	9.310 ± 4.129
GDET*	0.557 ± 0.118	0.450 ± 0.118	0.403 ± 0.112	0.359 ± 0.097	0.353 ± 0.106	0.404 ± 0.157
GIMP*	0.402 ± 0.257	0.413 ± 0.121	0.394 ± 0.112	0.429 ± 0.238	0.457 ± 0.176	0.509 ± 0.319
PRFT*	0.532 ± 0.082	0.410 ± 0.043	0.374 ± 0.075	0.342 ± 0.090	<b>0.299 ± 0.080</b>	0.365 ± 0.140
GBST*	0.534 ± 0.095	0.410 ± 0.043	0.359 ± 0.053	<b>0.326 ± 0.059</b>	0.355 ± 0.143	0.370 ± 0.169
LWST*	0.526 ± 0.091	0.432 ± 0.080	0.372 ± 0.074	0.353 ± 0.095	0.371 ± 0.167	<b>0.359 ± 0.227</b>
GWST*	0.526 ± 0.079	0.419 ± 0.092	0.378 ± 0.074	0.357 ± 0.185	0.344 ± 0.108	0.422 ± 0.302
TOFF*	0.655 ± 0.277	0.437 ± 0.109	0.375 ± 0.094	0.330 ± 0.065	0.357 ± 0.162	0.468 ± 0.245
GAIN*	1.272 ± 0.705	4.274 ± 2.774	10.31 ± 5.285	13.59 ± 5.365	16.22 ± 5.273	13.21 ± 5.179
PRNK	25.35 ± 8.844	68.76 ± 22.88	54.79 ± 15.78	42.75 ± 16.06	33.68 ± 11.29	20.93 ± 10.72
PSRT	24.67 ± 8.148	78.86 ± 24.97	61.31 ± 25.82	44.92 ± 18.78	34.60 ± 18.54	23.23 ± 11.42
AVER	0.816 ± 0.447	1.361 ± 0.676	2.720 ± 1.612	4.187 ± 1.823	9.170 ± 2.835	8.729 ± 3.198
MAXR	202.8 ± 59.34	87.95 ± 73.44	44.03 ± 47.36	19.44 ± 24.23	8.300 ± 7.467	5.492 ± 1.617
RECC	16.49 ± 8.461	5.688 ± 2.919	6.489 ± 2.680	8.865 ± 4.158	8.360 ± 3.049	12.32 ± 4.905
MEDC	0.801 ± 0.312	1.193 ± 0.900	1.968 ± 1.009	3.920 ± 2.055	10.23 ± 4.760	17.08 ± 9.668
CONC	0.694 ± 0.315	1.052 ± 0.537	1.571 ± 0.791	3.655 ± 1.953	8.651 ± 2.643	9.922 ± 4.084
ADOM	<b>0.326 ± 0.122</b>	0.400 ± 0.078	<b>0.333 ± 0.065</b>	0.359 ± 0.100	0.339 ± 0.145	0.382 ± 0.176
KDOM	0.752 ± 0.331	0.983 ± 0.631	1.685 ± 0.859	3.434 ± 1.627	8.502 ± 2.442	10.84 ± 2.743
VDOM	1.462 ± 1.729	1.442 ± 1.234	1.265 ± 0.948	0.939 ± 0.539	1.472 ± 1.068	1.397 ± 1.024
CEDM	0.501 ± 0.014	0.425 ± 0.072	0.373 ± 0.075	0.348 ± 0.092	0.371 ± 0.137	0.381 ± 0.160
WSUM	0.527 ± 0.078	0.433 ± 0.081	0.368 ± 0.067	0.342 ± 0.090	0.378 ± 0.212	0.379 ± 0.130
LDOM	0.547 ± 0.109	0.510 ± 0.181	0.470 ± 0.316	0.463 ± 0.165	0.449 ± 0.190	0.538 ± 0.299
FAVR	6.114 ± 14.16	7.898 ± 6.446	45.22 ± 21.32	39.40 ± 19.33	59.13 ± 23.22	58.44 ± 18.72
FDOM	0.758 ± 0.312	1.311 ± 0.699	1.531 ± 0.600	3.460 ± 1.640	9.171 ± 3.836	9.536 ± 3.016
PORD	14.90 ± 6.431	10.35 ± 5.571	6.417 ± 2.710	5.121 ± 2.162	-	-
GFAC	221.9 ± 39.53	102.5 ± 61.40	74.33 ± 31.68	53.11 ± 41.07	32.15 ± 22.98	33.60 ± 16.50
SVEC	13.88 ± 4.561	11.37 ± 5.261	8.142 ± 2.878	4.663 ± 2.732	3.238 ± 1.856	3.046 ± 1.295
MEPS	0.334 ± 0.235	0.923 ± 0.492	1.259 ± 0.829	1.563 ± 1.560	1.566 ± 0.945	1.230 ± 0.622
FPDM	17.37 ± 6.349	24.01 ± 9.149	30.66 ± 12.34	25.96 ± 8.695	17.76 ± 7.840	12.63 ± 5.520
SODC	2.195 ± 1.493	4.616 ± 2.710	5.099 ± 2.102	4.074 ± 1.992	2.926 ± 1.189	2.285 ± 1.069
WSCR	0.699 ± 0.260	1.443 ± 0.801	2.079 ± 1.231	4.349 ± 1.949	8.961 ± 2.998	9.533 ± 3.854

población final con características que favorecen el resultado para la métrica DGI. Estas características pueden incluir diversidad de las soluciones y/o convergencia hacia regiones donde existe un mejor compromiso entre objetivos, que tal como se mencionó en la Sección 2.6.2, son escenarios que favorecen los resultados para la métrica.

Un aspecto interesante que se observó en las tablas 4.8, 4.10 y 4.11, pertinentes a los problemas DTLZ2, DTLZ4 y DTLZ5, es que en cada columna existen muchas repeticiones de los resultados para la métrica. Este comportamiento se puede explicar observando que las correspondientes tablas en la Sección 4.4.1 presentan muchas líneas completas resaltadas por contener los mejores resultados para la métrica de convergencia (lo que sugiere que estos tres problemas no presentan tantas dificultades de convergencia). Retomando lo comentado anteriormente acerca de los escenarios que favorecen

Tabla 4.10: Media y desviación estándar de la métrica DGI. Problema DTLZ4.

Método	5 Obj.	10 Obj.	15 Obj.	20 Obj.	30 Obj.	50 Obj.
MINR*	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.300 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
SCOR*	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.300 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
SCOO*	0.430 ± 0.019	0.366 ± 0.008	0.316 ± 0.004	0.284 ± 0.003	0.239 ± 0.002	0.191 ± 0.001
GFLS*	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.001
GWIN*	0.502 ± 0.006	0.400 ± 0.002	0.339 ± 0.000	0.299 ± 0.001	0.249 ± 0.001	0.195 ± 0.001
GDET*	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
GIMP*	0.429 ± 0.073	0.305 ± 0.037	0.274 ± 0.017	0.262 ± 0.011	0.222 ± 0.007	0.185 ± 0.003
PRFT*	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
GBST*	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
LWST*	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
GWST*	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
TOFF*	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
GAIN*	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
PRNK	0.362 ± 0.032	0.411 ± 0.007	0.387 ± 0.029	0.362 ± 0.023	0.304 ± 0.026	0.245 ± 0.024
PSRT	0.361 ± 0.019	0.411 ± 0.007	0.387 ± 0.029	0.362 ± 0.023	0.304 ± 0.026	0.245 ± 0.024
AVER	0.503 ± 0.000	0.399 ± 0.003	0.339 ± 0.001	0.299 ± 0.001	0.249 ± 0.001	0.196 ± 0.000
MAXR	0.617 ± 0.066	0.488 ± 0.045	0.412 ± 0.036	0.362 ± 0.034	0.302 ± 0.026	0.245 ± 0.020
RECC	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.199 ± 0.005
MEDC	0.558 ± 0.033	0.440 ± 0.023	0.368 ± 0.014	0.319 ± 0.014	0.257 ± 0.007	0.197 ± 0.002
CONC	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.001	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
ADOM	0.432 ± 0.025	0.374 ± 0.016	0.327 ± 0.003	0.292 ± 0.002	0.245 ± 0.002	0.194 ± 0.001
KDOM	0.483 ± 0.031	0.399 ± 0.004	0.339 ± 0.003	0.299 ± 0.000	0.248 ± 0.002	0.196 ± 0.000
VDOM	0.487 ± 0.012	0.393 ± 0.005	0.336 ± 0.001	0.297 ± 0.000	0.248 ± 0.001	0.197 ± 0.002
CEDM	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
WSUM	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
LDOM	0.501 ± 0.012	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
FAVR	0.559 ± 0.045	0.434 ± 0.020	0.416 ± 0.040	0.343 ± 0.028	0.295 ± 0.024	0.240 ± 0.022
FDOM	0.489 ± 0.028	0.399 ± 0.004	0.339 ± 0.002	0.299 ± 0.001	0.249 ± 0.001	0.196 ± 0.001
PORD	0.362 ± 0.030	0.380 ± 0.009	0.332 ± 0.006	0.293 ± 0.003	-	-
GFAC	0.619 ± 0.053	0.499 ± 0.034	0.418 ± 0.033	0.372 ± 0.027	0.305 ± 0.028	0.247 ± 0.020
SVEC	0.505 ± 0.005	0.399 ± 0.006	0.338 ± 0.005	0.298 ± 0.003	0.245 ± 0.002	0.196 ± 0.001
MEPS	<b>0.152 ± 0.066</b>	<b>0.167 ± 0.037</b>	<b>0.161 ± 0.018</b>	<b>0.156 ± 0.014</b>	<b>0.158 ± 0.008</b>	<b>0.148 ± 0.005</b>
FPDM	0.225 ± 0.082	0.399 ± 0.004	0.339 ± 0.001	0.299 ± 0.000	0.321 ± 0.024	0.253 ± 0.025
SODC	0.503 ± 0.000	0.398 ± 0.003	0.337 ± 0.002	0.299 ± 0.001	0.248 ± 0.001	0.195 ± 0.001
WSCR	0.503 ± 0.000	0.399 ± 0.004	0.339 ± 0.001	0.299 ± 0.001	0.249 ± 0.001	0.196 ± 0.000

los resultados para la métrica DGI, este comportamiento nos indica que los diferentes enfoques en los que se observa esta característica lograron converger exactamente a alguno de los extremos del FP para estos problemas; v.g., cualquier extremo proporcionaría el mismo valor para la métrica.

Para dejar claro este punto consideremos el caso del problema DTLZ2 con 5 objetivos. Suponiendo que  $\mathbf{F}(\mathbf{X}) = [0, 0, 0, 0, 1]$  es el punto encontrado por el MOEA (uno de los extremos del problema). El conjunto de referencia utilizado para calcular la métrica DGI es:

$$P^* = \{[0, 0, 0, 0, 1], [0, 0, 0, 1, 0], [0, 0, 1, 0, 0], [0, 1, 0, 0, 0], [1, 0, 0, 0, 0], [\frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}]\}$$

La métrica DGI calcula y promedia la distancia de cada punto en  $P^*$  hacia el conjunto de soluciones encontradas por el MOEA, en este caso el punto  $\mathbf{F}(\mathbf{X})$ . De este modo, las distancias (distancia Euclidiana) de cada punto en  $P^*$  con respecto a  $\mathbf{F}(\mathbf{X})$  son:  $D = [0, \sqrt{2}, \sqrt{2}, \sqrt{2}, \sqrt{2}, \sqrt{1.1056}]$ . Por

lo tanto, el valor para la métrica DGI es:

$$DGI = \frac{\sqrt{\sum_{d \in D} d^2}}{|P^*|} = \frac{\sqrt{9.1056}}{6} = 0.503 \quad (4.4)$$

Como podrá notarse, 0.503 es un resultado que se repite constantemente en la primera columna de la tabla 4.8, y dichos resultados corresponden a algunos de los métodos que obtuvieron el mejor valor para la métrica de convergencia en la Sección 4.4.1. Estos resultados nos permiten observar que los métodos que inducen una fuerte presión de selección son robustos y capaces de lograr una buena convergencia; sin embargo, a la vez pierden la diversidad y convergen prácticamente hacia un sólo punto de la superficie compromiso.

Tabla 4.11: Media y desviación estándar de la métrica DGI. Problema DTLZ5.

Método	5 Obj.	10 Obj.	15 Obj.	20 Obj.	30 Obj.	50 Obj.
MINR*	0.498 ± 0.015	0.401 ± 0.001	0.341 ± 0.001	0.300 ± 0.000	0.250 ± 0.002	0.206 ± 0.008
SCOR*	0.498 ± 0.013	0.401 ± 0.000	0.341 ± 0.000	0.300 ± 0.000	0.251 ± 0.003	0.208 ± 0.007
SCOO*	0.406 ± 0.003	<b>0.360 ± 0.002</b>	<b>0.318 ± 0.003</b>	<b>0.286 ± 0.001</b>	<b>0.242 ± 0.001</b>	<b>0.193 ± 0.002</b>
GFLS*	0.503 ± 0.000	1.062 ± 0.001	0.891 ± 0.000	0.782 ± 0.000	0.647 ± 0.000	0.507 ± 0.000
GWIN*	0.503 ± 0.000	1.062 ± 0.001	0.891 ± 0.000	0.782 ± 0.000	0.647 ± 0.000	0.507 ± 0.000
GDET*	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
GIMP*	0.479 ± 0.021	0.396 ± 0.006	0.336 ± 0.005	0.299 ± 0.002	0.249 ± 0.000	0.196 ± 0.000
PRFT*	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
GBST*	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
LWST*	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
GWST*	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
TOFF*	0.514 ± 0.063	0.429 ± 0.026	0.360 ± 0.020	0.300 ± 0.003	0.249 ± 0.000	0.196 ± 0.000
GAIN*	0.503 ± 0.000	1.062 ± 0.000	0.891 ± 0.000	0.782 ± 0.000	0.647 ± 0.000	0.507 ± 0.000
PRNK	0.421 ± 0.019	0.402 ± 0.019	0.371 ± 0.026	0.326 ± 0.020	0.284 ± 0.020	0.231 ± 0.016
PSRT	0.426 ± 0.024	0.412 ± 0.028	0.361 ± 0.021	0.331 ± 0.022	0.284 ± 0.020	0.232 ± 0.015
AVER	0.503 ± 0.000	1.062 ± 0.001	0.891 ± 0.000	0.782 ± 0.000	0.647 ± 0.000	0.507 ± 0.000
MAXR	0.439 ± 0.077	0.437 ± 0.101	0.396 ± 0.097	0.357 ± 0.066	0.316 ± 0.040	0.279 ± 0.039
RECC	<b>0.352 ± 0.000</b>	0.361 ± 0.025	0.373 ± 0.043	0.378 ± 0.062	0.403 ± 0.067	0.372 ± 0.032
MEDC	0.503 ± 0.000	1.062 ± 0.001	0.891 ± 0.000	0.782 ± 0.000	0.647 ± 0.000	0.507 ± 0.000
CONC	0.503 ± 0.000	1.062 ± 0.000	0.891 ± 0.000	0.782 ± 0.000	0.647 ± 0.000	0.507 ± 0.000
ADOM	0.461 ± 0.030	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
KDOM	0.503 ± 0.000	1.061 ± 0.007	0.891 ± 0.000	0.782 ± 0.000	0.647 ± 0.000	0.507 ± 0.000
VDOM	0.491 ± 0.001	0.397 ± 0.000	0.338 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
CEDM	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
WSUM	0.503 ± 0.000	0.400 ± 0.000	0.339 ± 0.000	0.299 ± 0.000	0.249 ± 0.000	0.196 ± 0.000
LDOM	0.503 ± 0.000	0.407 ± 0.003	0.342 ± 0.002	0.301 ± 0.001	0.250 ± 0.001	0.196 ± 0.001
FAVR	0.503 ± 0.000	0.978 ± 0.029	0.815 ± 0.018	0.744 ± 0.009	0.622 ± 0.006	0.493 ± 0.003
FDOM	0.503 ± 0.000	1.062 ± 0.001	0.891 ± 0.000	0.782 ± 0.000	0.647 ± 0.000	0.507 ± 0.000
PORD	0.483 ± 0.052	0.986 ± 0.118	0.877 ± 0.039	0.782 ± 0.001	-	-
GFAC	0.430 ± 0.075	0.844 ± 0.169	0.732 ± 0.119	0.634 ± 0.098	0.570 ± 0.071	0.471 ± 0.037
SVEC	0.493 ± 0.038	0.998 ± 0.120	0.885 ± 0.010	0.782 ± 0.001	0.647 ± 0.000	0.507 ± 0.000
MEPS	0.375 ± 0.013	0.364 ± 0.009	0.334 ± 0.013	0.309 ± 0.018	0.262 ± 0.014	0.208 ± 0.006
FPDM	0.370 ± 0.013	0.839 ± 0.068	0.785 ± 0.025	0.713 ± 0.015	0.615 ± 0.011	0.496 ± 0.004
SODC	0.503 ± 0.001	0.400 ± 0.000	0.340 ± 0.000	0.301 ± 0.000	0.251 ± 0.001	0.198 ± 0.002
WSCR	0.503 ± 0.000	1.062 ± 0.001	0.891 ± 0.000	0.782 ± 0.000	0.647 ± 0.000	0.507 ± 0.000

Tabla 4.12: Media y desviación estándar de la métrica DGI. Problema DTLZ6.

