Comparação de Resultados do Algoritmo Genético Representação Binária vs. Representação Real

Rafael Dória Monteiro de Miranda
Professor: Roberto Limão
Faculdade de Computação - ICEN,
Universidade Federal do Pará (UFPA)

1 Introdução

Neste relatório, comparamos os resultados de dois códigos que implementam a mesma metaheurística para resolver um problema de otimização. Um dos códigos utiliza uma representação binária, enquanto o outro utiliza uma representação real. A metaheurística escolhida para esta comparação é o Algoritmo Genético, que é amplamente utilizado em problemas de otimização devido à sua capacidade de explorar grandes espaços de busca.

1.1 Objetivos

O objetivo deste relatório é analisar e comparar a eficácia das representações binária e real na resolução de um problema de otimização, avaliando a qualidade das soluções encontradas e o tempo de execução dos algoritmos.

2 Metodologia

2.1 Descrição da Metaheurística

O algoritmo genético (AG) é uma técnica de otimização inspirada no processo de seleção natural, baseada na ideia de evolução das espécies. Ele é amplamente utilizado para resolver problemas complexos de busca e otimização, especialmente quando o espaço de solução é muito vasto ou não pode ser explorado eficientemente por métodos tradicionais. O AG utiliza conceitos como seleção, crossover e mutação para evoluir soluções ao longo de sucessivas gerações.

2.2 Representação Binária

A representação binária é uma das formas mais comuns de codificação no AG. Nesse formato, cada solução é representada por uma cadeia de bits (0s e 1s). Por exemplo, um indivíduo pode ser representado como 10101001, onde cada bit corresponde a uma parte da solução. A manipulação dessas cadeias é realizada através de operadores genéticos, como crossover e mutação, permitindo a geração de novas soluções. A principal vantagem da representação binária é sua simplicidade e facilidade de implementação. No entanto, para problemas com variáveis contínuas, essa abordagem pode levar a dificuldades na conversão entre binário e valores reais, exigindo precisão e ajustes adequados.

2.3 Representação Real

Na representação real, os indivíduos são representados diretamente por valores reais em vez de cadeias binárias. Por exemplo, um indivíduo pode ser representado por [3.14, 2.71, -1.5], onde cada elemento do vetor é um valor dentro de um intervalo permitido. A vantagem desse tipo de representação é sua eficiência na manipulação de variáveis contínuas, evitando os problemas de conversão encontrados na codificação binária. Além disso, operadores genéticos podem ser mais adaptáveis a cada problema específico.

2.4 Configurações dos Experimentos

A função de Schaffer modificada F6 é uma função matemática utilizada em problemas de otimização e testes de algoritmos evolutivos. Sua fórmula é definida por:

$$f(x, y) = 0.5 - \frac{(\sin(\sqrt{x^2 + y^2}))^2 - 0.5}{(1 + 0.001(x^2 + y^2))^2} \tag{1}$$

Essa função apresenta características desafiadoras para algoritmos de otimização, pois possui múltiplos máximos locais, tornando a busca pela solução ótima mais complexa. O objetivo é maximizar, sendo que o máximo global ocorre em f(0,0) com um valor de 1. Devido à presença da função seno e do denominador não linear, essa função gera uma superfície ondulada com regiões de baixíssimos valores, exigindo métodos de otimização robustos, como os algoritmos genéticos, para encontrar a melhor solução de forma eficiente.

2.5 Requisitos Gerais

- Método de Seleção por meio da Roleta proporcional à medida de aptidão do indivíduo.
- Número de gerações: 500.
- Tamanho da população: 200.
- Taxa de mutação: 0.01.
- Taxa de Cruzamento: 0.85.
- Domínio das Variáveis: [-100, +100].

2.6 Requisitos para Representação Real

- Método de Cruzamento Aritmético com Alpha = 0.35.
- Método de Mutação Randômica Não Uniforme, com rand = normal de média zero e variância a ser definida pelo aluno.

2.7 Requisitos para Representação Binária

- Método de Cruzamento binário com um ponto de corte.
- Bits para representação: 25 bits cada variável.
- Precisão de 5 casas decimais para as duas variáveis.

3 Algoritmos

3.1 AG - Gerais

Inclusão de Bibliotecas

O código começa com a inclusão de bibliotecas essenciais para operações de entrada e saída, alocação de memória, funções matemáticas e manipulação de tempo.

Definição da Estrutura de Dados

A estrutura `Solution` é definida para representar uma solução no algoritmo genético, contendo os valores reais de `x` e `y`, além do valor da função F6.

Função de Cálculo de Fitness

A função `calculate_fitness` calcula o valor da função F6 para uma solução dada, utilizando os valores de `x` e `y` da estrutura `Solution`.

Seleção por Roleta

A função `roulette_selection` implementa a seleção por roleta, onde indivíduos são selecionados com base em sua aptidão. Um valor aleatório é gerado e o indivíduo correspondente à soma cumulativa de aptidão é retornado.

Execução do Algoritmo Genético

A função `run_genetic_algorithm` é o núcleo do algoritmo genético, onde a população é inicializada, as gerações são executadas, e os resultados são salvos em arquivos. A melhor solução encontrada em cada geração é rastreada e impressa.

3.2 AG - Binário

Cruzamento

A função `crossover_one_point` realiza o cruzamento de um ponto entre dois pais, onde um ponto de cruzamento é escolhido aleatoriamente e os filhos são gerados a partir das partes dos pais.

Mutação

A função `mutate` aplica mutação a uma solução, invertendo bits nas representações binárias de `x` e `y` com uma probabilidade definida pela taxa de mutação.

4 AG - Real

Cruzamento

A função de cruzamento é comentada, mas a ideia é que ela combina os valores de dois pais para gerar filhos. O cruzamento aritmético é utilizado, onde os filhos são gerados como uma combinação ponderada dos valores dos pais.

Mutação

A função `mutate_non_uniform` aplica mutação a uma solução de forma não uniforme, utilizando uma função gaussiana para perturbar os valores de `x` e `y`.

5 Resultados

Os resultados obtidos para as duas representações são apresentados nas Tabelas 1 e 2. A Tabela 1 mostra os resultados do algoritmo com representação binária, enquanto a Tabela 2 apresenta os resultados do algoritmo com representação real.

Tabela 1: Resultados do Algoritmo Genético com Representação Binária

Geração	Melhor Solução Encontrada
1	0.637924
10	0.921805
50	0.990282
100	0.990284
200	0.998505
300	0.995254
400	0.990216
500	0.991796

Tabela 2: Resultados do Algoritmo Genético com Representação Real

Geração	Melhor Solução Encontrada por Geração
1	0.916726
10	0.984949
50	0.989541
100	0.987226
200	0.988735
300	0.990227
400	0.990023
500	0.976845

Tabela 3: Resultados dos Algoritmo Genético

Representação	Geração	Melhor Solução	valor x	valor y
Binário	166	0.999913	-0.008062	0.004682
Real	192	0.999738	0.011228	0.011662

6 Análise dos Resultados

A análise dos resultados obtidos nas Tabelas 1 e 2 revela diferenças significativas entre as duas abordagens de representação no algoritmo genético.

Na Tabela 1, observamos que a melhor solução encontrada pelo algoritmo com representação binária evoluiu de 0.637924 na primeira geração para 0.998505 na geração 200. Essa melhoria contínua indica que o algoritmo foi capaz de explorar o espaço de soluções de forma eficaz, alcançando uma solução bastante próxima do ótimo. No entanto, a partir da geração 300, a qualidade da solução começou a oscilar, com um leve declínio na geração 500, onde a melhor solução foi de 0.991796. Isso sugere que o algoritmo pode ter encontrado um ótimo local, mas não conseguiu melhorar ainda mais a solução.

Por outro lado, a Tabela 2 mostra que a representação real começou com uma melhor solução inicial de 0.916726 na primeira geração e alcançou um pico de 0.990227 na geração 300. A solução na geração 500 caiu para 0.976845, o que indica uma instabilidade na busca por soluções ótimas. Apesar disso, a representação real apresentou soluções consistentemente mais altas em comparação com a representação binária, especialmente nas primeiras gerações.

Quando comparamos os resultados finais das duas representações na Tabela 3, notamos que a representação binária alcançou uma melhor solução de 0.999913 na geração 166, enquanto a representação real obteve uma solução de 0.999738 na geração 192. Embora a representação binária tenha encontrado uma solução ligeiramente melhor, a diferença é mínima, e ambas as representações demonstraram um desempenho robusto.

Além disso, os valores de x e y associados às melhores soluções também são relevantes. A representação binária apresentou um valor de $x = -0.008062$ e $y = 0.004682$, enquanto a representação real teve $x = 0.011228$ e $y = 0.011662$. Isso sugere que, embora a representação binária tenha encontrado uma solução ligeiramente melhor em termos de valor, a representação real pode ter explorado uma região diferente do espaço de soluções, resultando em valores de x e y que podem ser mais desejáveis dependendo do contexto do problema.