Método	5 Obj.	10 Obj.	15 Obj.	20 Obj.	30 Obj.	50 Obj.
MINR*	1.541 ± 0.219	1.809 ± 0.301	1.803 ± 0.353	1.613 ± 0.240	1.415 ± 0.159	1.173 ± 0.101
SCOR*	1.543 ± 0.207	1.786 ± 0.306	1.744 ± 0.282	1.635 ± 0.225	1.416 ± 0.158	1.178 ± 0.088
SCOO*	<b>0.415 ± 0.004</b>	<b>0.371 ± 0.005</b>	<b>0.329 ± 0.004</b>	<b>0.298 ± 0.003</b>	<b>0.254 ± 0.003</b>	0.210 ± 0.005
GFLS*	0.523 ± 0.006	3.294 ± 0.000	2.742 ± 0.000	2.398 ± 0.000	1.977 ± 0.000	1.544 ± 0.000
GWIN*	0.521 ± 0.008	3.294 ± 0.001	2.742 ± 0.000	2.398 ± 0.000	1.977 ± 0.000	1.544 ± 0.000
GDET*	0.525 ± 0.007	0.417 ± 0.006	0.355 ± 0.005	0.313 ± 0.005	0.259 ± 0.004	0.205 ± 0.003
GIMP*	0.505 ± 0.062	0.429 ± 0.061	0.363 ± 0.029	0.316 ± 0.027	0.279 ± 0.044	0.224 ± 0.054
PRFT*	0.524 ± 0.008	0.416 ± 0.006	0.353 ± 0.005	0.312 ± 0.005	0.260 ± 0.005	0.205 ± 0.003
GBST*	0.525 ± 0.010	0.416 ± 0.007	0.354 ± 0.006	0.311 ± 0.004	0.260 ± 0.004	<b>0.205 ± 0.002</b>
LWST*	0.512 ± 0.016	0.417 ± 0.007	0.354 ± 0.005	0.314 ± 0.005	0.261 ± 0.003	0.206 ± 0.003
GWST*	0.521 ± 0.008	0.416 ± 0.006	0.354 ± 0.006	0.313 ± 0.005	0.261 ± 0.004	0.206 ± 0.003
TOFF*	3.876 ± 0.210	1.621 ± 1.105	0.355 ± 0.006	0.315 ± 0.005	0.261 ± 0.005	0.205 ± 0.004
GAIN*	4.410 ± 0.000	3.294 ± 0.000	2.742 ± 0.000	2.398 ± 0.000	1.977 ± 0.001	1.543 ± 0.001
PRNK	2.547 ± 0.197	2.632 ± 0.125	2.235 ± 0.108	1.981 ± 0.101	1.678 ± 0.096	1.329 ± 0.054
PSRT	2.548 ± 0.225	2.642 ± 0.118	2.224 ± 0.113	1.978 ± 0.103	1.692 ± 0.072	1.333 ± 0.045
AVER	0.524 ± 0.009	3.294 ± 0.000	2.742 ± 0.000	2.398 ± 0.000	1.977 ± 0.000	1.544 ± 0.000
MAXR	2.895 ± 0.923	2.804 ± 0.265	2.356 ± 0.203	2.076 ± 0.178	1.668 ± 0.132	1.274 ± 0.081
RECC	2.280 ± 0.437	2.452 ± 0.187	2.235 ± 0.153	1.955 ± 0.160	1.752 ± 0.107	1.450 ± 0.082
MEDC	0.542 ± 0.016	3.294 ± 0.001	2.742 ± 0.000	2.398 ± 0.000	1.977 ± 0.000	1.544 ± 0.000
CONC	0.522 ± 0.006	3.294 ± 0.000	2.742 ± 0.000	2.398 ± 0.000	1.977 ± 0.000	1.544 ± 0.000
ADOM	0.428 ± 0.002	0.475 ± 0.023	0.429 ± 0.015	0.384 ± 0.014	0.327 ± 0.014	0.260 ± 0.009
KDOM	0.523 ± 0.007	3.294 ± 0.000	2.742 ± 0.000	2.398 ± 0.000	1.977 ± 0.000	1.544 ± 0.000
VDOM	0.431 ± 0.012	0.399 ± 0.007	0.355 ± 0.012	0.316 ± 0.010	0.288 ± 0.035	0.251 ± 0.041
CEDM	0.522 ± 0.007	0.416 ± 0.008	0.352 ± 0.005	0.312 ± 0.005	0.261 ± 0.004	0.206 ± 0.004
WSUM	0.524 ± 0.008	0.417 ± 0.005	0.355 ± 0.005	0.313 ± 0.005	0.260 ± 0.004	0.205 ± 0.003
LDOM	0.522 ± 0.007	2.309 ± 0.091	1.844 ± 0.093	1.572 ± 0.071	1.260 ± 0.078	0.908 ± 0.041
FAVR	2.531 ± 0.894	3.241 ± 0.058	2.587 ± 0.061	2.359 ± 0.018	1.922 ± 0.019	1.484 ± 0.019
FDOM	0.526 ± 0.007	3.294 ± 0.000	2.742 ± 0.000	2.398 ± 0.000	1.977 ± 0.000	1.544 ± 0.000
PORD	2.030 ± 0.250	3.226 ± 0.046	2.726 ± 0.009	2.392 ± 0.004	-	-
GFAC	3.381 ± 0.587	3.049 ± 0.080	2.640 ± 0.062	2.329 ± 0.048	1.946 ± 0.028	1.528 ± 0.019
SVEC	2.059 ± 0.246	3.223 ± 0.042	2.724 ± 0.014	2.392 ± 0.005	1.976 ± 0.001	1.543 ± 0.000
MEPS	0.643 ± 0.080	1.697 ± 0.112	1.660 ± 0.134	1.548 ± 0.124	1.316 ± 0.117	1.045 ± 0.090
FPDM	1.746 ± 0.168	2.883 ± 0.063	2.561 ± 0.032	2.275 ± 0.025	1.910 ± 0.015	1.513 ± 0.009
SODC	0.583 ± 0.017	0.559 ± 0.041	0.641 ± 0.083	0.723 ± 0.085	0.872 ± 0.089	0.930 ± 0.061
WSCR	0.523 ± 0.009	3.294 ± 0.000	2.742 ± 0.000	2.398 ± 0.000	1.977 ± 0.000	1.544 ± 0.000

### 4.4.3 Convergencia en línea

Un aspecto que sin duda es de gran importancia al evaluar un MOEA es su comportamiento en las diferentes etapas del proceso de búsqueda. En esta sección analizaremos cómo se desempeño el MOEA (en términos de convergencia) durante la búsqueda al utilizar los diferentes enfoques de jerarquización estudiados. Las figuras 4.5 a 4.10 muestran los resultados de este experimento para los problemas DTLZ1-DTLZ6, respectivamente. La interpretación de estos mosaicos de datos es la siguiente (ver figura 4.4): cada tercia de renglones corresponde al MOEA utilizando el método de jerarquización indicado a la izquierda, un renglón para cada tamaño de problema. Con la intención de aprovechar el espacio y al mismo tiempo presentar la mayor cantidad de información posible, sólo se

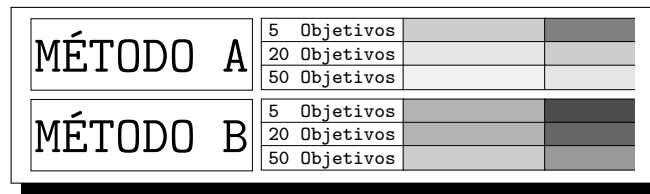


Figura 4.4: Interpretación de los mosaicos de datos.

incluyen las instancias con 5, 20 y 50 objetivos para cada uno de los problemas de prueba adoptados. Las columnas representan el progreso de la búsqueda en pasos de 10 generaciones. Los valores para la métrica de convergencia son representados por celdas en escala de grises, donde el tono más oscuro hará referencia a la mejor aproximación lograda. De esta manera, estos mosaicos permiten tanto comparar resultados entre distintos enfoques como analizar el deterioro de cada estrategia al aumentar el número de objetivos. Los datos contenidos en los elementos del mosaico corresponden a la media de 31 repeticiones para cada experimento, y están representados en escala logarítmica con la finalidad de resaltar las diferencias en los resultados obtenidos al utilizar cada alternativa.

La figura 4.5 confirma que **MAXR**, **MEDC** y **GFAC** son los métodos con el peor desempeño para el problema DTLZ1. Sin embargo, un aspecto importante que revela esta figura es que el desempeño de estos tres enfoques no solo es muy malo para este problema, si no que en lugar de guiar el proceso de búsqueda hacia el frente de Pareto (convergencia), estos métodos llevan a su población de soluciones lejos de él conforme avanza la búsqueda (divergencia). La figura 4.5 muestra que métodos como **GDET\***, **PRFT\***, **ADOM** y **LDOM**, entre otros, presentan una convergencia acelerada, ya que desde la primera mitad del proceso de búsqueda lograron valores relativamente bajos para la métrica de convergencia en la mayoría de las instancias.

Para el problema DTLZ2, la figura 4.6 muestra que muchos de los métodos presentaron un comportamiento divergente en alguna etapa de la búsqueda<sup>1</sup>. Los métodos que presentaron esta propiedad son los siguientes: nuestras técnicas **GFLS\***, **GWIN\*** y **GAIN\***, y los métodos de la literatura **AVER**, **MAXR**, **RECC**, **MEDC**, **CONC**, **KDOM**, **FAVR**, **FDOM**, **GFAC** y **WSCR**. Por otra parte, tanto nuestros métodos **GDET\***, **PRFT\***, **GBST\***, **LWST\*** y **GWST\***,

<sup>1</sup>Decimos que un método presenta un comportamiento divergente cuando en algún punto de la búsqueda sus resultados para la métrica de convergencia empeoran en lugar de mejorar (independientemente del resultado final).

como **CEDM**, **WSUM** y **LDOM**, mostraron claramente la mejor velocidad de convergencia, ya que al transcurrir las primeras 100 generaciones lograron una buena aproximación a sus mejores niveles para la métrica.

La coloración tenue de la figura 4.7 nos indica nuevamente que el problema DTLZ3 fue más difícil de optimizar por los diferentes enfoques. Al igual que para los problemas DTLZ1 y DTLZ2, es posible observar divergencia en el comportamiento de los métodos **MAXR** y **GFAC**, específicamente para la instancia más pequeña (primer renglón de cada método). Los métodos que lograron la mejor convergencia para este experimento requirieron de alrededor de 200 generaciones para aproximarse a estos resultados. Sin embargo, puede notarse que estos enfoques en ningún momento dejaron de mejorar, por lo que es posible que con un mayor número de generaciones les sea posible alcanzar la superficie compromiso para este problema.

El problema DTLZ4 no fue la excepción en cuanto a la divergencia mostrada por varios métodos (ver figura 4.8) en alguna etapa de la búsqueda. Esta característica puede observarse en **MAXR**, **GFAC**, **FAVR**, **FPDM**, **PRNK** y **PSRT**. Muchos de los métodos mostraron una convergencia acelerada para el problema DTLZ4, ya que en el primer tercio de la búsqueda lograron niveles relativamente bajos para la métrica de convergencia.

La mayor parte de los esquemas de jerarquización estudiados provocaron que el MOEA se alejara (divergiera) de la superficie óptima para los problemas DTLZ5 y DTLZ6 (figuras 4.9 y 4.10). Tal es la situación que es más sencillo mencionar los métodos que presentaron buenas características de convergencia para estos problemas. Por una parte, los métodos **GDET\***, **PRFT\***, **GBST\***, **LWST\***, **GWST\***, **CEDM** y **WSUM** se desempeñaron bien para ambos problemas. Por otro lado, **GIMP\*** y **SCOO\*** mostraron cada uno un buen desempeño sólo para alguno de estos problemas, DTLZ5 y DTLZ6, respectivamente. Para ambos problemas podría decirse que en las primeras 150 generaciones los métodos mencionados aproximaron sus mejores niveles de convergencia.

En general, la mejor velocidad de convergencia se observó al utilizar nuestros métodos **SCOO\***, **GDET\***, **PRFT\***, **GBST\***, **LWST\*** y **GWST\***, y los enfoques **CEDM**, **WSUM** y **LDOM** del estado del arte, ya que su desempeño fue consistente en la mayoría de las instancias de este experimento. Por otra parte, para cada uno de los problemas de prueba se identificó un conjunto de



alternativas que mostraron divergencia en alguna parte del proceso de búsqueda. La divergencia indica que bajo ciertas condiciones estas técnicas establecen erróneamente preferencias entre soluciones. Una característica en común de la mayoría de los enfoques que mostraron divergencia, es que jerarquizan utilizando como información únicamente la cantidad de objetivos en los que una solución es superior que otra. Sin embargo, no necesariamente porque un individuo  $X_i$  sea mejor que otro individuo  $X_j$  en más del 50 % de los objetivos quiere decir que  $X_i$  es una mejor alternativa que  $X_j$ , ya que estas mejoras pueden ser marginales y es posible que en realidad  $X_j$  sea significativamente mejor que  $X_i$  en el resto de los criterios. De este modo, una jerarquización de este tipo descarta información valiosa (como la magnitud de diferencia en los objetivos de un par de soluciones), lo que puede llevar a malas decisiones de discriminación.

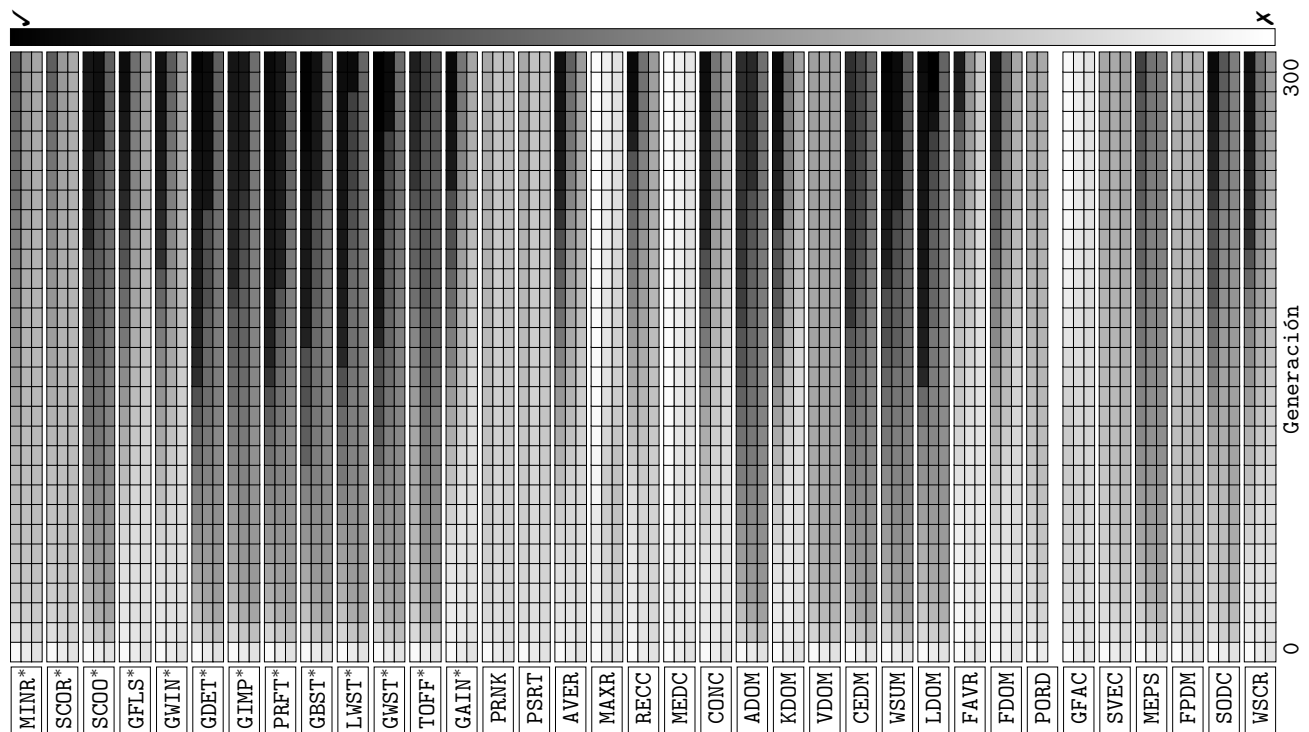


Figura 4.5: Convergencia en línea. Problema DTLZ1 con 5, 20 y 50 objetivos.

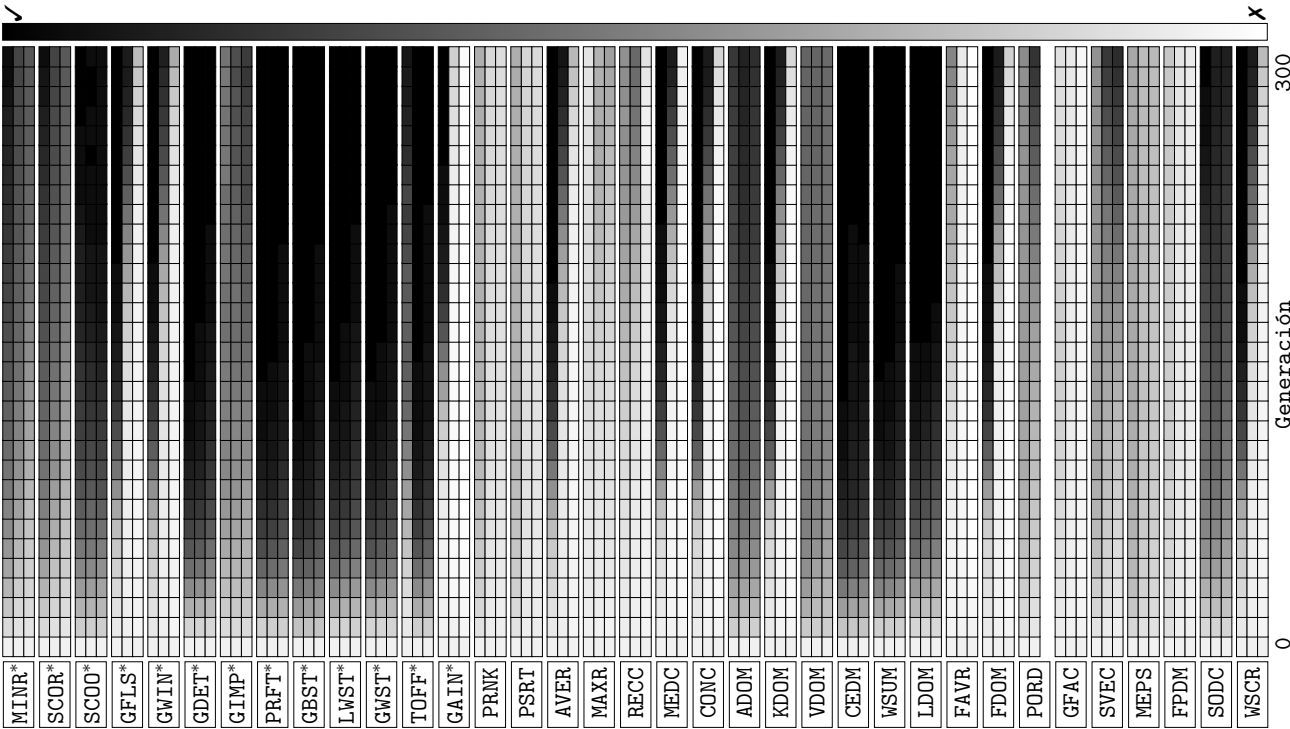


Figura 4.6: Convergencia en línea. Problema DTLZ2 con 5, 20 y 50 objetivos.

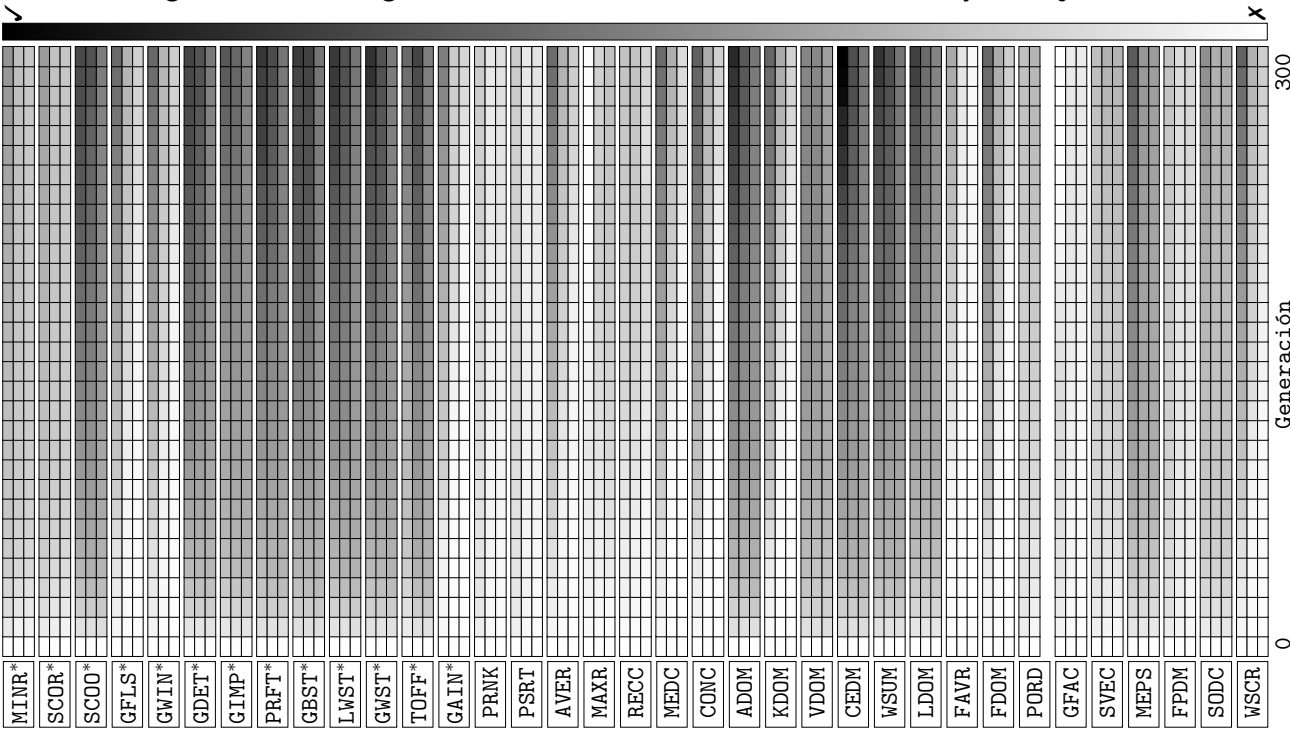


Figura 4.7: Convergencia en línea. Problema DTLZ3 con 5, 20 y 50 objetivos.