7 Discussão

Os resultados obtidos indicam que, embora ambas as representações do algoritmo genético tenham demonstrado eficácia, a representação real apresentou um desempenho mais consistente em termos de qualidade da solução nas gerações iniciais. A representação real começou com uma melhor solução inicial e, embora tenha mostrado uma leve instabilidade nas gerações posteriores, conseguiu manter soluções de qualidade superior em comparação com a representação binária. Por outro lado, a representação binária, embora tenha alcançado uma solução ligeiramente melhor em um número menor de gerações, apresentou oscilações significativas nas soluções ao longo do tempo, sugerindo que pode ter encontrado um ótimo local.

Uma análise mais detalhada dos gráficos da representação binária revela que, apesar de a maioria dos indivíduos convergir para um resultado, ainda existem vários indivíduos que se espalham por todo o gráfico. Isso resulta em um fenômeno onde o convergimento de todos os indivíduos não ocorre de maneira uniforme. Em contraste, na representação real, a primeira geração apresentava indivíduos muito espalhados, assim como na representação binária. No entanto, a cada geração subsequente, o convergimento de todos os indivíduos aumentava, demonstrando que, quanto mais gerações se passavam, mais agrupados estavam os indivíduos. Essa tendência sugere que a representação real é mais eficaz em promover a convergência global das soluções ao longo do tempo.

Essas observações sugerem que a escolha da representação deve ser cuidadosamente considerada com base nas características do problema em questão. A representação real pode

ser mais eficaz para problemas que exigem uma exploração mais contínua do espaço de soluções, enquanto a representação binária pode ser mais adequada para problemas onde as variáveis são discretas ou onde a estrutura do espaço de busca é mais complexa.

8 Conclusão

Este relatório apresentou uma comparação entre duas representações de um algoritmo genético aplicado a um problema de otimização. Os resultados mostraram que, embora a representação binária tenha alcançado uma solução ligeiramente melhor em um número menor de gerações, a representação real superou a binária em termos de qualidade da solução em várias gerações e consistência. A análise dos gráficos revelou que, na representação binária, embora a maioria dos indivíduos convergisse para um resultado, ainda havia uma dispersão significativa, o que impediu um convergimento uniforme. Em contraste, a representação real, que começou com indivíduos mais espalhados, demonstrou um aumento no agrupamento e convergência ao longo das gerações, indicando uma exploração mais eficaz do espaço de soluções.

Essa análise sugere que a representação real pode ser mais adequada para a maioria dos problemas de otimização, especialmente aqueles que se beneficiam de uma exploração mais contínua do espaço de soluções. Futuros trabalhos podem explorar outras metaheurísticas, diferentes configurações de parâmetros e a aplicação de técnicas híbridas para uma análise mais abrangente e robusta, considerando também a dinâmica de convergência das representações utilizadas.

A Apêndice

Neste apêndice, fornecemos links para os códigos e imagens utilizados neste relatório, que estão disponíveis no repositório do GitHub.

A.1 Códigos

O código fonte do algoritmo genético, incluindo as implementações para as representações binária e real. Assim como as imagens geradas durante a execução dos algoritmos, podem ser encontradas no seguinte link:

- `https://github.com/KuuhakuRD7/AlgoritmoGen-tico-TO.git`

B Plotes

B.1 Resultados da Representação Binária - Curva dos Indivíduos

Melhoria Inicial

Nos primeiros ciclos, observamos uma melhoria significativa na qualidade da solução. A fitness aumenta de 0.923999 na geração 0 para 0.962670 na geração 1. Essa rápida melhoria inicial sugere que o algoritmo está explorando efetivamente o espaço de soluções e encontrando soluções promissoras.

Estabilização e Oscilações

Após a geração 1, a taxa de melhoria começa a estabilizar, com a fitness oscilando em torno de valores próximos a 0.96. A partir da geração 4, a fitness se mantém em torno de 0.9626, indicando que o algoritmo pode estar se aproximando de um ótimo local. As oscilações menores

nas gerações subsequentes sugerem que as soluções estão se agrupando em torno de um valor de fitness relativamente alto.

Convergência para um Ótimo Local

A partir da geração 70, a fitness atinge valores muito próximos de 1, com algumas oscilações menores. A presença de valores como 0.999285 na geração 105 sugere que o algoritmo pode ter encontrado um ótimo local, onde as soluções não estão mais se movendo significativamente em direção a uma solução melhor.

Estabilidade das Variáveis

As colunas que representam os valores de x e y mostram um comportamento mais volátil nas primeiras gerações, mas tendem a se estabilizar em torno de certos valores à medida que o algoritmo avança. Por exemplo, os valores de x e y na geração 1 são 6.159082 e 1.155183, respectivamente, e se mantêm em uma faixa relativamente estreita nas gerações subsequentes. Isso indica que, enquanto a qualidade da solução (fitness) se estabiliza, as variáveis que compõem essa solução também estão se ajustando para valores que maximizam a função objetivo.

Exploração e Exploração

A oscilação nos valores de x e y sugere que o algoritmo está equilibrando a exploração de novas soluções e a exploração de soluções já conhecidas. Por exemplo, na geração 8, os valores de x e y mudam significativamente, indicando que o algoritmo está explorando novas áreas do espaço de soluções. Essa dinâmica é crucial em algoritmos genéticos, pois permite que o algoritmo não fique preso em ótimos locais, mas também não perca a qualidade das soluções.

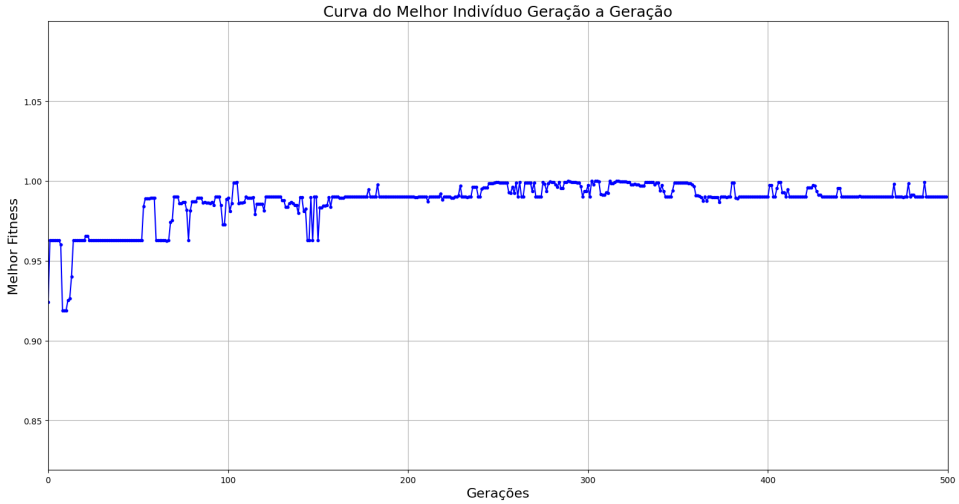

Figure 1: Individuos - Binário 0

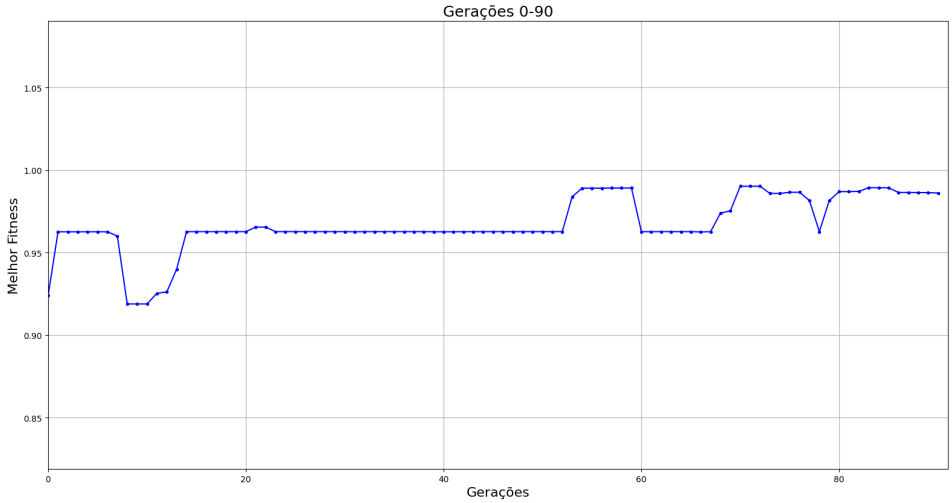

Figure 2: Individuos - Binário 1

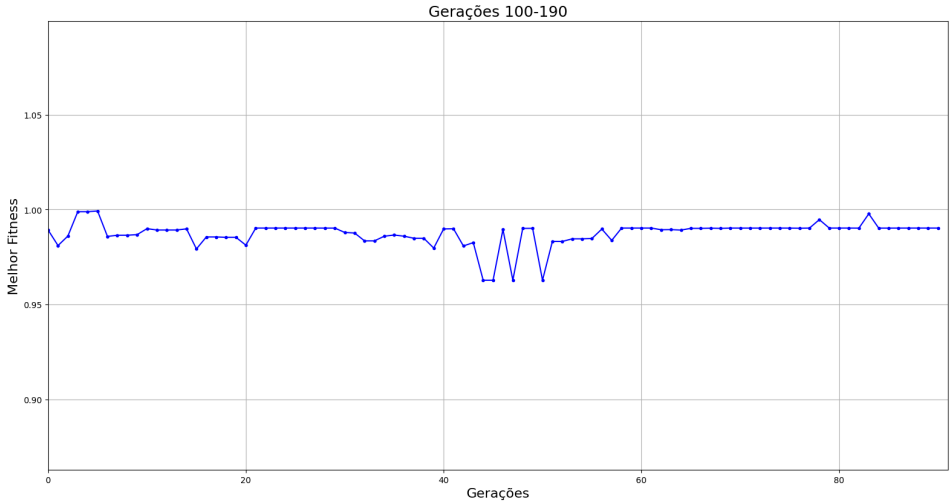

Figure 3: Individuos - Binário 2

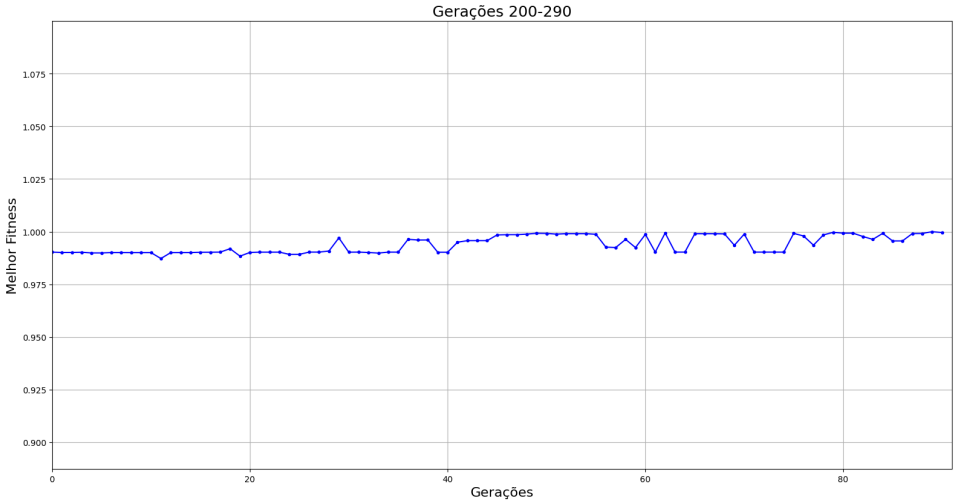

Figure 4: Individuos - Binário 3

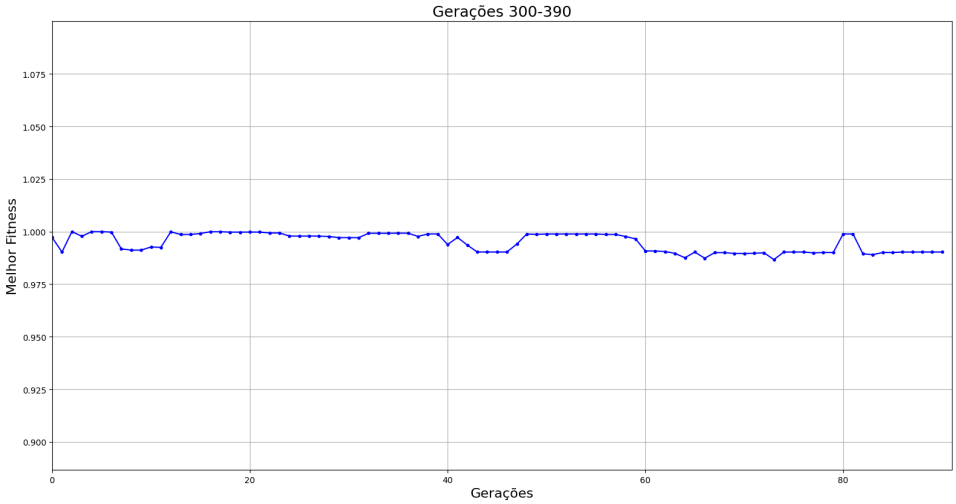

Figure 5: Individuos - Binário 4

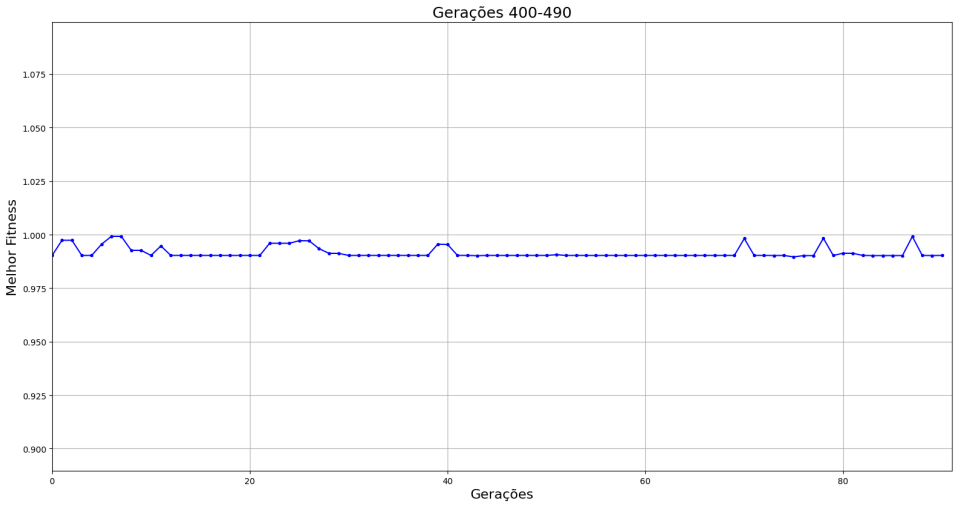

Figure 6: Individuos - Binário 5

B.2 Matrizes Das Populações por Geração

Nesta subseção, apresentamos as matrizes que representam as populações ao longo das gerações do algoritmo genético. Cada matriz ilustra a distribuição dos indivíduos em cada geração, permitindo uma análise visual da evolução das soluções ao longo do tempo. As matrizes ajudam a entender como a diversidade da população muda e como os indivíduos se agrupam em torno de soluções promissoras.