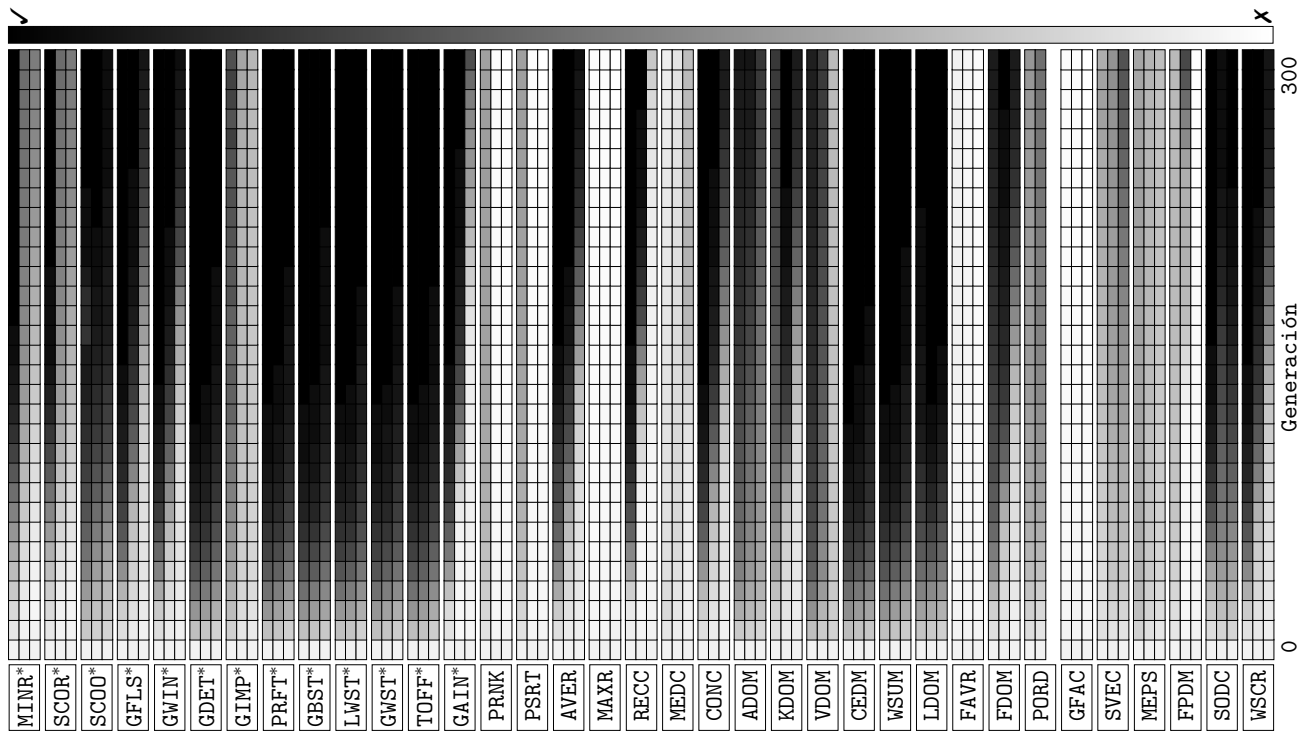


Figura 4.8: Convergencia en línea. Problema DTLZ4 con 5, 20 y 50 objetivos.

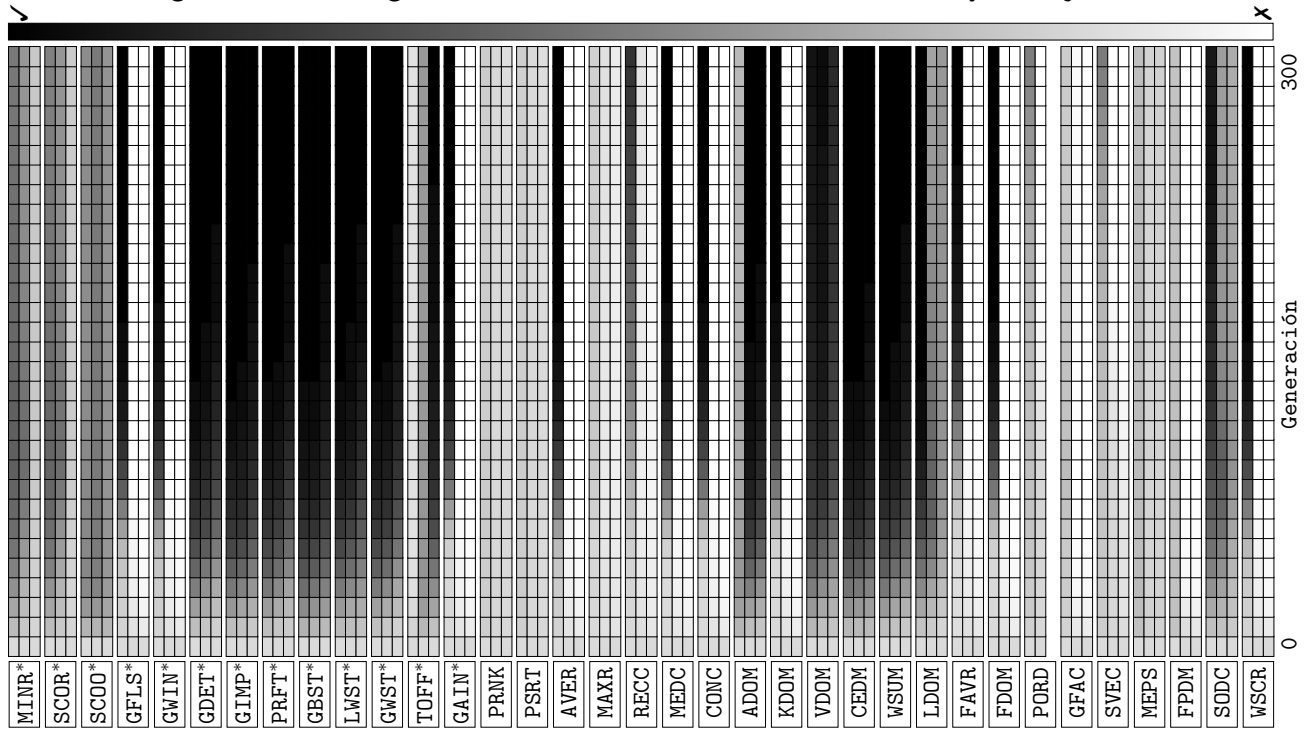


Figura 4.9: Convergencia en línea. Problema DTLZ5 con 5, 20 y 50 objetivos.

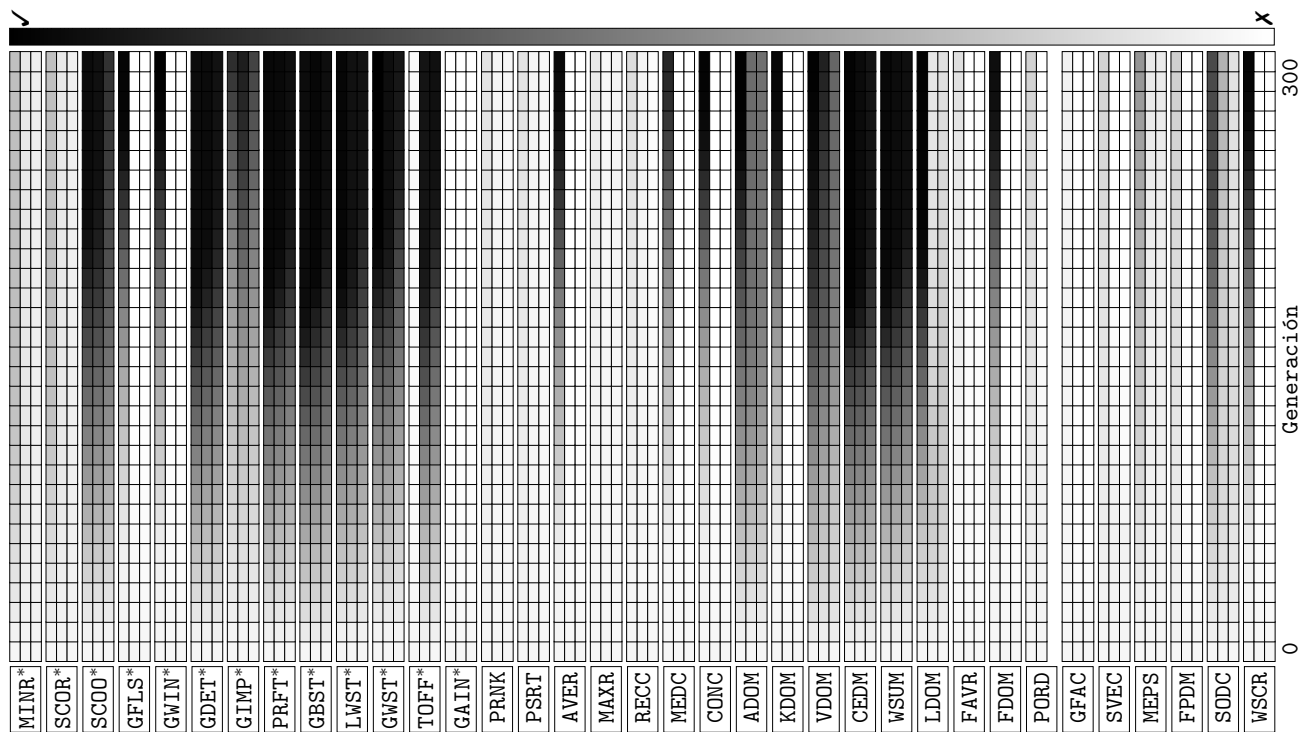


Figura 4.10: Convergencia en línea. Problema DTLZ6 con 5, 20 y 50 objetivos.

#### 4.4.4 Síntesis de resultados

Con la finalidad de facilitar y apoyar la interpretación de los resultados obtenidos en las secciones anteriores, en este apartado se presentan dichos resultados de manera gráfica y sintetizada.

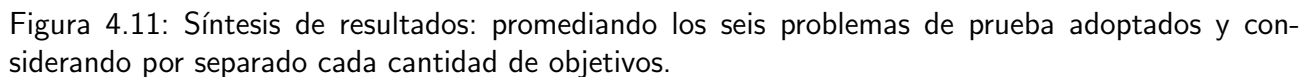
La figura 4.11 sintetiza los resultados obtenidos para la métrica de convergencia y métrica DGI. En estas gráficas se han promediado los resultados obtenidos para los seis problema de prueba adoptados (DTLZ1-DTLZ6) y se han considerado por separado los distintos tamaños de problema:  $M = \{5, 10, 15, 20, 30, 50\}$ . Al igual que la figura 4.11, la figura 4.12 representa los resultados obtenidos para las métricas de convergencia y DGI. Sin embargo, en este caso se han promediado los resultados obtenidos para los diferentes tamaños de problema ( $M = \{5, 10, 15, 20, 30, 50\}$ ) y se han considerado por separado los seis problema de prueba adoptados (DTLZ1-DTLZ6). En estas figuras, cada uno de los métodos estudiados se representa como un punto en un plano de dos dimensiones y sus coordenadas corresponden a los promedios obtenidos para cada una de las métricas utilizadas.

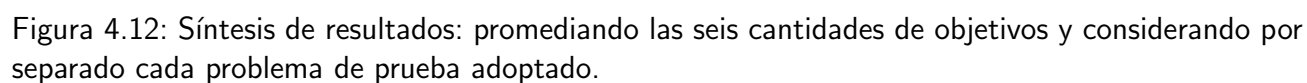
Es preciso recordar que tanto la métrica de convergencia, como la métrica DGI, son criterios de

minimización, de tal forma, se busca identificar aquellos métodos más cercanos a la esquina inferior izquierda en cada una de las gráficas presentadas.

A partir de las figuras 4.11 y 4.12 es posible agrupar los métodos estudiados en cuatro categorías diferentes de acuerdo al desempeño que mostraron:

1. Métodos con el mejor desempeño. En las diferentes gráficas de las figuras 4.11 y 4.12 es posible identificar un conjunto reducido de métodos que obtuvieron los mejores resultados: nuestros métodos **GBST\***, **GDET\***, **GWST\***, **LWST\***, **PRFT\***, y los métodos tomados de la literatura **WSUM**, **CEDM**. Por lo tanto, es este conjunto de métodos el que, de acuerdo con la experimentación realizada, provee el comportamiento más robusto en términos de convergencia y en relación con el incremento del número de objetivos.
2. Métodos con buen desempeño. Consideramos que nuestras propuestas **SCOO\***, **TOFF\*** y **GIMP\***, así como los métodos **ADOM** y **LDOM**, mostraron un buen desempeño, ya que sus resultados fueron también competitivos en la mayoría de los casos.
3. Métodos con un desempeño medio. Existe un conjunto de estrategias cuyo desempeño fue próximo al desempeño promedio en la mayoría de los casos. Este conjunto incluye: nuestros métodos **MINR\***, **SCOR\***, **GFLS\***, **GWIN\***, y los métodos tomados de la literatura **CONC**, **KDOM**, **FDOM**, **WSCR**, **MEPS**, **VDOM**, **SODC**, **RECC** y **AVER**.
4. Métodos con un mal desempeño. Los métodos cuyo mal desempeño se hizo presente en la mayoría de los casos son los siguientes: **MAXR**, **GFAC**, **MEDC**, **PRNK**, **PSRT**, **PORD**, **FAVR**, **FPDM**, **SVEC**, y así mismo nuestro enfoque **GAIN\***.





## 4.5 Conclusiones del capítulo

En este capítulo se estudió el comportamiento de distintos métodos de asignación de aptitud cuando son usados en un MOEA.

Uno de nuestros hallazgos principales fue que mediante un mecanismo de jerarquización que permita una discriminación estricta entre soluciones es posible converger en espacios de objetivos con alta dimensionalidad. Relacionando los experimentos realizados, los métodos que mostraron el mejor desempeño en términos de convergencia en la Sección 4.4.1, fueron también identificados previamente por proveer un alto nivel de discriminación en la Sección 4.2. Así mismo, una característica que comparten los diferentes enfoques que mostraron buen comportamiento convergente, es que consideran de alguna manera qué tan significativamente es mejor una solución que otra en cada dimensión (objetivo) del problema, lo que en nuestra opinión es un aspecto indispensable para que un método de asignación de aptitud sea efectivo y provea una discriminación adecuada. Contrastando un poco y en apoyo a lo anterior, en la Sección 4.4.3 se identificó un conjunto de alternativas que mostraron un comportamiento divergente en alguna etapa del proceso de búsqueda. Una característica en común de los enfoques que mostraron divergencia (salvo algunas excepciones), es que no toman en cuenta este aspecto al momento de establecer preferencias entre soluciones, de tal modo que descartar esta valiosa información puede llevar a malas decisiones de discriminación y, por lo tanto, afectar (negativamente) la habilidad de búsqueda del MOEA.

Por otra parte, tal como los resultados de la Sección 4.4.2 lo indican, cuando se ejerce una máxima presión de selección se tiende a sacrificar la diversidad. De tal forma, los métodos que fueron identificados por proveer una buena capacidad de convergencia por lo general convergieron hacia un único punto de la superficie compromiso, ya que desde etapas tempranas de la búsqueda la mayor parte de su población se compone de individuos duplicados o muy similares. Este fenómeno de converger hacia un sólo punto de la superficie compromiso, o hacia una pequeña región de ésta, se conoce como *deriva genética* (*genetic drift*) y es en parte consecuencia de una fuerte presión de selección y del uso de elitismo en los algoritmos evolutivos [2]. Cuando la población se satura de soluciones duplicadas, las capacidades de exploración de un algoritmo evolutivo se deterioran



claramente, ya que los operadores de cruce no tienen el efecto deseado y la generación de nuevos individuos (diferentes) recae únicamente en el proceso de mutación. La existencia de soluciones duplicadas en la población también tiene efectos negativos sobre el proceso de selección [2], debido a que la probabilidad de que un determinado individuo sea seleccionado se multiplica por el número de duplicados que éste tiene, de modo que los individuos repetidos tendrán ciertas ventajas selectivas y, por lo tanto, el proceso de selección, que es el encargado de guiar la búsqueda, no estará únicamente basado en la aptitud de las soluciones. La clave para evitar el problema de la deriva genética es encontrar una forma para preservar la diversidad en la población [11].

Converger de una manera distribuida (con diversidad) permitirá proveer al tomador de decisiones una aproximación suficientemente representativa de la superficie compromiso. Sin embargo, se reconoce que satisfacer los requerimientos de un optimizador multiobjetivo es en sí un problema con objetivos múltiples, ya que la convergencia y diversidad son dos criterios de optimización que claramente se encuentran en conflicto. De modo que al ejercer la mayor presión de selección para lograr la mejor convergencia se sacrifica diversidad, mientras que la incorporación de mecanismos para preservar diversidad tiende a deteriorar la convergencia.

Generalmente los MOEAs reportados en la literatura utilizan la relación de dominancia de Pareto como criterio de convergencia, incorporando adicionalmente algún mecanismo de preservación de diversidad como criterio secundario para discriminar entre soluciones que comparten la misma jerarquía. Sin embargo, se sabe que al incrementarse la cantidad de objetivos la mayor parte de la población con la que trabaja un MOEA se satura rápidamente de soluciones no dominadas (ver Sección 2.4.2), de tal forma que el proceso de guiar la búsqueda recae prácticamente en los mecanismos de diversidad. Bajo estas condiciones, los mecanismos que promueven diversidad pueden ser nocivos para los MOEAs, ya que tienden a preferir soluciones con una convergencia pobre, dirigiendo la búsqueda lejos de la superficie compromiso [61, 1]. De esta manera, el conflicto entre los requerimientos de convergencia y diversidad se acentúa con el aumento en el número de objetivos a optimizar.

Desarrollar algoritmos capaces de converger en espacios de objetivos con alta dimensionalidad y que al mismo tiempo lo hagan manteniendo diversidad es un reto claro de investigación. No obstante, el requerimiento de convergencia debe ser la prioridad en todos los casos, ya que en realidad es

preferible converger hacia un único punto de la superficie compromiso, que obtener un conjunto bien distribuido de soluciones que se encuentren lejos del óptimo [38]. En el Capítulo 5 se proponen algunas aproximaciones que intentan dar solución a este problema considerando el conocimiento obtenido en esta experimentación.

# 5

## Algoritmos Propuestos

### 5.1 Introducción

El estudio reportado en el Capítulo 4 demostró que para lograr una convergencia robusta es indispensable contar con un esquema de jerarquización que provea una fuerte presión de selección mediante una discriminación estricta entre soluciones. Sin embargo, también se observó que una presión de selección elevada tiende a sacrificar la diversidad y por lo general se converge hacia un único punto de la superficie compromiso. Este problema, conocido como deriva genética (ver Sección 4.5), puede evitarse mediante la incorporación de mecanismos para promover diversidad en la población.

Al igual que la convergencia, la diversidad es un requerimiento cuando se optimizan múltiples objetivos; sin embargo, estos requerimientos están en conflicto y la mejor diversidad esta comúnmente asociada con una convergencia pobre [61]. Por lo tanto, la preservación de diversidad debe hacerse de un modo tal que no afecte negativamente la convergencia.

Generalmente la diversidad se promueve como criterio secundario para discriminar entre soluciones que comparten la misma jerarquía. No obstante, cuando se tiene una máxima discriminación de

soluciones, dado que cada solución tendría por lo regular una jerarquía diferente, utilizar mecanismos de promoción de diversidad como criterio secundario claramente no tendría ningún efecto, por lo que resulta necesario desarrollar enfoques con una metodología más adecuada para solventar este tipo de escenarios.

En este capítulo proponemos dos nuevos MOEAs cuya característica principal es que mantienen como prioridad el requerimiento de convergencia al mismo tiempo que incorporan mecanismos para preservar la existencia de diversidad en la población. Para la validación estadística de estas propuestas seleccionamos un conjunto de MOEAs representativos del estado del arte y realizamos un estudio comparativo. Los resultados de dicho estudio demuestran que los algoritmos propuestos son significativamente superiores a los enfoques tomados de la literatura, al menos para los experimentos realizados.