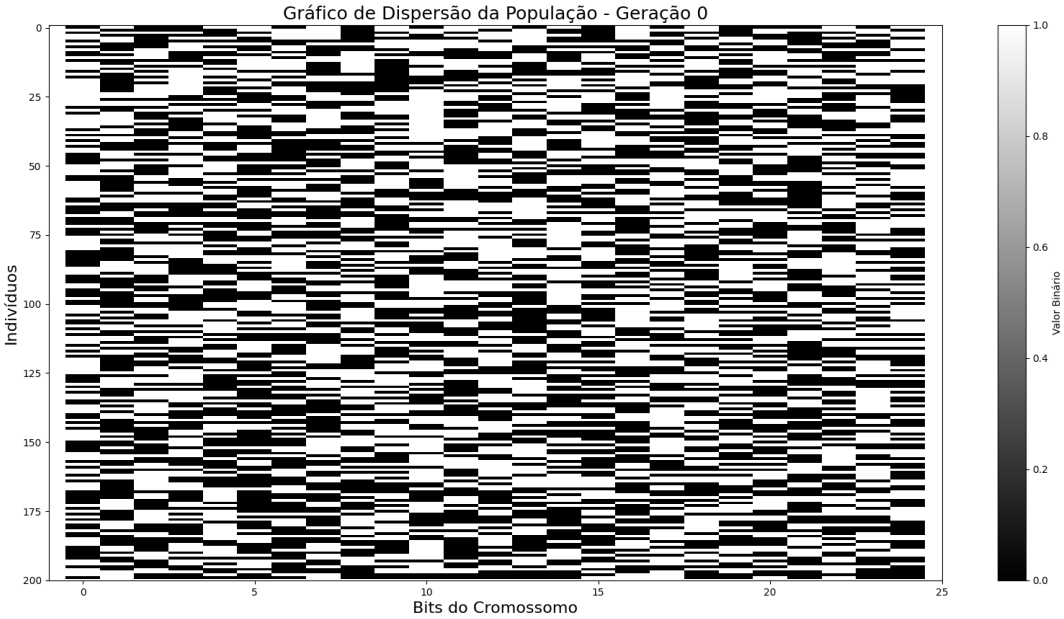

Figure 7: Matrizes 0

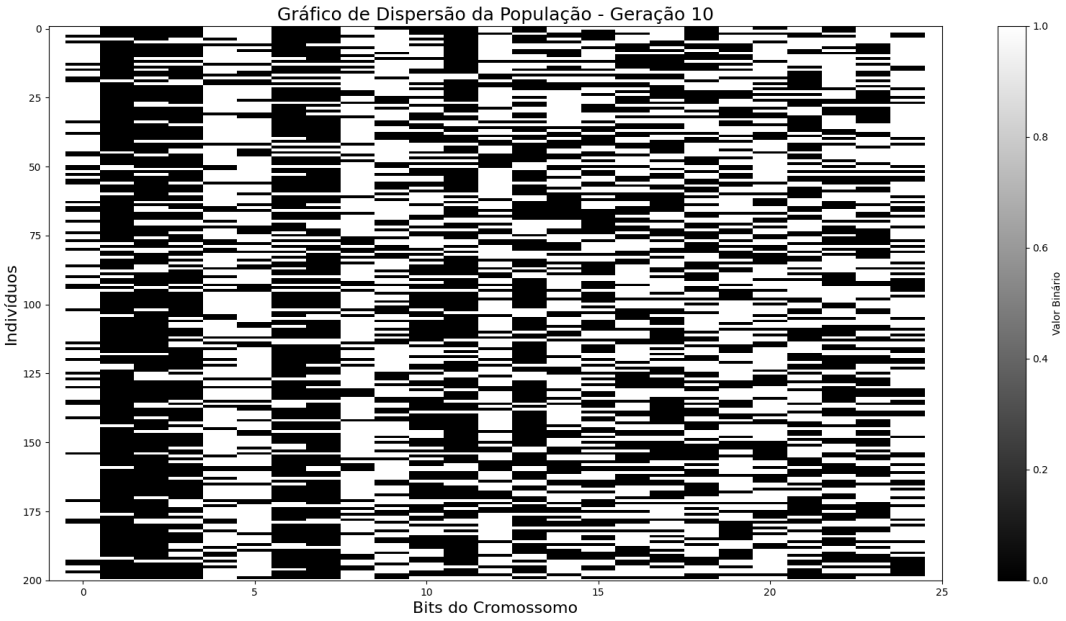

Figure 8: Matrizes 10

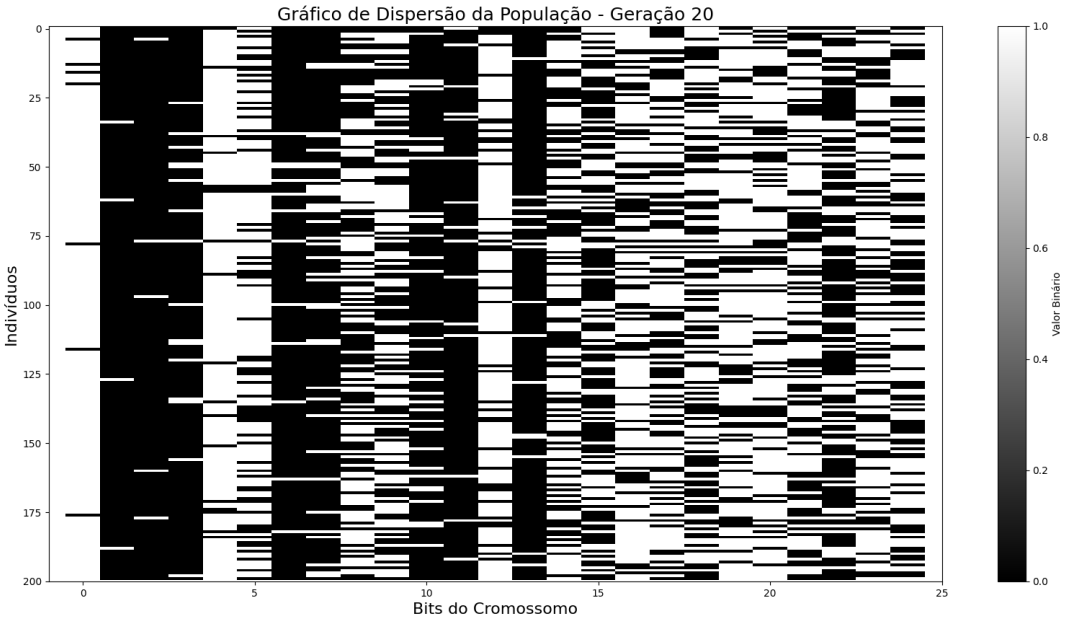

Figure 9: Matrizes 20

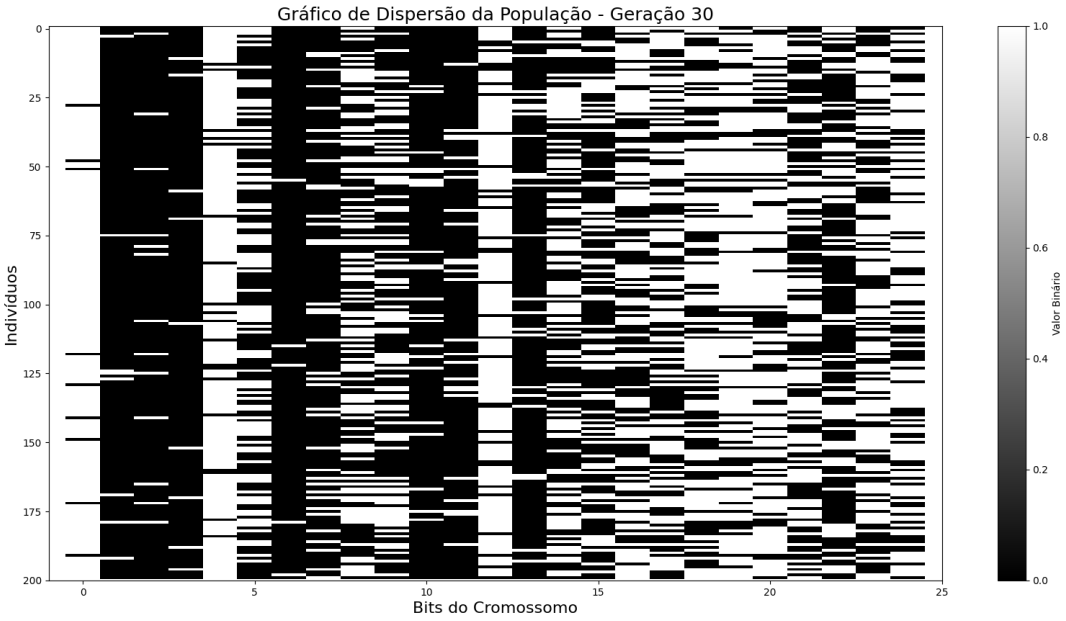

Figure 10: Matrizes 30

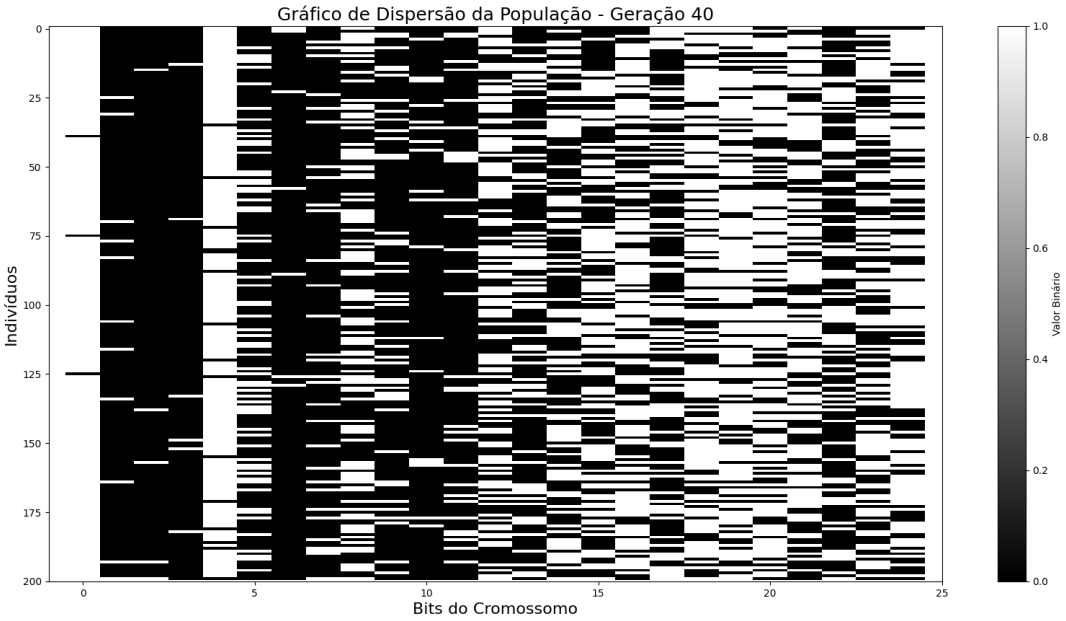