## 5.2 Algoritmo base

Los dos enfoques propuestos se han implementado sobre un algoritmo base cuyo flujo de ejecución se describe mediante la figura 5.1. Inicialmente se genera de manera aleatoria una población de  $N$  padres. Sobre esta población se realiza un proceso de selección para identificar los padres que podrán reproducirse. Con los padres seleccionados se aplican los operadores de variación (cruza y mutación) para generar una población de  $N$  hijos. La población de padres y la nueva población de hijos son ahora combinadas, lo que da como resultado una población de tamaño  $2N$ , de donde serán seleccionados los  $N$  individuos que formarán la población de padres para la siguiente generación (MOEA elitista [18]). Este proceso se repite durante un número determinado de generaciones.

Los operadores implementados son: selección por torneo binario (*binary tournament selection*) basada en la jerarquía (aptitud) de las soluciones. Cruza binaria simulada (*simulated binary crossover*, SBX) ( $\eta_c = 15$ ) con probabilidad de 1.0 [21]. Mutación polinomial (*polynomial mutation*) ( $\eta_m = 20$ ) con probabilidad de  $1/n$ , donde  $n$  es el número de variables de decisión [19].

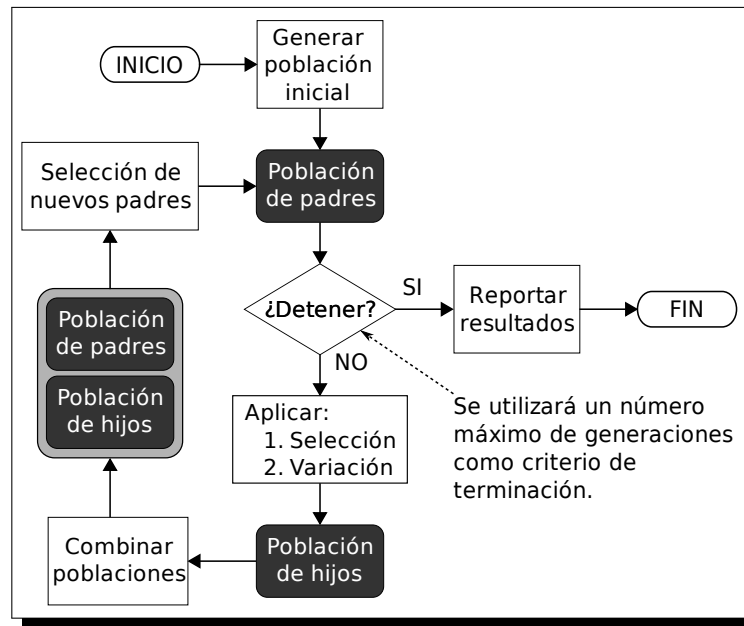


Figura 5.1: Algoritmo base: Flujo de ejecución.

### 5.3 Algoritmo genético elitista basado en agrupamiento

El agrupamiento (o *clustering* como es más comúnmente conocido<sup>1</sup>) permite la clasificación de una colección de elementos con base en sus similitudes. El algoritmo genético elitista basado en *clustering* (*clustering-based elitist genetic algorithm*, o **CEGA\*** por brevedad) es un MOEA que incorpora *clustering* como mecanismo para preservación de diversidad<sup>2</sup>, al mismo tiempo que utiliza una estrategia de jerarquización estricta para promover la convergencia.

**CEGA\*** se implementó sobre el algoritmo base descrito en la Sección 5.2. El flujo básico de ejecución del algoritmo resultante se muestra en la figura 5.2. En la figura 5.2 se encuentran resaltados mediante recuadros punteados los elementos clave de **CEGA\***. Por un lado, un mecanismo de jerarquización que será utilizado para asignar a los individuos un valor (aptitud) para competir durante el proceso de selección. Por otro lado, una estrategia basada en *clustering* que en conjunción con el mecanismo de jerarquización permitirá la selección de los  $N$  individuos que sobrevivirán y formarán la población de padres para la siguiente generación. Estos elementos serán descritos por separado a

<sup>1</sup>En esta tesis se utilizarán indistintamente los términos agrupamiento y *clustering*.

<sup>2</sup>La idea de utilizar *clustering* para promover diversidad en MOEAs no es nueva (ver por ejemplo [75, 55]).

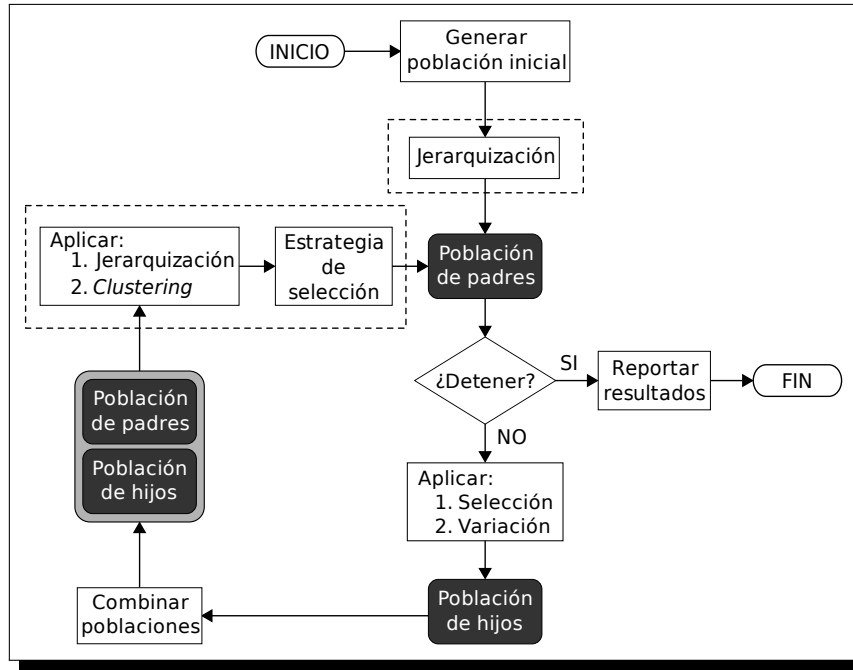


Figura 5.2: CEGA\*: Flujo de ejecución.

continuación.

### 5.3.1 Jerarquización

CEGA\* incorpora el método GDET\* descrito en la Sección 3.3.2 como mecanismo discriminante. GDET\* consiste en penalizar a cada individuo, por la magnitud de la diferencia con que es superado al compararse, objetivo por objetivo, con respecto al resto de la población. De acuerdo con este método la aptitud de un determinado individuo  $\mathbf{X}_i$  se calcula de la siguiente manera:

$$\text{GDET}^*(\mathbf{X}_i) = \sum_{\mathbf{X}_j \neq \mathbf{X}_i} \sum_{m=1}^M \max(f_m(\mathbf{X}_i) - f_m(\mathbf{X}_j), 0) \quad (5.1)$$

GDET\* es un criterio de minimización, por lo que se dice que un individuo  $\mathbf{X}_i$  domina a otro individuo  $\mathbf{X}_j$  si y sólo si  $\text{GDET}^*(\mathbf{X}_i) < \text{GDET}^*(\mathbf{X}_j)$ . Consultar Sección 3.3.2 para mayores detalles de este método.

### 5.3.2 Proceso de agrupamiento

El algoritmo de *clustering* implementado para CEGA\* sigue un enfoque *jerárquico aglomerativo* (ver figura 5.3). Inicialmente cada individuo de la población representa un *cluster* diferente y de manera iterativa se combinan los dos *clusters* más similares hasta completar la cantidad deseada de grupos.

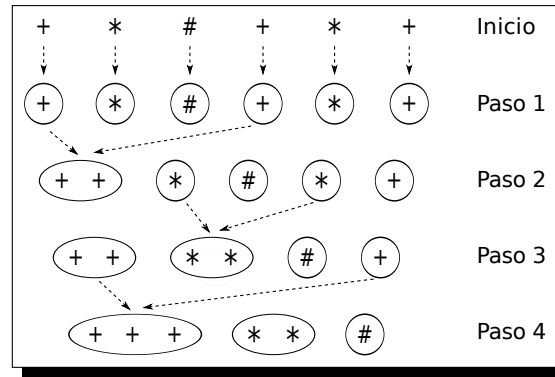


Figura 5.3: Agrupamiento (*Clustering*) jerárquico aglomerativo.

El algoritmo de *clustering* se aplicará en el espacio de las variables de decisión. El propósito de utilizar *clustering* es asistir en la selección de los individuos mejor distribuidos en el espacio de las variables de decisión, de esta manera promover la diversidad en la población y, por lo tanto, favorecer la capacidad exploratoria del algoritmo. Como medida de similitud entre *clusters* se utilizará la distancia promedio (*average linkage*): la distancia entre dos *clusters*  $c_1$  y  $c_2$  es el promedio de las distancias entre cada pareja de soluciones  $\mathbf{X}_i$  y  $\mathbf{X}_j$  tales que  $\mathbf{X}_i \in c_1$  y  $\mathbf{X}_j \in c_2$ <sup>3</sup>. El algoritmo de *clustering* se aplicará sobre los  $2N$  individuos que forman la unión de las poblaciones de padres e hijos (ver figura 5.2). Se seleccionarán  $N$  individuos para formar la población de padres de la siguiente generación, de tal manera que la cantidad de *clusters* necesaria para este enfoque se encuentra en el rango  $[2, N]$ . Todos nuestros experimentos involucran una población de  $N = 100$  individuos y, por razones que se discutirán posteriormente en la Sección 5.3.4, para este trabajo se utilizarán un total de  $C = N/2 = 50$  *clusters*.

<sup>3</sup>En este trabajo se usó la distancia Euclidiana para calcular la cercanía entre un par de soluciones.

### 5.3.3 Estrategia de selección

Esta etapa consiste en seleccionar  $N$  de los  $2N$  individuos en la población combinada de padres e hijos para que sobrevivan y formen la población de padres de la siguiente generación. Este proceso de selección se realizará de la siguiente manera:

1. Calcular la aptitud de cada individuo en la población combinada utilizando el método **GDET\*** descrito en la Sección 5.3.1.
2. Generar  $C$  clusters sobre los  $2N$  individuos en la población combinada de la forma descrita en la Sección 5.3.2.
3. Seleccionar el mejor representante de cada uno de los  $C$  clusters. Para este propósito se jerarquizan de manera local los individuos dentro de un mismo cluster de acuerdo con el método **GBST\*** descrito en la Sección 3.3.2 (ver figura 5.4). Si se presenta el caso en que haya más de un individuo en el mismo cluster con la mejor aptitud de acuerdo al método **GBST\***, se elegirá uno de manera aleatoria.

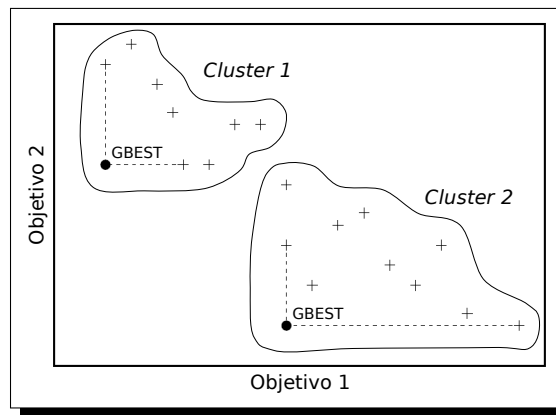


Figura 5.4: **GBST\*** aplicado localmente para elegir el individuo representante de cada cluster.

4. En el paso 3 se seleccionaron  $C$  individuos. En caso de que  $C < N$ , seleccionar los  $N - C$  individuos restantes considerando únicamente su jerarquía global obtenida con **GDET\*** en el paso 1 (omitiendo los individuos ya seleccionados previamente).



De esta manera, la estrategia de selección propuesta mantiene como prioridad la convergencia al mismo tiempo que promueve la existencia de diversidad en la población.

### 5.3.4 Justificación de ajuste de parámetros y operadores

El desempeño de **CEGA\*** depende de distintos factores:

1. Estrategia utilizada ( $E$ ): se implementaron 14 estrategias diferentes para el algoritmo,  $E = \{\text{CEGA1}, \text{CEGA2}, \dots, \text{CEGA14}\}$ . Estas alternativas a grandes rasgos difieren en la forma de utilizar la jerarquización y el procedimiento de *clustering* para guiar el proceso de búsqueda.
2. Medida de similitud entre *clusters* ( $S$ ): Se implementaron tres enfoques para caracterizar la similitud de un par de *clusters* ( $|S| = 3$ ): distancia promedio (*average linkage*), mínima distancia (*single linkage*) y máxima distancia (*complete linkage*).
3. Número de *clusters* ( $C$ ): Tal como se mencionó en la Sección 5.3.2, el número de *clusters* posibles para nuestro enfoque se encuentra en el rango  $[2, N]$ . Sin embargo, fue necesario discretizar este rango y, dado que para nuestra experimentación utilizamos una población de tamaño  $N=100$ , los valores considerados para este parámetro son  $C = \{25, 50, 75\}$ .
4. Método de asignación de aptitud ( $A$ ): Distintos de nuestros métodos propuestos en el Capítulo 3.3 demostraron proveer buena capacidad de convergencia para la experimentación reportada en el Capítulo 4. Una vez más fue necesario discretizar nuestras opciones y decidimos experimentar con  $A = \{\text{PRFT}^*, \text{GBST}^*, \text{GDET}^*\}$ .

Los diferentes ajustes de parámetros descritos anteriormente generan un total de  $|E| \times |S| \times |C| \times |A| = 14 \times 3 \times 3 \times 3 = 378$  configuraciones posibles para el algoritmo. Las distintas configuraciones proveen comportamientos muy diferentes para **CEGA\***, de tal forma, se tomó la decisión de verificar estadísticamente cada una de estas alternativas para seleccionar las características que permitieran un mejor desempeño. De esta manera, las 378 configuraciones de **CEGA\*** fueron evaluadas para los seis problemas de prueba descritos en la Sección 2.5, utilizando instancias de  $M = \{5, 10, 15, 20, 30, 50\}$

objetivos y con 31 repeticiones independientes para cada experimento. Por motivos de espacio los resultados de esta extensa experimentación no son mostrados en este documento. Sin embargo, la descripción presentada previamente para **CEGA\*** corresponde a la configuración más prometedora de acuerdo con los resultados obtenidos.

## 5.4 Asignación de aptitud en direcciones múltiples

La estrategia de asignación de aptitud en direcciones múltiples (*Multi-Directional Fitness Assignment*, **MDFA\***), es un mecanismo de asignación de aptitud para guiar el proceso de búsqueda en diferentes direcciones simultáneamente. Para establecer diferentes direcciones de búsqueda se utiliza un conjunto de vectores de pesos. Un vector de pesos es un conjunto de valores o coeficientes que denotan la importancia relativa de cada uno de los objetivos. De este modo, asumimos que diferentes vectores de pesos permitirían dirigir la búsqueda hacia diferentes regiones del espacio de objetivos<sup>4</sup>.

**MDFA\*** se implementó utilizando como base el algoritmo descrito en la Sección 5.2. La figura 5.5 describe el flujo de ejecución del MOEA resultante. El proceso de jerarquización es específicamente nuestra propuesta (este elemento ha sido resaltado con un rectángulo punteado en la figura 5.5) y será descrito a continuación.

### 5.4.1 Jerarquización MDFA\*

La jerarquización se realiza con un doble propósito: por un lado, asignar a cada individuo un valor con el que competirá al momento de seleccionar los que podrán reproducirse. Por otro lado, discriminar entre individuos al momento de seleccionar los que sobrevivirán y formarán la población de padres para la siguiente generación. Considerando una población de tamaño  $N$ , el proceso de jerarquización **MDFA\*** requiere la especificación de un conjunto  $V$  con  $N$  vectores de pesos diferentes ( $|V| = N$ ).

El proceso de jerarquización **MDFA\*** se describe mediante el Algoritmo 4. Todos los individuos

---

<sup>4</sup>El uso de múltiples vectores de pesos para asignar aptitud a un conjunto de soluciones no es una idea nueva, tal es el caso del algoritmo **MSOPS** [36] que se describe más adelante en la Sección 5.5.1.4.

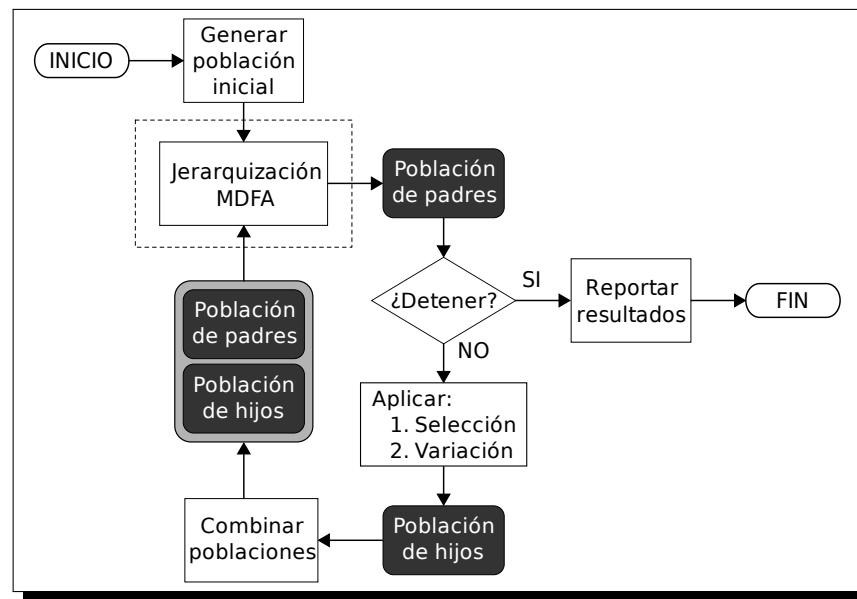


Figura 5.5: MDFA\*: Flujo de ejecución.

se evalúan para cada uno de los vectores de pesos utilizando algún enfoque de agregación. En este trabajo se implementó la suma ponderada (WSUM, ver Sección 3.2.3.2). Para cada vector de pesos se identifica el individuo con el mejor desempeño. Si un individuo fue identificado por tener el mejor desempeño para alguno de los vectores, asignar este desempeño como aptitud para dicho individuo. Dado que un mismo individuo puede ser el mejor para varios vectores de pesos, su aptitud será la mejor de todas las obtenidas con ese conjunto de vectores. Los individuos a los que no se les asignó aptitud (puesto que no mostraron el desempeño más alto para ningún vector), recibirán como aptitud el peor desempeño que obtuvieron para los diferentes vectores. Adicionalmente, la aptitud de estos individuos será penalizada sumándole la peor de las aptitudes asignadas en la primera etapa. Al asignar a estos individuos la peor de sus aptitudes se dará preferencia a los individuos con un mejor desempeño promedio. Penalizarlos con la peor aptitud de la primera etapa permitirá asegurar que estos tendrán menor preferencia (son inferiores) que el primer conjunto de individuos.

De esta manera, MDFA\* mantiene como prioridad el requerimiento de convergencia, al mismo tiempo que promueve dirigir la búsqueda en diferentes direcciones simultáneamente y preserva diversidad en la población.

**Algoritmo 4** Asignación de aptitud MDFA\*.**DEFINE** MDFA()

```

1:  $aptitud[X_i] \leftarrow \infty \quad \forall X_i \in P$ 
2: - Etapla 1:
3: for all  $v \in V$  do
4:   Encontrar  $X_i \in P : WSUM(X_i, v) < WSUM(X_j, v) \quad \forall X_j \in P : X_j \neq X_i$ 
5:   if  $WSUM(X_i, v) < aptitud[X_i]$  then
6:      $aptitud[X_i] \leftarrow WSUM(X_i, v)$ 
7: - Etapla 2:
8:  $peor\_etapa1 \leftarrow \max(aptitud[X_i]) \quad \forall X_i \in P : aptitud[X_i] \neq \infty$ 
9: for all  $X_i \in P : aptitud[X_i] = \infty$  do
10:   $aptitud[X_i] \leftarrow \max_{v \in V}(WSUM(X_i, v)) + peor\_etapa1$ 

```

**END****DEFINE** WSUM( $X_i, v$ )

```

11: return  $\sum_{m=1}^M v_m f_m(X_i)$ 

```

**END****5.4.2 Justificación de ajuste de parámetros y operadores**

Los factores que pueden afectar el comportamiento de MDFA\* se listan a continuación:

1. Estrategia utilizada,  $E$ : Se implementaron 5 estrategias diferentes para este enfoque,  $E = \{MDFA1, MDFA2, \dots, MDFA5\}$ . Estas estrategias difieren principalmente en la forma de utilizar el conjunto de vectores de pesos para asignar la aptitud de los individuos y qué hacer con los individuos que no recibieron aptitud a través de este proceso.
2. Conjunto de vectores de pesos,  $VS$ : Se experimentó con conjuntos vectores de pesos generados de 3 formas distintas,  $VS = \{VS1, VS2, VS3\}$ :
  - a)  $VS1$ : Vectores de pesos generados con una herramienta desarrollada por Hughes para su algoritmo MSOPS [36], disponible para uso académico en [39].
  - b)  $VS2$ : Vectores de pesos generados aleatoriamente en el rango  $[0.5, 1]$ .
  - c)  $VS3$ : Vectores de pesos generados aleatoriamente en el rango  $[0.25, 1]$ .

Estos factores generan un total de  $|E| \times |VS| = 5 \times 3 = 15$  configuraciones para **MDFA\***. Del mismo modo que con **CEGA\***, evaluamos estas 15 configuraciones utilizando los seis problemas de prueba adoptados, con  $M = \{5, 10, 15, 20, 30, 50\}$  objetivos y considerando 31 repeticiones independientes para cada experimento. El procedimiento de jerarquización **MDFA\*** detallado anteriormente corresponde a la estrategia que mostró el mejor desempeño de acuerdo con las estadísticas obtenidas con esta experimentación. El conjunto de vectores de pesos que aportó el mejor comportamiento para el algoritmo es *VS3*, es decir, vectores de pesos generados aleatoriamente en el rango  $[0.25, 1]$ .

## 5.5 Experimentación

El propósito de esta fase de experimentación es estudiar el comportamiento de los algoritmos propuestos en términos de convergencia y diversidad, así como también la escalabilidad de estos enfoques con respecto al número de objetivos. Para esta experimentación se consideraron los diferentes problemas de prueba descritos en la Sección 2.5, utilizando instancias de tamaño  $M = \{5, 10, 15, 20, 30, 50\}$ , donde  $M$  denota el número de objetivos a optimizar.

Para estos experimentos hemos incluido como punto de referencia y comparación 4 MOEAs tomados de la literatura especializada. Algunos de estos MOEAs han sido propuestos específicamente para lidiar con el aumento en el número de objetivos, mientras que otros son optimizadores multiobjetivo representativos que han sido ampliamente utilizados en estudios comparativos. Tanto nuestros algoritmos como los de la literatura han sido evaluados utilizando una población de  $N = 100$  individuos durante 300 generaciones.

En la Sección 5.5.1 se describen de manera introductoria los diferentes enfoques de la literatura considerados. Finalmente, la Sección 5.5.2 presenta los resultados obtenidos con esta experimentación.

### 5.5.1 Algoritmos del estado del arte tomados para realizar la comparativa

#### 5.5.1.1. NSGA-II

El *Elitist Non-Dominated Sorting Genetic Algorithm* (NSGA-II), propuesto por Deb *et al.* en 2000 [17, 22], se ha convertido en el MOEA más representativo de la literatura y ha sido utilizado ampliamente como punto de referencia para comparativas en nuevos desarrollos. Por esta razón, y aunque se ha demostrado que este algoritmo no trabaja bien al incrementarse la cantidad de objetivos [60, 42, 40], hemos considerado a NSGA-II para este estudio comparativo. NSGA-II implementa el *Non-Dominated Sorting* (NDS), que es un método de jerarquización basado en la relación de dominancia de Pareto (ver Sección 3.2.1) y fue propuesto originalmente por Goldberg en 1989 [33]. Una de las características principales de NSGA-II es la incorporación de un mecanismo explícito para preservación de diversidad, al que los autores nombraron *crowding distance* (CWD). CWD estima la densidad de la región del espacio de objetivos donde se encuentra cada solución. Este operador se utiliza como criterio secundario para discriminar entre soluciones con la misma jerarquía, dando preferencia a aquellas ubicadas en regiones menos densas. El Algoritmo 5 describe el flujo de ejecución de NSGA-II. Para mayor detalles del algoritmo referirse a [17, 22].

---

#### Algoritmo 5 NSGA-II: Flujo de ejecución.

---

- 1:  $gen \leftarrow 0$
  - 2: Generar aleatoriamente población inicial de  $N$  padres:  $P_{gen}$
  - 3: Jerarquizar  $P_{gen}$  usando *Non-Dominated Sorting* (ver Sección 3.2.1):  $NDS(P_{gen})$
  - 4: **while**  $gen < GMAX$  (repetir durante  $GMAX$  generaciones) **do**
  - 5: Torneo binario basado en la jerarquía de las soluciones. Si dos soluciones tienen la misma jerarquía, utilizar CWD como criterio de preferencia.
  - 6: Aplicar cruza y mutación para generar la población de  $N$  hijos:  $Q_{gen}$
  - 7: Combinar poblaciones de padres e hijos:  $R_{gen} \leftarrow P_{gen} \cup Q_{gen}$  ( $|R_{gen}| = 2N$ )
  - 8: Jerarquizar  $R_{gen}$  usando *Non-Dominated Sorting* (ver Sección 3.2.1):  $NDS(R_{gen})$
  - 9: Seleccionar  $N$  individuos de  $R_{gen}$  (con respecto a la jerarquía) para formar la población de padres para la siguiente generación,  $P_{gen+1}$ . Si existen más soluciones con la misma jerarquía que la capacidad disponible en  $P_{gen+1}$ , seleccionar sólo  $N - |P_{gen+1}|$  individuos utilizando CWD como criterio de discriminación.
  - 10:  $gen \leftarrow gen + 1$
-

## 5.5.1.2. DMO

Adra y Fleming [1] propusieron en 2009 un operador para administrar el uso de mecanismos de preservación de diversidad dentro de un MOEA. Los autores nombraron a este operador *Diversity Management Operator* (DMO) y lo presentaron como una alternativa para promover diversidad en problemas donde muchos (más de tres) objetivos requieren ser optimizados. DMO es una estrategia adaptativa: los mecanismos de preservación de diversidad son activados únicamente cuando las condiciones de la población así lo requieran. Para conocer el estado de la población (con respecto a la diversidad) en determinada etapa de la búsqueda, DMO utiliza la métrica de *máxima extensión* (*maximum spread*) propuesta por Zitzler en [73]. Esta métrica se define mediante la siguiente ecuación:

$$D = \sqrt{\sum_{m=1}^M \left( \max_{\mathbf{X}_i \in Z} \{f_m(\mathbf{X}_i)\} - \min_{\mathbf{X}_i \in Z} \{f_m(\mathbf{X}_i)\} \right)^2} \quad (5.2)$$

donde  $D$  corresponde a la medida de la diagonal del hipercubo formado por los extremos de  $Z$ , que representa el subconjunto de soluciones no dominadas (con respecto a la dominancia de Pareto) en la población actual. De este modo,  $D$  usa la máxima extensión en cada dimensión del espacio de objetivos para estimar el rango en que las soluciones de  $Z$  se encuentran distribuidas. Una vez que el indicador  $D$  ha sido calculado, éste se normaliza con respecto a un valor óptimo para la métrica,  $D^*$ , sugerido por el tomador de decisiones.  $D^*$  puede calcularse reemplazando  $Z$  por  $P^*$  en la ecuación (5.2), donde  $P^*$  es un conjunto de referencia que representa el FP para el problema a optimizar. El indicador normalizado de diversidad,  $I_s$ , se define entonces de la siguiente manera:

$$I_s = D/D^* \quad (5.3)$$

$I_s$  puede tomar cualquier valor positivo, del que depende la activación o desactivación de los mecanismos de preservación de diversidad. Por un lado,  $I_s = 1 - \varepsilon$  corresponde a una diversidad ideal y  $I_s > 1 - \varepsilon$  indica una dispersión excesiva de las soluciones en el espacio de objetivos, situaciones

en las que los autores proponen desactivar la promoción de diversidad. Por otro lado,  $I_s < 1 - \varepsilon$  se interpreta como escasez de diversidad en la población, por lo que los mecanismos de diversidad son activados.  $\varepsilon$  es un valor opcional de tolerancia, el cual ha sido omitido en este trabajo ( $\varepsilon = 0$ ). En [1] se propone la incorporación de **DMO** en el algoritmo **NSGA-II** descrito en la Sección 5.5.1.1. El enfoque resultante de esta hibridación se describe en el Algoritmo 6.

---

**Algoritmo 6** DMO incorporado en **NSGA-II**: Flujo de ejecución.

---

- 1:  $gen \leftarrow 0$
  - 2: Generar aleatoriamente población inicial de  $N$  padres:  $P_{gen}$
  - 3: Jerarquizar  $P_{gen}$  usando *Non-Dominated Sorting* (ver Sección 3.2.1):  $NDS(P_{gen})$
  - 4: **while**  $gen < GMAX$  (repetir durante  $GMAX$  generaciones) **do**
  - 5:    $Z \leftarrow soluciones\_no\_dominadas(P_{gen})$
  - 6:    $I_s \leftarrow diversidad\_normalizada(Z)$
  - 7:   Torneo binario basado en la jerarquía de las soluciones. Cuando dos soluciones tienen la misma jerarquía: si  $I_s < 1 - \varepsilon$  utilizar CWD como criterio de preferencia, de otro modo elegir aleatoriamente.
  - 8:   Aplicar cruza y mutación para generar la población de hijos:  $Q_{gen}$
  - 9:   Combinar poblaciones de padres e hijos:  $R_{gen} \leftarrow P_{gen} \cup Q_{gen}$  ( $|R_{gen}| = 2N$ )
  - 10:   Jerarquizar  $R_{gen}$  usando *Non-Dominated Sorting* (ver Sección 3.2.1):  $NDS(R_{gen})$
  - 11:   Seleccionar  $N$  individuos de  $R_{gen}$  (con respecto a la jerarquía) para formar la población de padres para la siguiente generación,  $P_{gen+1}$ . Si existen más soluciones con la misma jerarquía que la capacidad disponible en  $P_{gen+1}$ , seleccionar sólo  $N - |P_{gen+1}|$  individuos utilizando CWD como criterio de discriminación si se cumple que  $I_s < 1 - \varepsilon$ , de otro modo hacerlo de manera aleatoria.
  - 12:  $gen \leftarrow gen + 1$
- 

### 5.5.1.3. HypE

Bader y Zitzler [3] propusieron **HypE** (*Hypervolume Estimation Algorithm for Multiobjective Optimization*), un MOEA basado en la métrica de hipervolumen. La métrica de hipervolumen ha sido ampliamente utilizada para evaluar y comparar el desempeño de MOEAs. **HypE** implementa esta métrica como parte de su mecanismo de asignación de aptitud para guiar el proceso de búsqueda. La aptitud de una determinada solución  $\mathbf{X}_i$  está dada por el hipervolumen que se atribuye a  $\mathbf{X}_i$  con respecto al hipervolumen global de todo el conjunto de soluciones. Dado que el costo computacional requerido para calcular el hipervolumen se incrementa exponencialmente con el número de objetivos,



**HypE** aproxima los valores de esta métrica mediante simulación de Monte Carlo. Para **HypE** no es tan importante el valor real de la métrica, sino la jerarquización que puede inducirse mediante este indicador sobre una población de soluciones candidatas. De esta manera, los autores presentan a **HypE** como una alternativa capaz de explotar las ventajas de este indicador y al mismo tiempo escalable con respecto al número de objetivos. El Algoritmo 7 describe el flujo básico de ejecución de **HypE**. Sin embargo, para profundizar en los detalles de la estimación del hipervolumen y el cálculo de la aptitud de las soluciones refiérase a [3].

---

**Algoritmo 7 HypE:** Flujo de ejecución.

---

- 1:  $gen \leftarrow 0$
  - 2: Generar aleatoriamente población inicial de  $N$  padres:  $P_{gen}$
  - 3: Jerarquizar  $P_{gen}$  usando *Non-Dominated Sorting* (ver Sección 3.2.1):  $NDS(P_{gen})$
  - 4: **while**  $gen < GMAX$  (repetir durante GMAX generaciones) **do**
  - 5:   Torneo binario basado en la contribución estimada de las soluciones para la métrica del hipervolumen.
  - 6:   Aplicar cruce y mutación para generar la población de hijos:  $Q_{gen}$
  - 7:   Combinar poblaciones de padres e hijos:  $R_{gen} \leftarrow P_{gen} \cup Q_{gen}$  ( $|R_{gen}| = 2N$ )
  - 8:   Jerarquizar  $R_{gen}$  usando *Non-Dominated Sorting* (ver Sección 3.2.1):  $NDS(R_{gen})$
  - 9:   Formar la población de padres para la siguiente generación,  $P_{gen+1}$ , seleccionando  $N$  individuos de  $R_{gen}$  con base en la jerarquización obtenida en el paso 8. Si existen más soluciones con la misma jerarquía que la capacidad disponible en  $P_{gen+1}$  proceder de la siguiente manera: iterativamente estimar la aptitud con respecto al hipervolumen de dicho conjunto de soluciones y eliminar la solución con el peor desempeño, repitiendo este proceso hasta que solo queden los  $N - |P_{gen+1}|$  individuos necesarios.
  - 10:  $gen \leftarrow gen + 1$
- 

#### 5.5.1.4. MSOPS

En 2003, Hughes [36] propuso el *Multiple Single Objective Pareto Sampling* (**MSOPS**) como una alternativa para optimizar problemas con muchos objetivos. **MSOPS** es una técnica que realiza en paralelo múltiples optimizaciones convencionales basadas en vectores de pesos, lo que permite buscar simultáneamente en diferentes direcciones. **MSOPS** no fue propuesto como un MOEA, sino como un mecanismo de jerarquización de soluciones que puede ser incorporado en una variedad de ellos. Este mecanismo consiste en utilizar  $T$  vectores de pesos y evaluar las soluciones para cada

uno de ellos con alguna técnica de agregación. Una matriz  $S$ , con  $N$  renglones y  $T$  columnas, almacenará el desempeño de cada uno de los  $N$  individuos con respecto a los  $T$  vectores de pesos. Posteriormente, cada columna de  $S$  se jerarquiza de un modo tal que el individuo con el mejor desempeño tendrá jerarquía 1, mientras que la peor solución tendrá jerarquía  $N$ . La jerarquización resultante de las diferentes columnas será almacenada en una matriz  $R$ . Los renglones de  $R$  son ahora ordenados ascendentemente, de manera que en la primera columna se encontrará la mejor jerarquía de cada individuo. Finalmente, a partir de la matriz  $R$  es posible comparar y jerarquizar las soluciones. El Algoritmo 8 ilustra el procedimiento que puede utilizarse para comparar dos soluciones  $X_i$  y  $X_j$  utilizando la matriz  $R$ .

---

**Algoritmo 8** MSOPS: Comparación de dos soluciones utilizando la matriz  $R$ .

---

**DEFINE function** comparación\_MSOPS( $X_i, X_j$ )

**for**  $t = 1$  to  $T$  **do**

**if**  $R[X_i][t] < R[X_j][t]$  **then**

**return**  $X_i$

**else if**  $R[X_j][t] < R[X_i][t]$  **then**

**return**  $X_j$

**return** *iguales*

**END function**

---

En [36] se propone utilizar Min-Max con pesos (*Weighted Min-Max*) en combinación con un enfoque al que Hughes nombró *Vector Angle Distance Scaling* (VADS): *optimización dual* (*dual optimization*). De tal forma se utilizó MSOPS para este trabajo. Utilizar estas dos técnicas en conjunción implica que las matrices  $S$  y  $R$  tendrán  $2T$  columnas, una para cada vector de pesos con cada uno de los dos enfoques. De acuerdo con Hughes [36], el número y calidad de los vectores de pesos son factores que determinarán la capacidad del algoritmo para aproximar la superficie compromiso. Para este trabajo se utilizaron 100 vectores de pesos generados con una herramienta desarrollada por Hughes disponible para uso académico en [39]. Sin embargo, por limitantes de tiempo y por conveniencia se utilizó la versión de MSOPS implementada y disponible en el marco

de trabajo PISA [5].

### 5.5.2 Resultados

Los resultados de esta etapa de experimentación se presentan mediante diagramas de cajas (boxplots), puesto que son gráficas que muestran simultáneamente diferentes medidas descriptivas; v.g., valores que describen ciertos aspectos importantes de un conjunto de datos (en nuestro caso el conjunto de datos corresponde a las diferentes repeticiones de cada experimento). La interpretación de los diagramas utilizados se muestra en la figura 5.6. Aunque todos los datos mostrados en estas

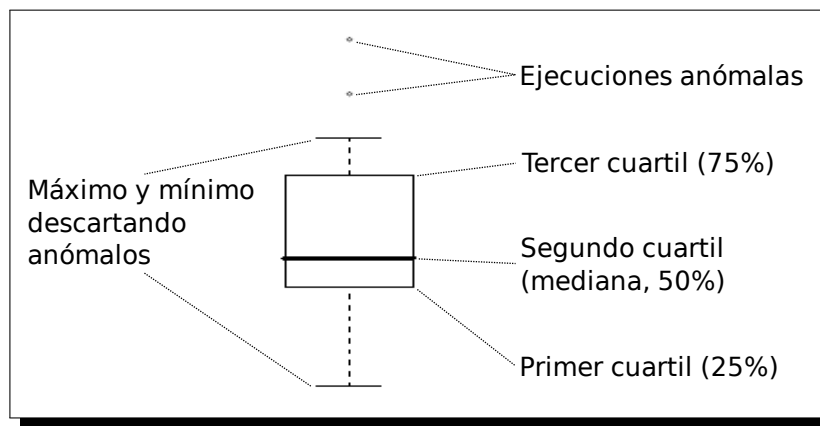


Figura 5.6: Interpretación de los diagramas de cajas (boxplots).

gráficas son importantes, de particular interés son los límites de las cajas y la barra dentro de éstas, siendo los principales puntos de referencia para comparaciones. Los límites inferior y superior de las cajas representan el primer y tercer cuartil respectivamente. La longitud de la caja es, por lo tanto, el rango intercuartil, de manera que cuanto más grande sea la caja, mayor será la dispersión de los datos. De este modo, la caja contendrá el 50 % central de los valores de un grupo. La barra horizontal que aparece dentro de la caja identifica la mediana del conjunto de valores (segundo cuartil), siendo una medida de tendencia central que, a diferencia de la media, no es tan susceptible a valores anómalos (atípicos).

#### 5.5.2.1. Convergencia

El propósito de este experimento es investigar la habilidad de los diferentes algoritmos estudiados para converger en espacios de objetivos con alta dimensionalidad. Los resultados obtenidos por los diferentes enfoques fueron evaluados utilizando la métrica de convergencia descrita en la Sección 2.6.1.

Las figuras 5.7 a 5.12 muestran los resultados obtenidos para los problemas DTLZ1 a DTLZ6 respectivamente. Los datos graficados corresponden a 31 corridas independientes para cada configuración del experimento. La métrica de convergencia es un criterio de minimización, por lo que mientras mas bajo sea el valor para la métrica (menor altura de las cajas) mejor será la convergencia lograda por un algoritmo.

De manera general, está claro que nuestros algoritmos **CEGA\*** y **MDFA\*** mostraron el mejor desempeño para este experimento, puesto que en todas las instancias de los diferentes problemas de prueba alcanzaron los niveles más bajos para la métrica de convergencia, superando significativamente a los 4 enfoques del estado del arte elegidos para comparación. Comparando específicamente nuestros enfoques, consideramos que en términos de convergencia **CEGA\*** es superior a **MDFA\***, ya que en la mayoría de los casos obtuvo mejores resultados. Únicamente en los problemas DTLZ3 y DTLZ6 (figuras 5.9 y 5.12, respectivamente) las diferencias no son tan significativas como para considerar que alguno de estos dos enfoques es superior a otro.