Figure 11: Matrizes 40

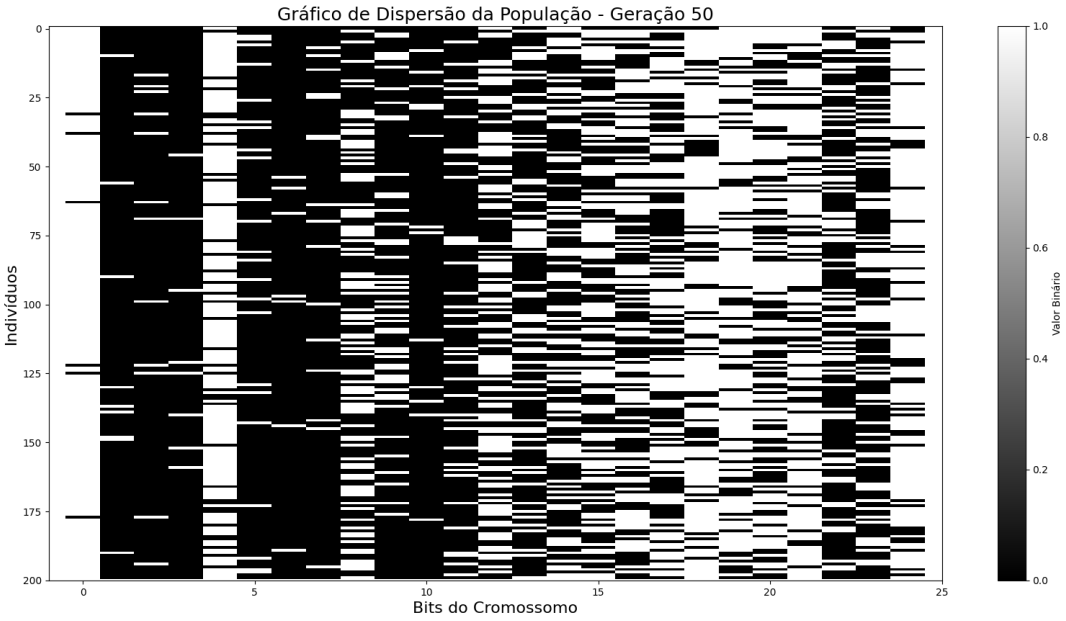

Figure 12: Matrizes 50

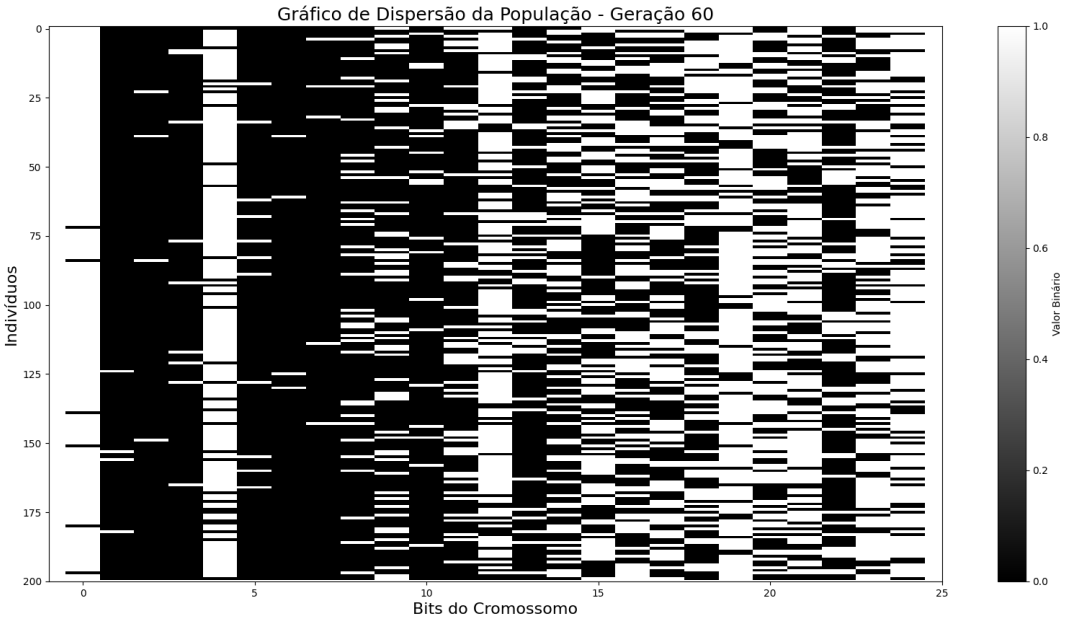

Figure 13: Matrizes 60

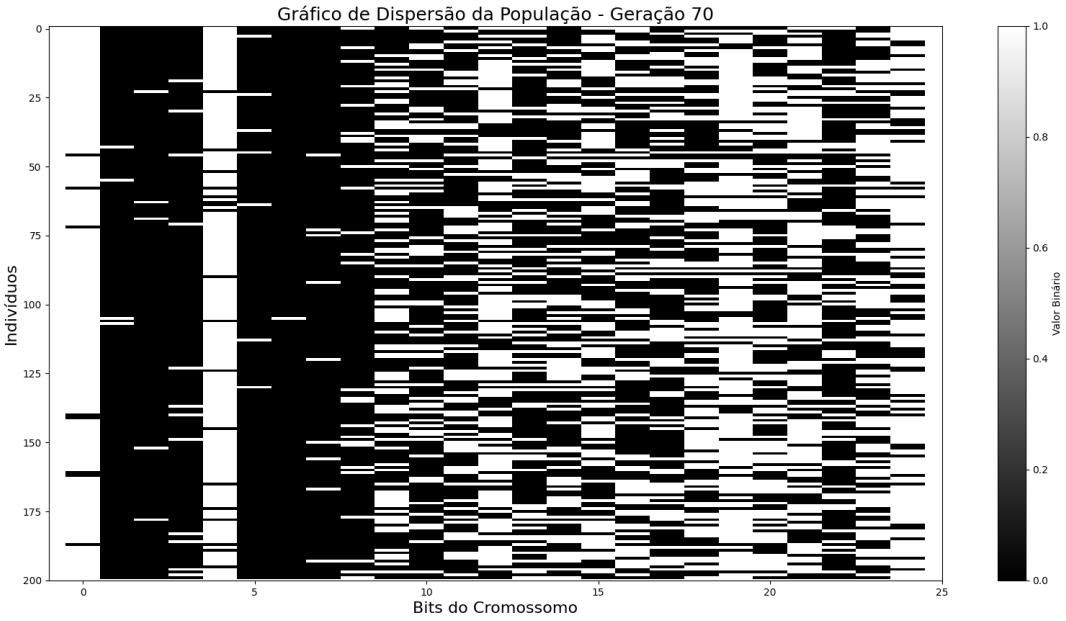

Figure 14: Matrizes 70

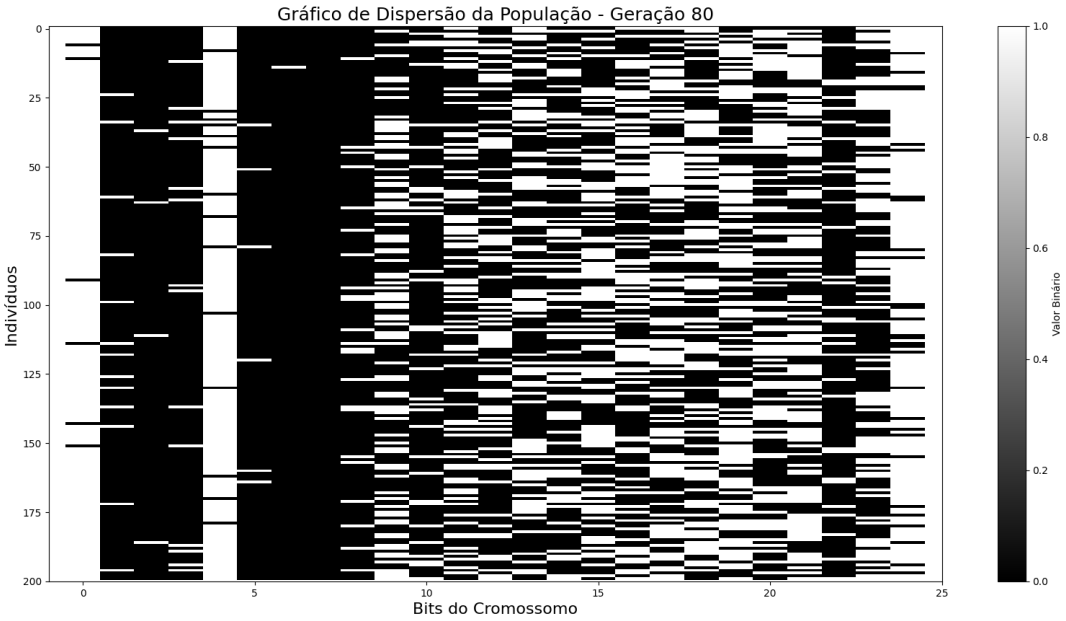

Figure 15: Matrizes 80

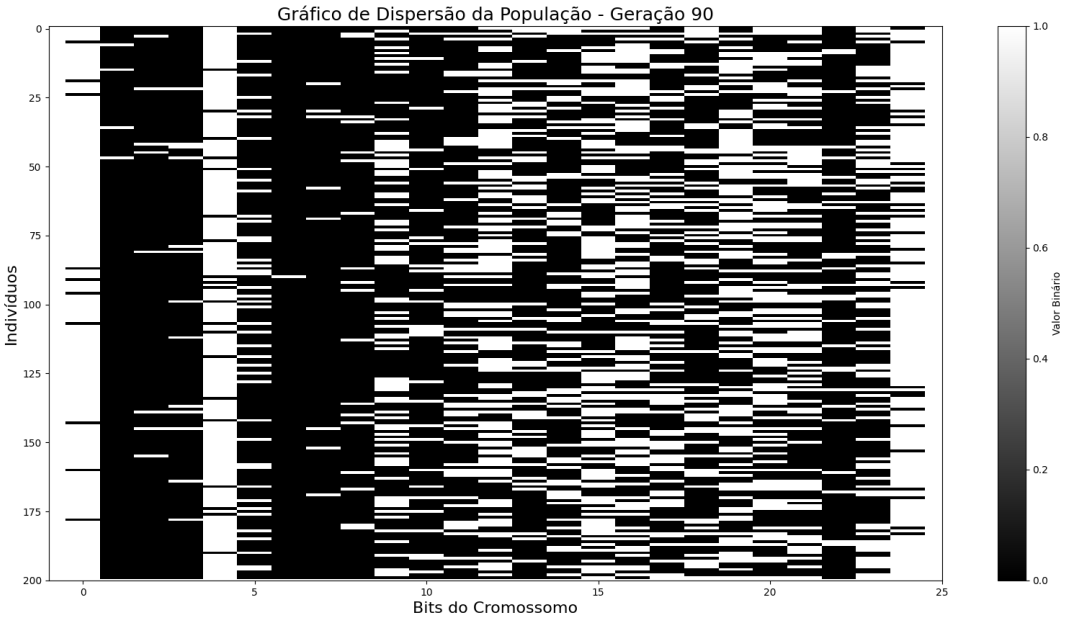

Figure 16: Matrizes 90

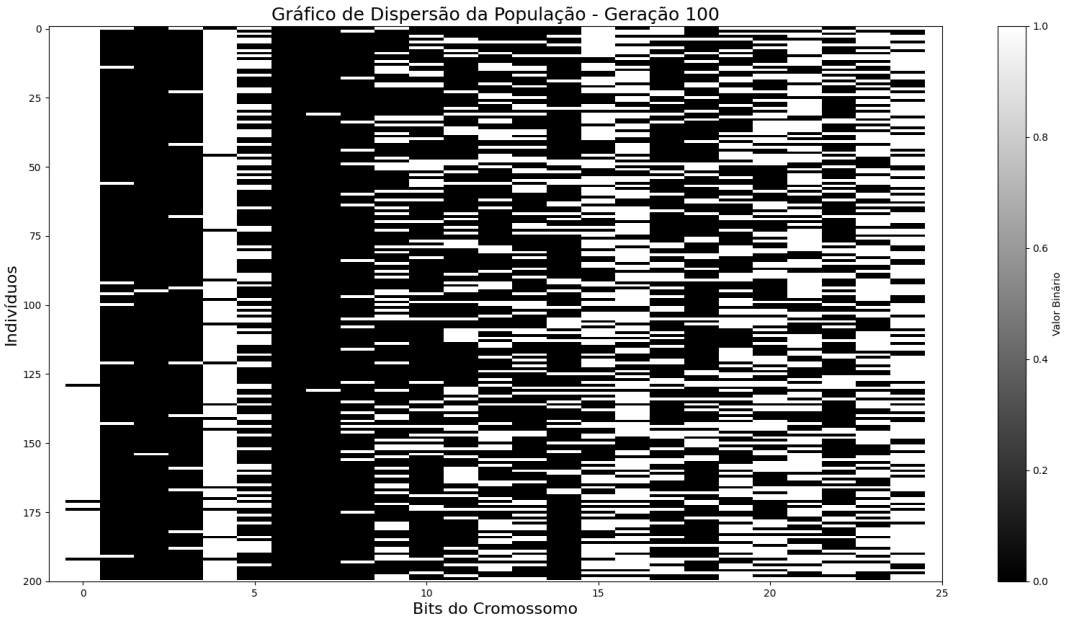

Figure 17: Matrizes 100

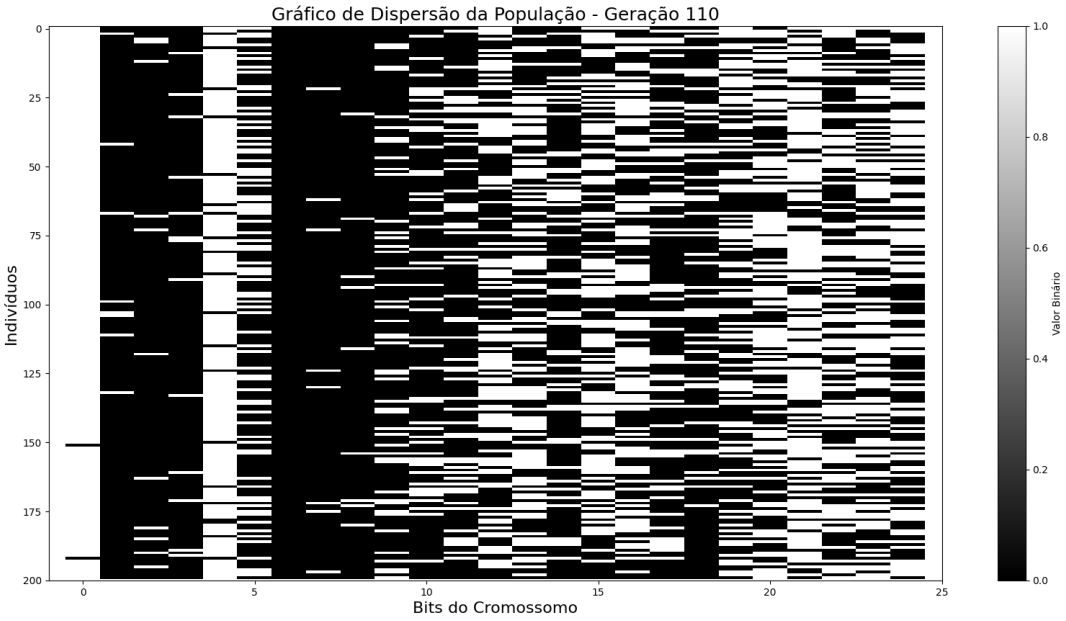

Figure 18: Matrizes 110

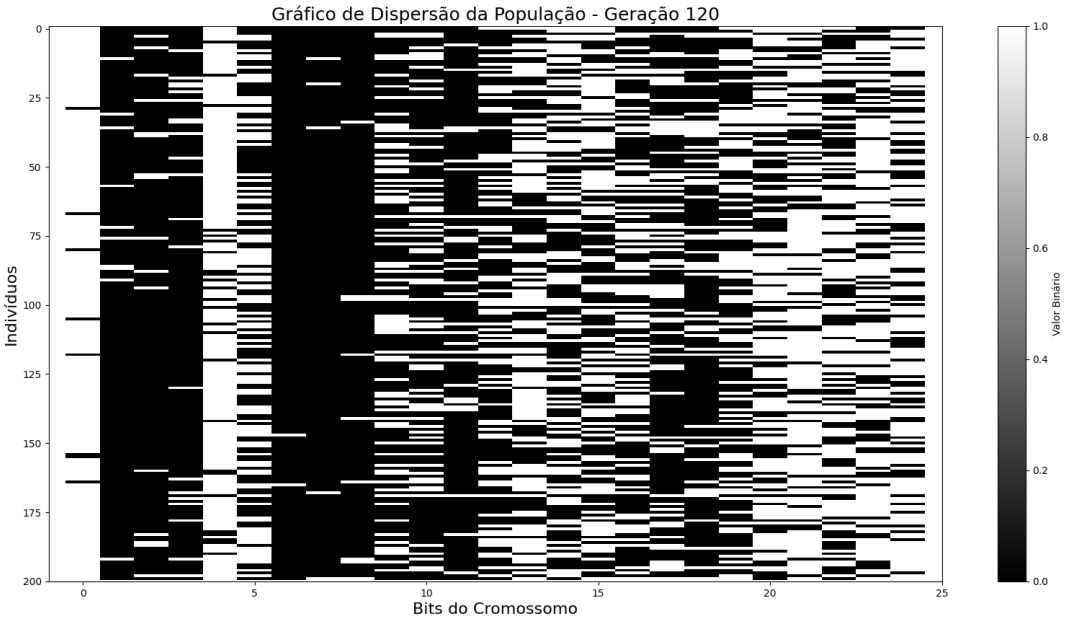

Figure 19: Matrizes 120