Al analizar el comportamiento de nuestros enfoques en los diferentes casos de prueba se puede observar que el desempeño de **CEGA\*** mejora al incrementarse la cantidad de objetivos. Este comportamiento se presenta en el caso de los problemas DTLZ1, DTLZ2, DTLZ4 y DTLZ5. Para DTLZ3 sólo se cumple en las tres instancias más pequeñas, mientras que para el problema DTLZ6 el desempeño de este enfoque es similar en todas las instancias. Este comportamiento de **CEGA\*** nos permite establecer tendencias en su escalabilidad con respecto al número de objetivos y su robustez ante diferentes problemas de prueba. En términos de convergencia es posible argumentar que, al menos para los problemas de prueba utilizados, incrementar la cantidad de objetivos no tuvo efectos negativos sobre **CEGA\***. Por otro lado, en la mayoría de los casos el desempeño de **MDFA\***

empeora ligeramente de forma gradual al incrementarse la cantidad de objetivos, lo que nos indica que aunque su habilidad de convergencia no es mala, es un enfoque más susceptible a este factor al compararse con **CEGA\***.

**NSGA-II** fue inferior a **DMO** en la mayoría de las instancias de este experimento. Estos resultados confirman que el uso de mecanismos para promover diversidad puede resultar nocivo en escenarios con muchos objetivos. Sin embargo, la mejora de **DMO** sobre **NSGA-II** no es tan significativa como para alcanzar buenos niveles de convergencia, ya que sigue siendo un método que usa como criterio de convergencia la relación de dominancia de Pareto que, como ya se ha discutido previamente, pierde sus efectos de discriminación al incrementarse la cantidad de objetivos provocando que el proceso de búsqueda sea guiado prácticamente de manera aleatoria. Por otra parte, de los cuatro algoritmos del estado del arte elegidos para la comparación, **MSOPS** fue el enfoque que mostró una mejor convergencia en la mayoría de los casos.

El rango de los resultados mostrados en cada figura confirma que los problemas DTLZ1, DTLZ3 y DTLZ6 involucran mayores dificultades de convergencia que el resto de los casos de prueba, tal como la definición de estos problemas lo sugiere (ver Sección 2.5). Para estos tres problemas la diferencia entre la convergencia lograda por nuestros enfoques y los algoritmos de la literatura es notablemente significativa, de tal forma se puede destacar la robustez y habilidad de convergencia de **CEGA\*** y **MDFA\***.

Figura 5.7: Métrica de convergencia. Problema DTLZ1.

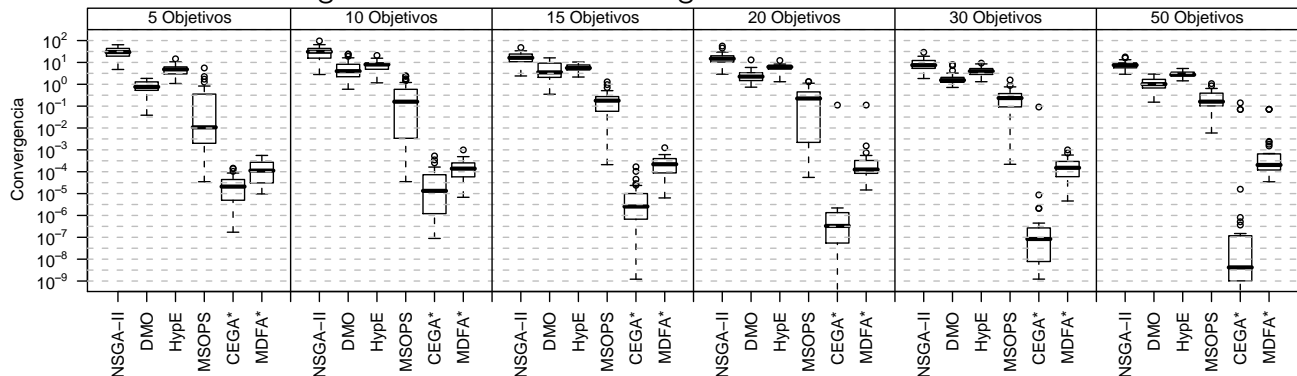


Figura 5.8: Métrica de convergencia. Problema DTLZ2.

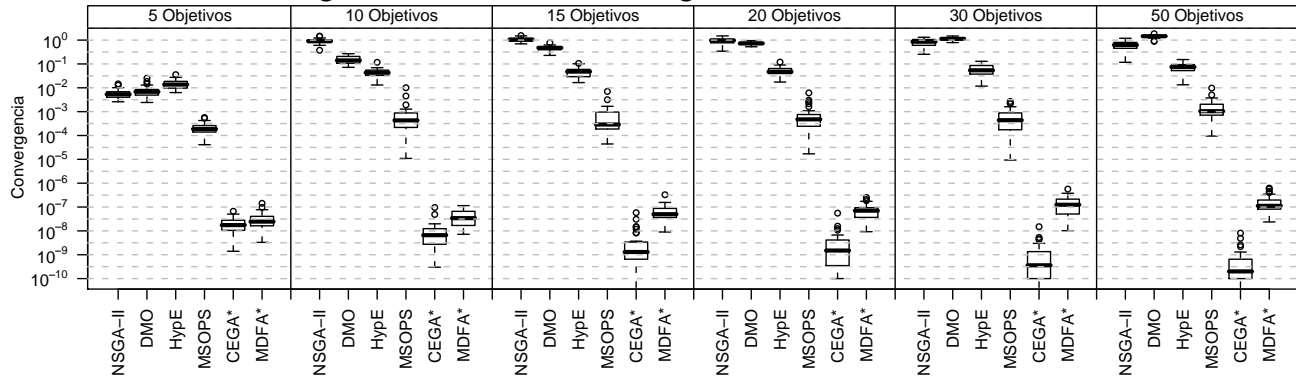


Figura 5.9: Métrica de convergencia. Problema DTLZ3.

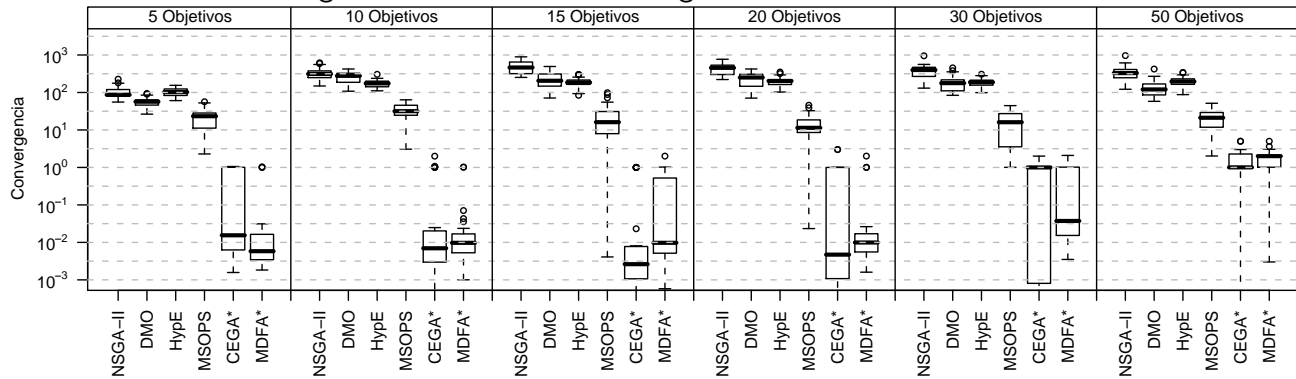


Figura 5.10: Métrica de convergencia. Problema DTLZ4.

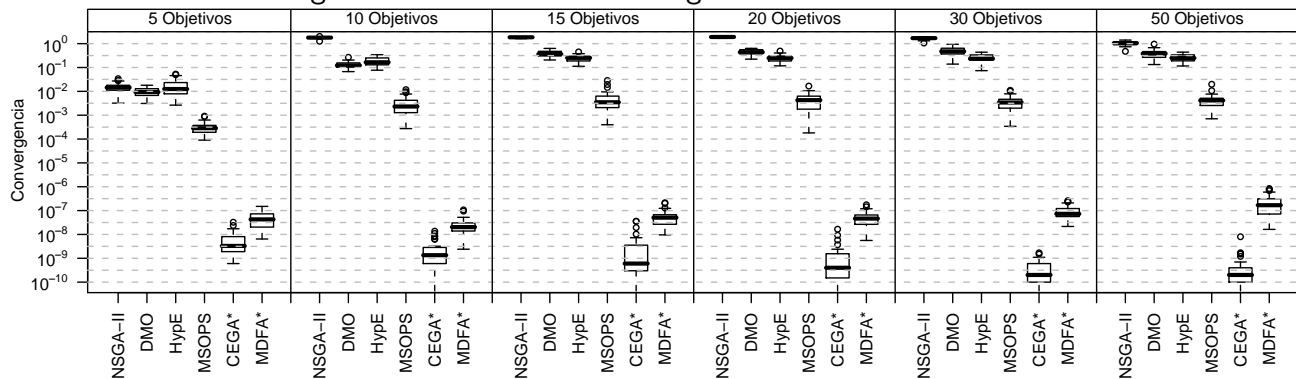


Figura 5.11: Métrica de convergencia. Problema DTLZ5.

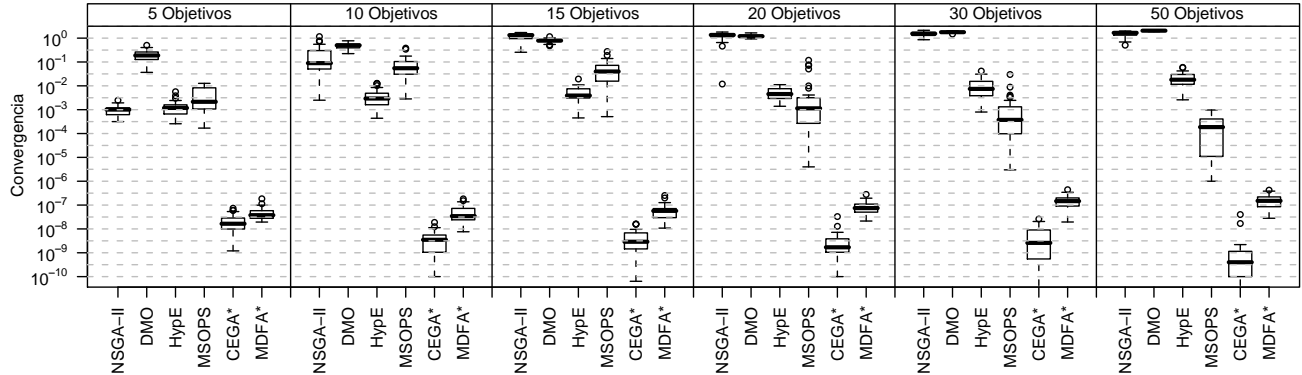
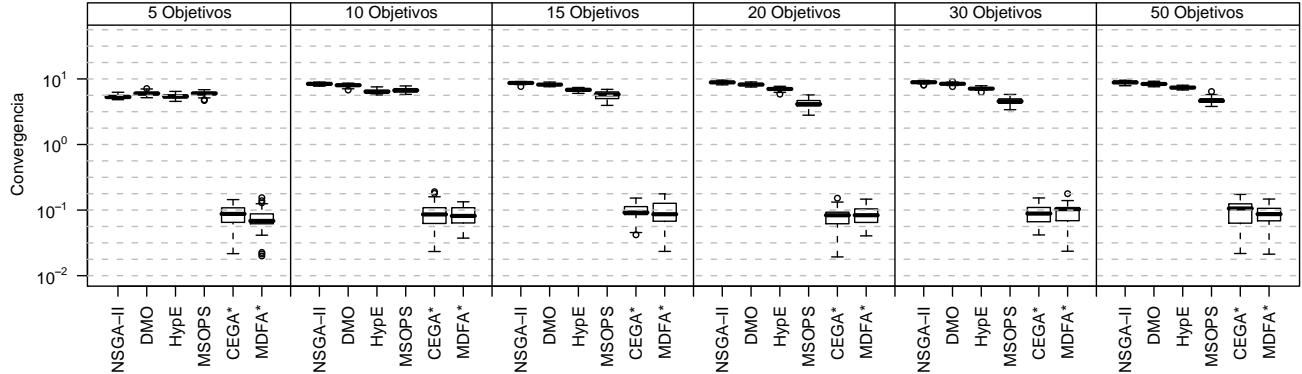


Figura 5.12: Métrica de convergencia. Problema DTLZ6.



### 5.5.2.2. Distancia generacional invertida

La distancia generacional invertida (DGI) es una métrica que permite evaluar simultáneamente la habilidad de convergencia de los MOEAs y la capacidad que estos tienen para encontrar un conjunto diverso de soluciones. DGI es un criterio de minimización, de manera que cuánto más bajo sea el valor para esta métrica, mejor será el desempeño del MOEA evaluado. Para mayores detalles de esta métrica referirse a la Sección 2.6.2.

Las figuras 5.13 a 5.18 muestran los resultados de este experimento para los problemas DTLZ1 a DTLZ6, respectivamente. Los datos en estas figuras corresponden a 31 ejecuciones independientes para cada instancia del experimento, y la interpretación de estas figuras es aquella descrita al inicio de la Sección 5.5 mediante la figura 5.6.

Para este experimento podemos distinguir principalmente dos conjuntos de problemas en los que pueden agruparse los comportamientos observados. Por una parte están los problemas DTLZ1, DTLZ3 y DTLZ6, en los que nuestros algoritmos **CEGA\*** y **MDFA\*** mostraron superioridad para sus diferentes instancias. En el caso del problema DTLZ1 los resultados de **MSOPS** son muy cercanos a los de nuestros enfoques. Sin embargo, la amplitud de su correspondiente caja y las ejecuciones atípicas para las diferentes instancias de este problema son factores que nos indican que presenta un desempeño más inconsistente. Contrastando estas observaciones con los resultados de la Sección 5.5.2.1 es posible remarcar que, aunque **MSOPS** no alcanza los mejores niveles de convergencia, logra obtener un conjunto de soluciones suficientemente distribuido como para mostrar resultados competitivos en este experimento. Por otro lado, en el caso de los problemas DTLZ3 y DTLZ6 la superioridad de nuestros enfoques fue más evidente. Estos tres problemas, de acuerdo con su definición (ver Sección 2.5) y con los resultados de la Sección 5.5.2.1, son los que involucran mayores dificultades de convergencia.

El segundo conjunto de observaciones corresponde a los problemas DTLZ2, DTLZ4 y DTLZ5, en los que se presenta un comportamiento parecido para los distintos algoritmos. En general podemos considerar que **HypE**, **MSOPS**, y nuestras propuestas **CEGA\*** y **MDFA\*** son los cuatro algoritmos que mejor se desempeñaron para estos problemas de prueba. En el caso del problema DTLZ2, el peor desempeño para la instancia más pequeña ( $M = 5$ ) se le atribuye a nuestro algoritmo **MDFA\***, pero mejoró gradualmente al incrementarse el número de objetivos. Por el contrario, **MSOPS** comenzó mostrando el mejor desempeño, pero esta diferencia se redujo para instancias más grandes del problema. A partir de  $M = 15$  los algoritmos **HypE**, **MSOPS**, **CEGA\*** y **MDFA\*** se comportaron de manera similar obteniendo los mejores resultados. Para el problema DTLZ4, en la instancia más pequeña **CEGA\*** y **MDFA\*** obtuvieron los peores resultados, mientras que **MSOPS** nuevamente mostró el mejor desempeño. Sin embargo, con el incremento en las dimensiones del problema estas diferencias se aminoraron y a partir de  $M = 15$  la diferencia entre **HypE**, **MSOPS**, **CEGA\*** y **MDFA\*** fueron mínimas. Finalmente, en todas las instancias del problema DTLZ5 los algoritmos **HypE**, **MSOPS** y **CEGA\*** mostraron el mejor desempeño, mientras que nuevamente **MDFA\*** comenzó con los peores resultados pero para instancias más grandes se equiparó con este



grupo de algoritmos.

En nuestra opinión, para este segundo conjunto de casos de prueba las diferencias entre el desempeño de los seis algoritmos evaluados no son tan significativas como para establecer aseveraciones fuertes sobre la superioridad de alguno de ellos. Sin embargo, es preciso señalar que en las diferentes instancias de estos tres problemas, nuestros enfoques **CEGA\*** y **MDFA\*** fueron igualados o superados por algunos de los algoritmos del estado del arte. Estos resultados nos indican que **CEGA\*** y **MDFA\*** son inferiores en cuanto a la habilidad para mantener un conjunto diversificado de soluciones. Aunque los diferentes algoritmos del estado del arte no alcanzaron los mejores niveles de convergencia, su buena capacidad para mantener diversidad en la población y los paisajes de aptitud (notablemente menos complicados) de estos tres problemas de prueba son factores que les permitieron mostrar resultados competitivos para este experimento.

Figura 5.13: Métrica DGI. Problema DTLZ1.

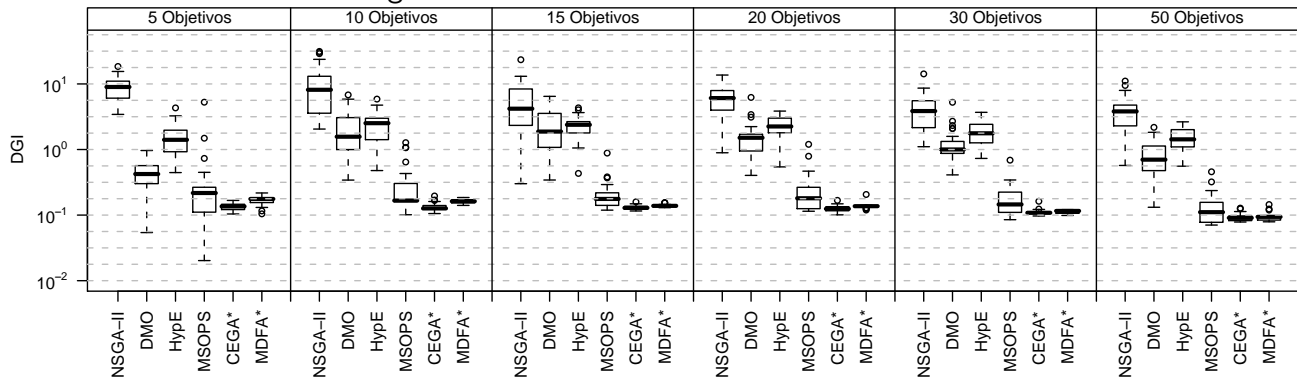


Figura 5.14: Métrica DGI. Problema DTLZ2.

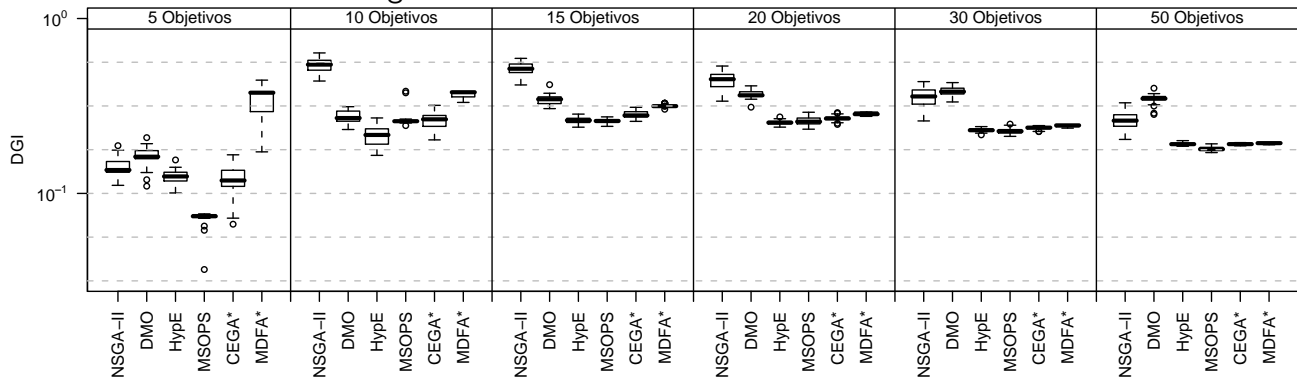


Figura 5.15: Métrica DGI. Problema DTLZ3.