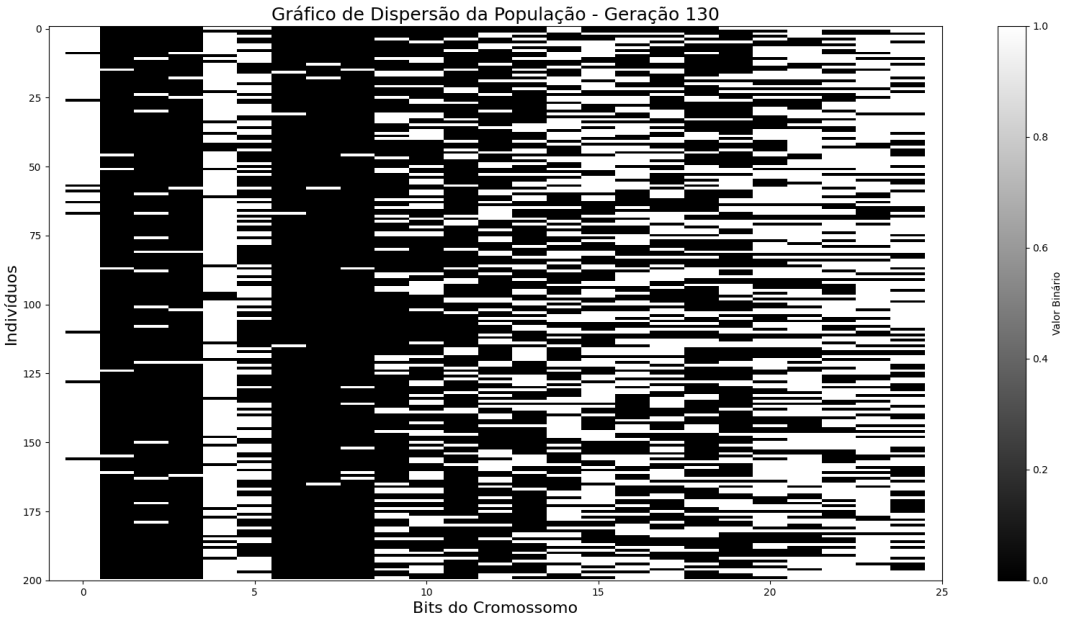

Figure 20: Matrizes 130

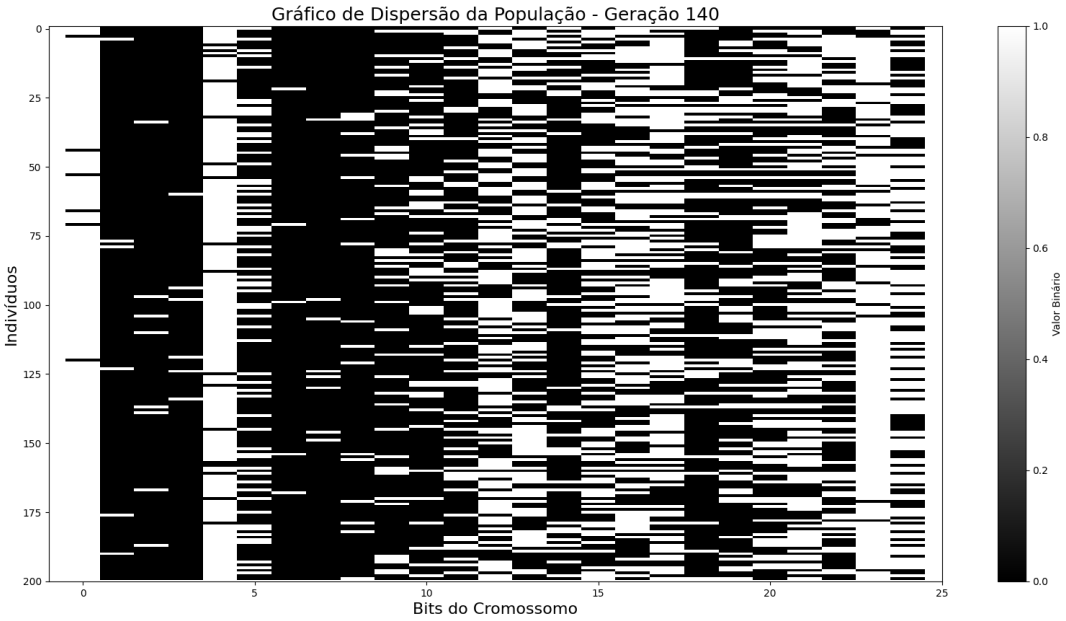

Figure 21: Matrizes 140

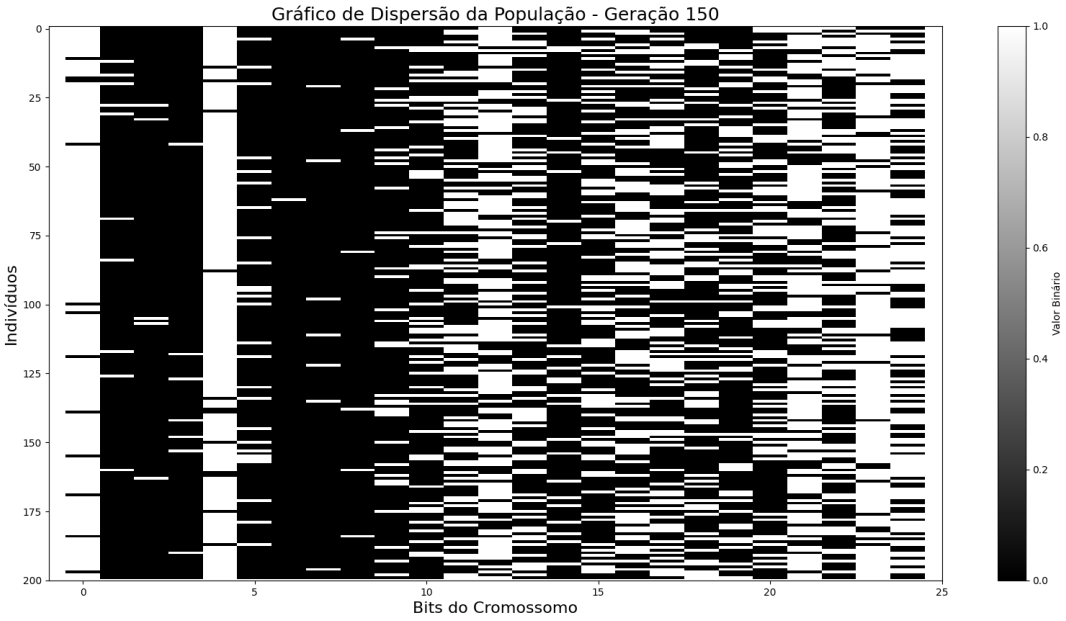

Figure 22: Matrizes 150

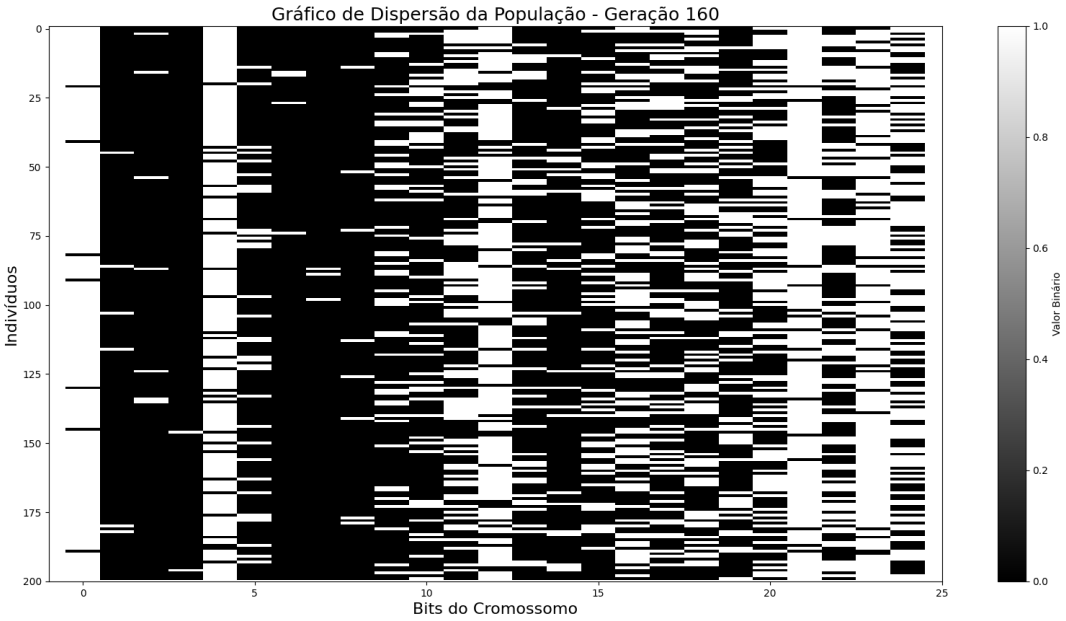

Figure 23: Matrizes 160

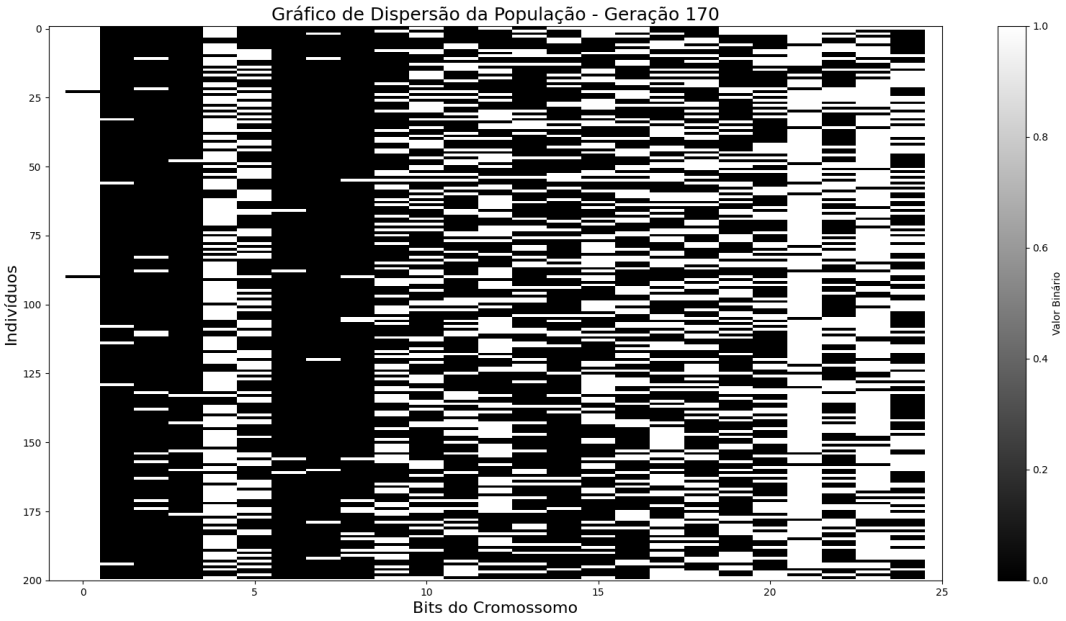