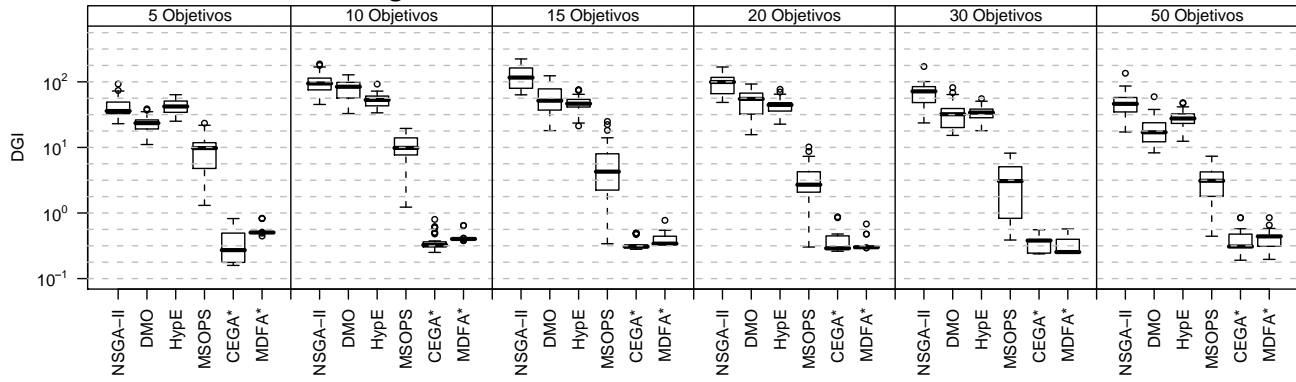


Figura 5.16: Métrica DGI. Problema DTLZ4.

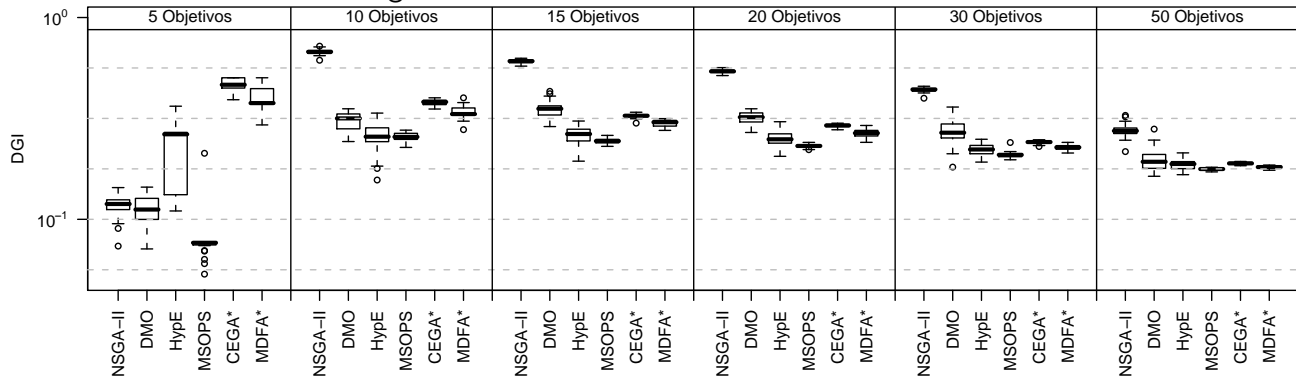
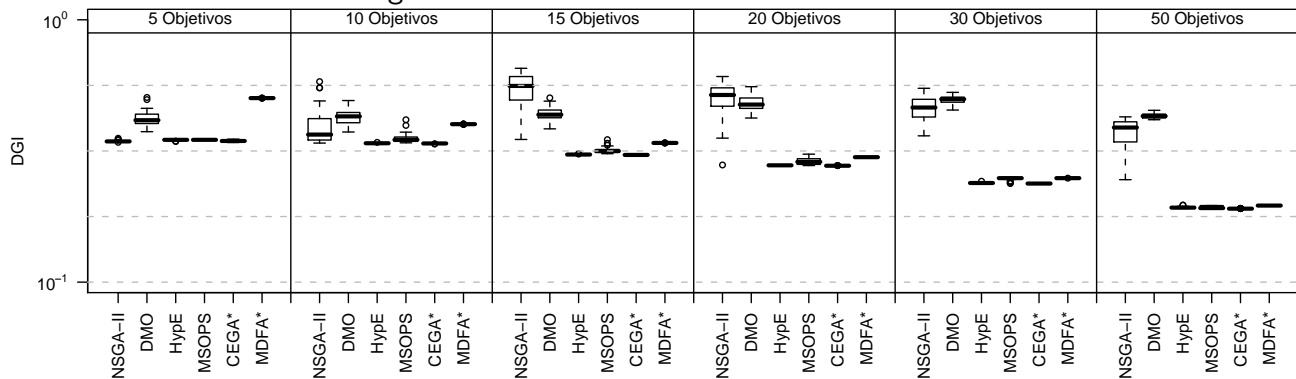
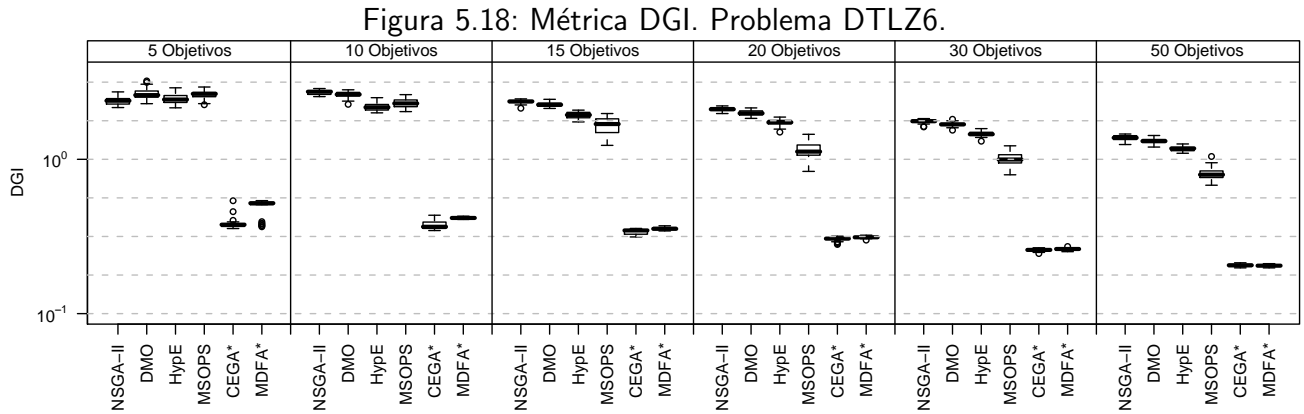


Figura 5.17: Métrica DGI. Problema DTLZ5.





## 5.6 Conclusiones del capítulo

En este capítulo se presentaron dos nuevos MOEAs: **CEGA\*** y **MDFA\***. Estos enfoques se propusieron como alternativas capaces de converger en espacios de objetivos con alta dimensionalidad, y que al mismo tiempo incorporen mecanismos que promuevan la existencia de diversidad en la población sin afectar negativamente la convergencia.

En este capítulo también se presentaron los resultados de un estudio comparativo donde se evaluaron los dos nuevos MOEAs y cuatro enfoques representativos de la literatura. Para estos experimentos se consideró un conjunto de seis problemas de prueba escalables con instancias que involucran  $M = \{5, 10, 15, 20, 30, 50\}$  objetivos.

Los resultados de este estudio indican que en términos de convergencia los dos enfoques propuestos son significativamente superiores a los cuatro MOEAs tomados de la literatura. El éxito de nuestros enfoques en términos de convergencia es principalmente el resultado de las estrategias de discriminación adoptadas. Comparando específicamente nuestras propuestas, **CEGA\*** mostró un mejor desempeño que **MDFA\*** en la mayoría de los casos de prueba. Una observación importante acerca de **CEGA\*** es que el incremento en la cantidad de objetivos no tiene efectos negativos sobre este enfoque, ya incluso sus resultados mejoraron cuando se realizó un incremento en la dimensión de la mayoría de los problemas de prueba considerados.

Por otra parte, aún cuando nuestras propuestas demostraron ser superiores en términos de con-

vergencia, los resultados indican que los enfoques del estado del arte proveen una mejor preservación de diversidad. Aunque los enfoques de la literatura no lograron la mejor convergencia, la diversidad de su conjunto aproximación fue suficiente para lograr resultados competitivos de la métrica utilizada. A pesar de que nuestros enfoques incorporan mecanismos para promover la diversidad, es notorio que no son capaces de converger de una manera suficientemente distribuida.

Estos resultados hacen clara la necesidad de desarrollar mecanismos más efectivos para preservación de diversidad y que puedan trabajar en conjunto con una jerarquización estricta de las soluciones, de tal manera que se logre un balance adecuado entre los requerimientos de convergencia y diversidad.

# 6

## Conclusiones y trabajo futuro

### 6.1 Conclusiones

La dominancia de Pareto (PD) ha sido la principal base para el diseño de algoritmos evolutivos habilitados para optimizar problemas con 2 o 3 objetivos. No obstante, se ha demostrado en el pasado y en este trabajo que PD pierde su eficacia ante el incremento de la cantidad de objetivos, y como consecuencia, el desempeño de tales enfoques se deteriora significativamente. Esta realidad hace necesario identificar métodos alternativos a PD para establecer preferencias entre soluciones en el contexto multiobjetivo.

En este trabajo se estudiaron 35 métodos de asignación de aptitud para optimización multiobjetivo, de los cuales 22 fueron tomados de la literatura especializada y los 13 restantes son nuevas propuestas resultantes de este proyecto. En dicho estudio se realizaron experimentos para analizar la habilidad de búsqueda de un determinado MOEA al incorporar cada una de las 35 alternativas estudiadas. Para estos experimentos se consideraron 6 problemas de prueba escalables utilizando 6 cantidades de objetivos:  $M = \{5, 10, 15, 20, 30, 50\}$ . Las principales conclusiones a partir de este estudio son las siguientes:

- La mayoría de los métodos propuestos demostraron ser efectivos al guiar el proceso de búsqueda en los diferentes escenarios de optimización considerados.
- Es indispensable ejercer una fuerte presión de selección mediante una jerarquización estricta de las soluciones candidatas para lograr una convergencia robusta.
- Para que un método de jerarquización sea efectivo, se debe considerar la magnitud de la diferencia en cada objetivo al momento de comparar un par de soluciones; descartar esta información puede llevar a malas decisiones en la discriminación y consecuentemente perjudicar la habilidad de búsqueda del MOEA.
- Incrementar la presión de selección tiende a sacrificar la diversidad y por lo general conlleva a converger hacia un único punto de la superficie compromiso. Los requerimientos de convergencia y diversidad son dos criterios que se encuentran en conflicto, por lo que un reto claro de investigación es desarrollar MOEAs que mantengan un balance adecuado entre ellos.

Basándonos en las observaciones y resultados obtenidos con dicho estudio, como parte de este proyecto de tesis también se propusieron dos nuevos MOEAs: **CEGA\*** y **MDFA\***. La motivación detrás de estos dos nuevos enfoques fue proveer alternativas capaces de converger robustamente en espacios de objetivos con alta dimensionalidad e incorporar mecanismos apropiados para mantener la diversidad en la población sin perjudicar la convergencia. Un reto claro al diseñar estos enfoques fue combinar una estrategia de jerarquización estricta con algún mecanismo de preservación de diversidad, ya que con una máxima discriminación no tendría ningún efecto preservar diversidad como criterio secundario para establecer preferencias entre soluciones con la misma jerarquía (como comúnmente se hace). El desempeño de estos MOEAs se comparó contra cuatro MOEAs representativos del estado del arte. Las principales conclusiones a partir de este estudio comparativo son las siguientes:

- Las estadísticas mostraron que en términos de convergencia los dos enfoques propuestos son significativamente superiores a los cuatro MOEAs tomados de la literatura especializada.
- **CEGA\*** mostró un mejor desempeño que **MDFA\*** en la mayoría de los casos de prueba.

- De acuerdo con los resultados obtenidos, el incremento en la cantidad de objetivos no tiene efectos negativos sobre CEGA\* (al menos para los casos de prueba considerados).
- Aún cuando nuestros enfoques demostraron ser superiores en términos de convergencia, los resultados indican que no fueron capaces de converger de una manera suficientemente distribuida; las aproximaciones obtenidas por estos enfoques incluyen relativamente pocos puntos diferentes.

Como conclusión final, en este trabajo se demostró que mediante una discriminación adecuada es posible converger de una manera robusta ante el incremento del número de criterios de optimización. Sin embargo, es necesario desarrollar mecanismos de preservación de diversidad más efectivos, que puedan combinarse con una jerarquización estricta y que permitan lograr aproximaciones representativas de la superficie compromiso.

## 6.2 Contribuciones

Las principales contribuciones de este trabajo de tesis son las siguientes:

- Recopilación y documentación de 22 métodos diferentes de asignación de aptitud para optimización multiobjetivo que han sido reportados en la literatura especializada.
- 13 nuevos métodos de asignación de aptitud para optimización multiobjetivo propuestos.
- Un estudio comparativo que contempla una serie de experimentos para evaluar la capacidad de cada una de las 35 estrategias de asignación de aptitud estudiadas para guiar el proceso de búsqueda, así como la escalabilidad de estas alternativas con respecto al número de objetivos. De acuerdo con la revisión de la literatura realizada, a la fecha, éste es el estudio que contempla más métodos de jerarquización.
- Dos nuevos MOEAs que de acuerdo con la experimentación realizada son capaces de converger robustamente en espacios de objetivos con alta dimensionalidad.

- Un estudio comparativo donde se incluyen los dos nuevos MOEAs y cuatro enfoques representativos del estado del arte. El propósito de dicho estudio fue validar la capacidad de los diferentes MOEAs considerados para optimizar diferentes problemas de prueba y su comportamiento en relación con el aumento en el número de criterios de optimización.
- Algunos resultados derivados de este trabajo de tesis son:
  - Mario Garza Fabre, Gregorio Toscano Pulido and Carlos A. Coello Coello. *Alternative Fitness Assignment Methods for Many-Objective Optimization Problems*, in 9th International Conference on Artificial Evolution (EA 2009), Springer, Strasbourg, France, October 2009.
  - Mario Garza Fabre, Gregorio Toscano Pulido and Carlos A. Coello Coello. *Ranking Methods for Many-Objective Optimization*, in 8th Mexican International Conference on Artificial Intelligence (MICA 2009), Springer, Guanajuato, Gto., México, November 2009.

## 6.3 Trabajo futuro

Como trabajo futuro se pueden distinguir ciertas rutas importantes:

- Desarrollar mecanismos de preservación de diversidad que puedan combinarse con una estrategia de jerarquización estricta y que permitan promover más efectivamente la diversidad sin deteriorar la convergencia. Por ejemplo, en [2], Aguirre y Tanaka demostraron que el simple hecho de prevenir la existencia de individuos duplicados en la población puede mejorar notablemente el desempeño de un MOEA. Esta idea podría ser directamente incorporada en cualquiera de nuestros dos enfoques.
- Para generalizar más rigurosamente las observaciones es necesario considerar más casos de prueba con características diferentes. Así mismo, sería interesante aplicar el conocimiento obtenido a problemas del mundo real.



- Sería de gran importancia incrementar el conjunto de métricas utilizadas para evaluar los resultados de los diferentes experimentos.
- Incluir pruebas estadísticas en la experimentación para realizar mejores inferencias a partir de los resultados. Por ejemplo, Mann-Whitney, Wilcoxon, Kruskal-Wallis, Friedman, entre otras [44].
- Esta tesis utilizó como heurística base un algoritmo genético elitista. Sin embargo, otras heurísticas, tales como *Evolución Diferencial*, *Optimización mediante cúmulos de Partículas* (PSO, por sus siglas en Inglés), podrían ser exploradas con resultados diferentes.
- Algunas de las estrategias de jerarquización estudiadas dependen de un adecuado ajuste de parámetros. En este trabajo se utilizaron los valores que se consideraron más convenientes de acuerdo a lo reportado en cada una de sus respectivas referencias bibliográficas, o siguiendo recomendaciones directas de las autores. Sin embargo, sería de gran importancia estudiar cómo afecta el ajuste de parámetros a cada alternativa con la finalidad de identificar los valores más apropiados.



# Bibliografía

- [1] Salem F. Adra and Peter F. Fleming. A Diversity Management Operator for Evolutionary Many-Objective Optimisation. In Matthias Ehrgott, Carlos M. Fonseca, Xavier Gandibleux, Jin-Kao Hao, and Marc Sevaux, editors, *Evolutionary Multi-Criterion Optimization. 5th International Conference, EMO 2009*, pages 81–94. Springer. Lecture Notes in Computer Science Vol. 5467, Nantes, France, April 2009.
- [2] Hernán Aguirre and Kiyoshi Tanaka. Selection, Drift, Recombination, and Mutation in Multi-objective Evolutionary Algorithms on Scalable MNK-Landscapes. In Carlos A. Coello Coello, Arturo Hernández Aguirre, and Eckart Zitzler, editors, *Evolutionary Multi-Criterion Optimization. Third International Conference, EMO 2005*, pages 355–369, Guanajuato, México, March 2005. Springer. Lecture Notes in Computer Science Vol. 3410.
- [3] J. Bader and E. Zitzler. HypE: An Algorithm for Fast Hypervolume-Based Many-Objective Optimization. TIK Report 286, Computer Engineering and Networks Laboratory (TIK), ETH Zurich, November 2008.
- [4] P. J. Bentley and J. P. Wakefield. Finding Acceptable Solutions in the Pareto-Optimal Range using Multiobjective Genetic Algorithms. In P. K. Chawdhry, R. Roy, and R. K. Pant, editors, *Soft Computing in Engineering Design and Manufacturing*, Part 5, pages 231–240, London, June 1997. Springer Verlag London Limited. (Presented at the 2nd On-line World Conference on Soft Computing in Design and Manufacturing (WSC2)).
- [5] Stefan Bleuler, Marco Laumanns, Lothar Thiele, and Eckart Zitzler. PISA—A Platform and Programming Language Independent Interface for Search Algorithms. In Carlos M. Fonseca, Peter J. Fleming, Eckart Zitzler, Kalyanmoy Deb, and Lothar Thiele, editors, *Evolutionary Multi-Criterion Optimization. Second International Conference, EMO 2003*, pages 494–508, Faro, Portugal, April 2003. Springer. Lecture Notes in Computer Science. Volume 2632.

- [6] D. Brockhoff and E. Zitzler. Offline and Online Objective Reduction in Evolutionary Multi-objective Optimization Based on Objective Conflicts. TIK Report 269, Institut für Technische Informatik und Kommunikationsnetze, ETH Zürich, April 2007.
- [7] Dima Brockhoff and Eckart Zitzler. Are All Objectives Necessary? On Dimensionality Reduction in Evolutionary Multiobjective Optimization. In Thomas Philip Runarsson, Hans-Georg Beyer, Edmund Burke, Juan J. Merelo-Guervós, L. Darrell Whitley, and Xin Yao, editors, *Parallel Problem Solving from Nature - PPSN IX, 9th International Conference*, pages 533–542. Springer. Lecture Notes in Computer Science Vol. 4193, Reykjavik, Iceland, September 2006.
- [8] Dima Brockhoff and Eckart Zitzler. Dimensionality Reduction in Multiobjective Optimization: The Minimum Objective Subset Problem. In Karl Heinz Waldmann and Ulrike M. Stocker, editors, *Operations Research Proceedings 2006*, pages 423–429, Saarbücken, Germany, 2007. Springer.
- [9] E.K. Burke and J.D. Landa Silva. The influence of the fitness evaluation method on the performance of multiobjective search algorithms. *European Journal of Operational Research*, 169(3):875–897, March 2006.
- [10] Carlos A. Coello Coello. 20 Years of Evolutionary Multi-Objective Optimization: What Has Been Done and What Remains to be Done. In Gary Y. Yen and David B. Fogel, editors, *Computational Intelligence: Principles and Practice*, chapter 4, pages 73–88. IEEE Computational Intelligence Society, Vancouver, Canada, 2006, ISBN 0-9787135-0-8.
- [11] Carlos A. Coello Coello, Gary B. Lamont, and David A. Van Veldhuizen. *Evolutionary Algorithms for Solving Multi-Objective Problems*. Springer, New York, second edition, September 2007. ISBN 978-0-387-33254-3.
- [12] David Corne and Joshua Knowles. Techniques for Highly Multiobjective Optimisation: Some Nondominated Points are Better than Others. In Dirk Thierens, editor, *2007 Genetic and*

- Evolutionary Computation Conference (GECCO'2007)*, volume 1, pages 773–780, London, UK, July 2007. ACM Press.
- [13] David W. Corne, Nick R. Jerram, Joshua D. Knowles, and Martin J. Oates. PESA-II: Region-based Selection in Evolutionary Multiobjective Optimization. In Lee Spector, Erik D. Goodman, Annie Wu, W.B. Langdon, Hans-Michael Voigt, Mitsuo Gen, Sandip Sen, Marco Dorigo, Shahram Pezeshk, Max H. Garzon, and Edmund Burke, editors, *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO'2001)*, pages 283–290, San Francisco, California, 2001. Morgan Kaufmann Publishers.
- [14] David W. Corne, Joshua D. Knowles, and Martin J. Oates. The Pareto Envelope-based Selection Algorithm for Multiobjective Optimization. In Marc Schoenauer, Kalyanmoy Deb, Günter Rudolph, Xin Yao, Evelyne Lutton, Juan Julian Merelo, and Hans-Paul Schwefel, editors, *Proceedings of the Parallel Problem Solving from Nature VI Conference*, pages 839–848, Paris, France, 2000. Springer. Lecture Notes in Computer Science No. 1917.
- [15] D.B. Das and C. Patvardhan. New multi-objective stochastic search technique for economic load dispatch. *IEE Proceedings on Generation, Transmission and Distribution*, 145(6):747–752, November 1998.
- [16] Kalyanmoy Deb. *Multi-Objective Optimization using Evolutionary Algorithms*. John Wiley & Sons, Chichester, UK, 2001. ISBN 0-471-87339-X.
- [17] Kalyanmoy Deb, Samir Agrawal, Amrit Pratab, and T. Meyarivan. A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-Objective Optimization: NSGA-II. In Marc Schoenauer, Kalyanmoy Deb, Günter Rudolph, Xin Yao, Evelyne Lutton, Juan Julian Merelo, and Hans-Paul Schwefel, editors, *Proceedings of the Parallel Problem Solving from Nature VI Conference*, pages 849–858, Paris, France, 2000. Springer. Lecture Notes in Computer Science No. 1917.

- [18] Kalyanmoy Deb, Samir Agrawal, Amrit Pratab, and T. Meyarivan. A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-Objective Optimization: NSGA-II. KanGAL report 200001, Indian Institute of Technology, Kanpur, India, 2000.
- [19] Kalyanmoy Deb and Mayank Goyal. A Combined Genetic Adaptive Search (GeneAS) for Engineering Design. *Computer Science and Informatics*, 26:30–45, 1996.
- [20] Kalyanmoy Deb and Sachin Jain. Running Performance Metrics for Evolutionary Multi-Objective Optimization. In Lipo Wang, Kay Chen Tan, Takeshi Furuhashi, Jong-Hwan Kim, and Xin Yao, editors, *Proceedings of the 4th Asia-Pacific Conference on Simulated Evolution and Learning (SEAL '02)*, volume 1, pages 13–20, Orchid Country Club, Singapore, November 2002. Nanyang Technical University.
- [21] Kalyanmoy Deb and Amarendra Kumar. Real-coded Genetic Algorithms with Simulated Binary Crossover: Studies on Multimodal and Multiobjective Problems. *Complex Systems*, 9:431–454, 1995.
- [22] Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and T. Meyarivan. A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, April 2002.
- [23] Kalyanmoy Deb and Dhish Kumar Saxena. Searching for Pareto-optimal solutions through dimensionality reduction for certain large-dimensional multi-objective optimization problems. In *2006 IEEE Congress on Evolutionary Computation (CEC'2006)*, pages 3353–3360, Vancouver, BC, Canada, July 2006. IEEE.
- [24] Kalyanmoy Deb, Lothar Thiele, Marco Laumanns, and Eckart Zitzler. Scalable Multi-Objective Optimization Test Problems. In *Congress on Evolutionary Computation (CEC'2002)*, volume 1, pages 825–830, Piscataway, New Jersey, May 2002. IEEE Service Center.
- [25] Kalyanmoy Deb, Lothar Thiele, Marco Laumanns, and Eckart Zitzler. Scalable Test Problems for Evolutionary Multiobjective Optimization. In Ajith Abraham, Lakhmi Jain, and Robert Goldberg,

- editors, *Evolutionary Multiobjective Optimization. Theoretical Advances and Applications*, pages 105–145. Springer, USA, 2005.
- [26] Francesco di Pierro, Shoon-Thiam Khu, and Dragan A. Savić. An Investigation on Preference Order Ranking Scheme for Multiobjective Evolutionary Optimization. *IEEE Transactions on Evolutionary Computation*, 11(1):17–45, February 2007.
- [27] Nicole Drechsler, Rolf Drechsler, and Bernd Becker. Multi-objective Optimisation Based on Relation *favour*. In Eckart Zitzler, Kalyanmoy Deb, Lothar Thiele, Carlos A. Coello Coello, and David Corne, editors, *First International Conference on Evolutionary Multi-Criterion Optimization*, pages 154–166. Springer-Verlag. Lecture Notes in Computer Science No. 1993, 2001.
- [28] M. Farina and P. Amato. On the Optimal Solution Definition for Many-criteria Optimization Problems. In *Proceedings of the NAFIPS-FLINT International Conference'2002*, pages 233–238, Piscataway, New Jersey, June 2002. IEEE Service Center.
- [29] M. Farina and P. Amato. A fuzzy definition of “optimality” for many-criteria optimization problems. *IEEE Transactions on Systems, Man, and Cybernetics Part A—Systems and Humans*, 34(3):315–326, May 2004.
- [30] L. J. Fogel, A. J. Owens, and M. J. Walsh. *Artificial Intelligence through Simulated Evolution*. John Wiley, New York, USA, 1966.
- [31] Carlos M. Fonseca and Peter J. Fleming. Genetic Algorithms for Multiobjective Optimization: Formulation, Discussion and Generalization. In Stephanie Forrest, editor, *Proceedings of the Fifth International Conference on Genetic Algorithms*, pages 416–423, San Mateo, California, 1993. University of Illinois at Urbana-Champaign, Morgan Kauffman Publishers.
- [32] Rafael Giusti, Gustavo E. A. P. A. Batista, and Ronaldo Cristiano Prati. Evaluating ranking composition methods for multi-objective optimization of knowledge rules. In *HIS '08: Proceedings of the 2008 8th International Conference on Hybrid Intelligent Systems*, pages 537–542, Washington, DC, USA, 2008. IEEE Computer Society.

- [33] David E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Publishing Company, Reading, Massachusetts, 1989.
- [34] John H. Holland. Genetic algorithms and the optimal allocation of trials. *SIAM J. Comput.*, 2(2):88–105, 1973.
- [35] Jeffrey Horn and Nicholas Nafpliotis. Multiobjective Optimization using the Nicheed Pareto Genetic Algorithm. Technical Report IlliGAI Report 93005, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA, 1993.
- [36] Evan J. Hughes. Multiple Single Objective Pareto Sampling. In *Proceedings of the 2003 Congress on Evolutionary Computation (CEC'2003)*, volume 4, pages 2678–2684, Canberra, Australia, December 2003. IEEE Press.
- [37] Evan J. Hughes. Evolutionary Many-Objective Optimisation: Many Once or One Many? In *2005 IEEE Congress on Evolutionary Computation (CEC'2005)*, volume 1, pages 222–227, Edinburgh, Scotland, September 2005. IEEE Service Center.
- [38] Evan J. Hughes. Fitness Assignment Methods for Many-Objective Problems. In Joshua Knowles, David Corne, and Kalyanmoy Deb, editors, *Multi-Objective Problem Solving from Nature: From Concepts to Applications*, pages 307–329. Springer, Berlin, 2008. ISBN 978-3-540-72963-1.
- [39] Evan J. Hughes. Evan's Software, August 2009. <http://code.evanhughes.org/>.
- [40] Hisao Ishibuchi, Noritaka Tsukamoto, and Yusuke Nojima. Evolutionary many-objective optimization: A short review. In *2008 Congress on Evolutionary Computation (CEC'2008)*, pages 2424–2431, Hong Kong, June 2008. IEEE Service Center.
- [41] Antonio López Jaimes, Carlos A. Coello Coello, and Jesús E. Urías Barrientos. Online Objective Reduction to Deal with Many-Objective Problems. In Matthias Ehrgott, Carlos M. Fonseca, Xavier Gandibleux, Jin-Kao Hao, and Marc Sevaux, editors, *Evolutionary Multi-Criterion Optimization. 5th International Conference, EMO 2009*, pages 423–437. Springer. Lecture Notes in Computer Science Vol. 5467, Nantes, France, April 2009.



- [42] V. Khare, X. Yao, and K. Deb. Performance Scaling of Multi-objective Evolutionary Algorithms. In Carlos M. Fonseca, Peter J. Fleming, Eckart Zitzler, Kalyanmoy Deb, and Lothar Thiele, editors, *Evolutionary Multi-Criterion Optimization. Second International Conference, EMO 2003*, pages 376–390, Faro, Portugal, April 2003. Springer. Lecture Notes in Computer Science. Volume 2632.
- [43] Joshua Knowles and David Corne. Quantifying the Effects of Objective Space Dimension in Evolutionary Multiobjective Optimization. In Shigeru Obayashi, Kalyanmoy Deb, Carlo Poloni, Tomoyuki Hiroyasu, and Tadahiko Murata, editors, *Evolutionary Multi-Criterion Optimization, 4th International Conference, EMO 2007*, pages 757–771, Matshushima, Japan, March 2007. Springer. Lecture Notes in Computer Science Vol. 4403.
- [44] Joshua Knowles, Lothar Thiele, and Eckart Zitzler. A Tutorial on the Performance Assessment of Stochastic Multiobjective Optimizers. 214, Computer Engineering and Networks Laboratory (TIK), ETH Zurich, Switzerland, February 2006. revised version.
- [45] Joshua D. Knowles and David W. Corne. The Pareto Archived Evolution Strategy: A New Baseline Algorithm for Multiobjective Optimisation. In *1999 Congress on Evolutionary Computation*, pages 98–105, Washington, D.C., July 1999. IEEE Service Center.
- [46] Ikeda Kokolo, Kita Hajime, and Kobayashi Shigenobu. Failure of Pareto-based MOEAs: Does Non-dominated Really Mean Near to Optimal? In *Proceedings of the Congress on Evolutionary Computation 2001 (CEC'2001)*, volume 2, pages 957–962, Piscataway, New Jersey, May 2001. IEEE Service Center.
- [47] Mario Köppen and Kaori Yoshida. Substitute Distance Assignments in NSGA-II for Handling Many-Objective Optimization Problems. In Shigeru Obayashi, Kalyanmoy Deb, Carlo Poloni, Tomoyuki Hiroyasu, and Tadahiko Murata, editors, *Evolutionary Multi-Criterion Optimization, 4th International Conference, EMO 2007*, pages 727–741, Matshushima, Japan, March 2007. Springer. Lecture Notes in Computer Science Vol. 4403.

- [48] Marco Laumanns, Lothar Thiele, Kalyanmoy Deb, and Eckart Zitzler. Combining Convergence and Diversity in Evolutionary Multi-objective Optimization. *Evolutionary Computation*, 10(3):263–282, Fall 2002.
- [49] Khoi Le and Dario Landa-Silva. Obtaining Better Non-Dominated Sets Using Volume Dominance. In *2007 IEEE Congress on Evolutionary Computation (CEC'2007)*, pages 3119–3126, Singapore, September 2007. IEEE Press.
- [50] Antonio López Jaimes and Carlos A. Coello Coello. Study of Preference Relations in Many-Objective Optimization. In *2009 Genetic and Evolutionary Computation Conference (GECOCO'2009)*, pages 611–618, Montreal, Canada, July 8–12 2009. ACM Press. ISBN 978-1-60558-325-9.
- [51] Antonio López Jaimes, Carlos A. Coello Coello, and Debrup Chakraborty. Objective Reduction Using a Feature Selection Technique. In *2008 Genetic and Evolutionary Computation Conference (GECCO'2008)*, pages 674–680, Atlanta, USA, July 2008. ACM Press. ISBN 978-1-60558-131-6.
- [52] Antonio López Jaimes, Luis Vicente Santana Quintero, and Carlos A. Coello Coello. Ranking methods in many-objective evolutionary algorithms. In Raymond Chiong, editor, *Nature-Inspired Algorithms for Optimisation*, pages 413–434. Springer, Berlin, 2009. ISBN 978-3-642-00266-3.
- [53] Kuntinee Maneeratana, Kittipong Boonlong, and Nachol Chaiyaratana. Compressed-Objective Genetic Algorithm. In Thomas Philip Runarsson, Hans-Georg Beyer, Edmund Burke, Juan J. Merelo-Guervós, L. Darrell Whitley, and Xin Yao, editors, *Parallel Problem Solving from Nature - PPSN IX, 9th International Conference*, pages 473–482. Springer. Lecture Notes in Computer Science Vol. 4193, Reykjavik, Iceland, September 2006.
- [54] K. Miettinen. *On the Methodology of Multiobjective Optimization with Applications, Report 60 (doctoral thesis)*. PhD thesis, University of Jyväskylä, Department of Mathematics, Jyväskylä, 1994.

- [55] A.K. Molyneaux, G.B. Leyland, and D.Favrat. A New, Clustering Evolutionary Multi-Objective Optimisation Technique. In *Proceedings of the Third International Symposium on Adaptive Systems—Evolutionary Computation and Probabilistic Graphical Models*, pages 41–47, Havana, Cuba, March 19–23 2001. Institute of Cybernetics, Mathematics and Physics.
- [56] A. Osyczka. *Multicriterion Optimization in Engineering with FORTRAN Programs*. Ellis Horwood Limited, 1984.
- [57] Vilfredo Pareto. *Cours d'Economie Politique*. Droz, Genève, 1896.
- [58] Kata Praditwong and Xin Yao. How Well Do Multi-Objective Evolutionary Algorithms Scale to Large Problems. In *2007 IEEE Congress on Evolutionary Computation (CEC'2007)*, pages 3959–3966, Singapore, September 2007. IEEE Press.
- [59] Robin C. Purshouse and Peter J. Fleming. Conflict, Harmony, and Independence: Relationships in Evolutionary Multi-criterion Optimisation. In Carlos M. Fonseca, Peter J. Fleming, Eckart Zitzler, Kalyanmoy Deb, and Lothar Thiele, editors, *Evolutionary Multi-Criterion Optimization. Second International Conference, EMO 2003*, pages 16–30, Faro, Portugal, April 2003. Springer. Lecture Notes in Computer Science. Volume 2632.
- [60] Robin C. Purshouse and Peter J. Fleming. Evolutionary Multi-Objective Optimisation: An Exploratory Analysis. In *Proceedings of the 2003 Congress on Evolutionary Computation (CEC'2003)*, volume 3, pages 2066–2073, Canberra, Australia, December 2003. IEEE Press.
- [61] Robin C. Purshouse and Peter J. Fleming. On the Evolutionary Optimization of Many Conflicting Objectives. *IEEE Transactions on Evolutionary Algorithms*, 11(6):770–784, December 2007.
- [62] Robin Charles Purshouse. *On the Evolutionary Optimisation of Many Objectives*. PhD thesis, Department of Automatic Control and Systems Engineering, The University of Sheffield, Sheffield, UK, September 2003.
- [63] S.S Rao. *Engineering Optimization: Theory and Practice*. 3rd ed. John Wiley and Sons, New York, 1996. ISBN 0-471-55034-5.

- [64] Ingo Rechenberg. *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*. Frommann-Holzboog, Stuttgart, Alemania, 1973.
- [65] R. S. Rosenberg. *Simulation of genetic populations with biochemical properties*. PhD thesis, University of Michigan, Ann Harbor, Michigan, 1967.
- [66] Hiroyuki Sato, Hernán E. Aguirre, and Kiyoshi Tanaka. Controlling Dominance Area of Solutions and Its Impact on the Performance of MOEAs. In Shigeru Obayashi, Kalyanmoy Deb, Carlo Poloni, Tomoyuki Hiroyasu, and Tadahiko Murata, editors, *Evolutionary Multi-Criterion Optimization, 4th International Conference, EMO 2007*, pages 5–20, Matsushima, Japan, March 2007. Springer. Lecture Notes in Computer Science Vol. 4403.
- [67] J. David Schaffer. *Multiple Objective Optimization with Vector Evaluated Genetic Algorithms*. PhD thesis, Vanderbilt University, 1984.
- [68] J. David Schaffer. Multiple Objective Optimization with Vector Evaluated Genetic Algorithms. In *Genetic Algorithms and their Applications: Proceedings of the First International Conference on Genetic Algorithms*, pages 93–100. Lawrence Erlbaum, 1985.
- [69] N. Srinivas and Kalyanmoy Deb. Multiobjective optimization using nondominated sorting in genetic algorithms. Technical report, Department of Mechanical Engineering, Indian Institute of Technology, Kanpur, India, 1993.
- [70] N. Srinivas and Kalyanmoy Deb. Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms. *Evolutionary Computation*, 2(3):221–248, Fall 1994.
- [71] David A. Van Veldhuizen. *Multiobjective Evolutionary Algorithms: Classifications, Analyses, and New Innovations*. PhD thesis, Department of Electrical and Computer Engineering. Graduate School of Engineering. Air Force Institute of Technology, Wright-Patterson AFB, Ohio, May 1999.
- [72] Tobias Wagner, Nicola Beume, and Boris Naujoks. Pareto-, Aggregation-, and Indicator-Based Methods in Many-Objective Optimization. In Shigeru Obayashi, Kalyanmoy Deb, Carlo Poloni,

- Tomoyuki Hiroyasu, and Tadahiko Murata, editors, *Evolutionary Multi-Criterion Optimization, 4th International Conference, EMO 2007*, pages 742–756, Matshushima, Japan, March 2007. Springer. Lecture Notes in Computer Science Vol. 4403.
- [73] E. Zitzler. *Evolutionary Algorithms for Multiobjective Optimization: Methods and Applications*. PhD thesis, ETH Zurich, Switzerland, 1999.
- [74] Eckart Zitzler, Marco Laumanns, and Lothar Thiele. SPEA2: Improving the Strength Pareto Evolutionary Algorithm. Technical Report 103, Computer Engineering and Networks Laboratory (TIK), Swiss Federal Institute of Technology (ETH) Zurich, Gloriastrasse 35, CH-8092 Zurich, Switzerland, May 2001.
- [75] Eckart Zitzler and Lothar Thiele. Multiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach. *IEEE Transactions on Evolutionary Computation*, 3(4):257–271, November 1999.
- [76] Xiufen Zou, Yu Chen, Minzhong Liu, and Lishan Kang. A New Evolutionary Algorithm for Solving Many-Objective Optimization Problems. *IEEE Transactions on Systems, Man, and Cybernetics–Part B: Cybernetics*, 38(5):1402–1412, October 2008.