Figure 24: Matrizes 170

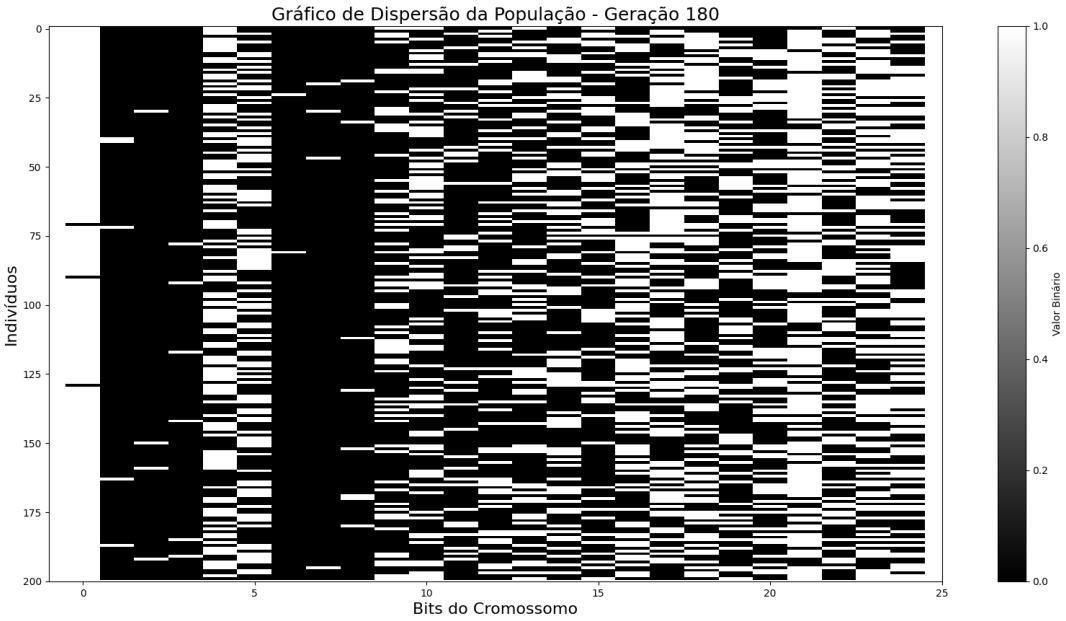

Figure 25: Matrizes 180

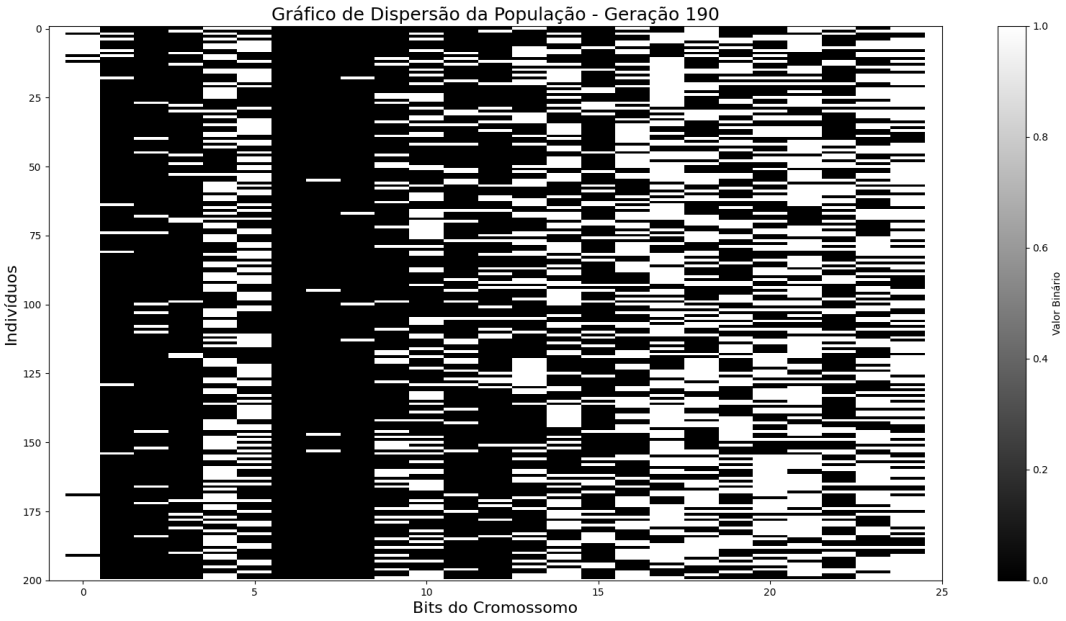

Figure 26: Matrizes 190

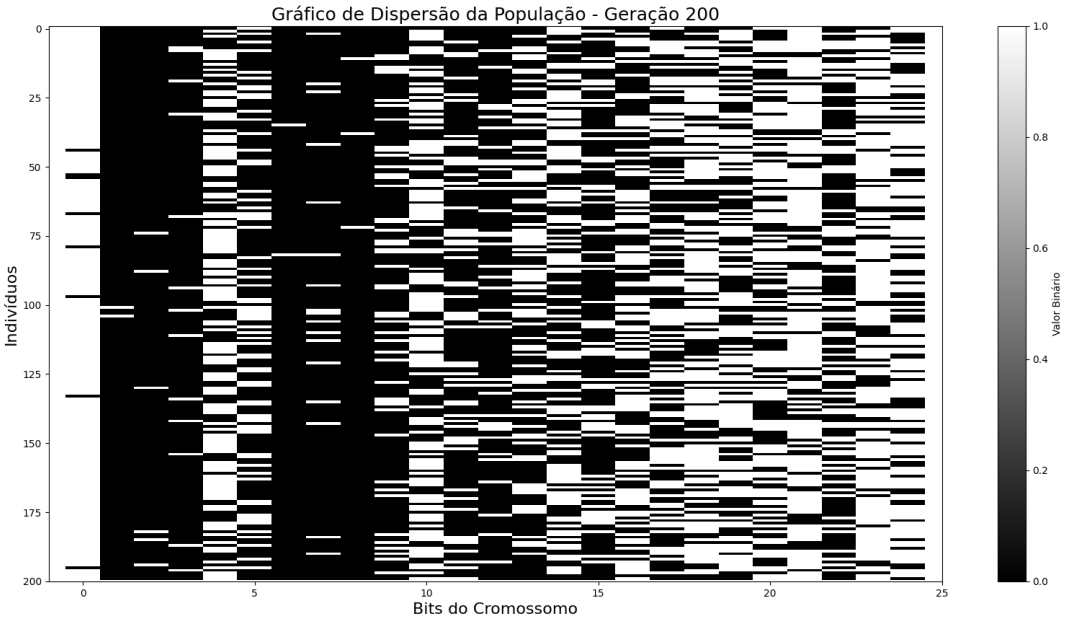

Figure 27: Matrizes 200

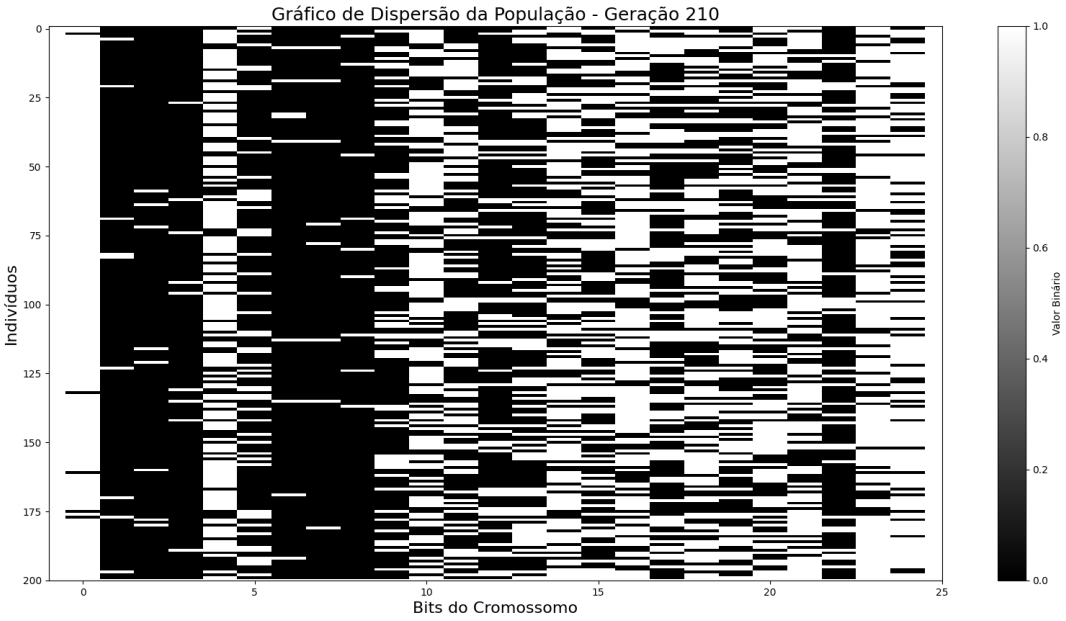

Figure 28: Matrizes 210

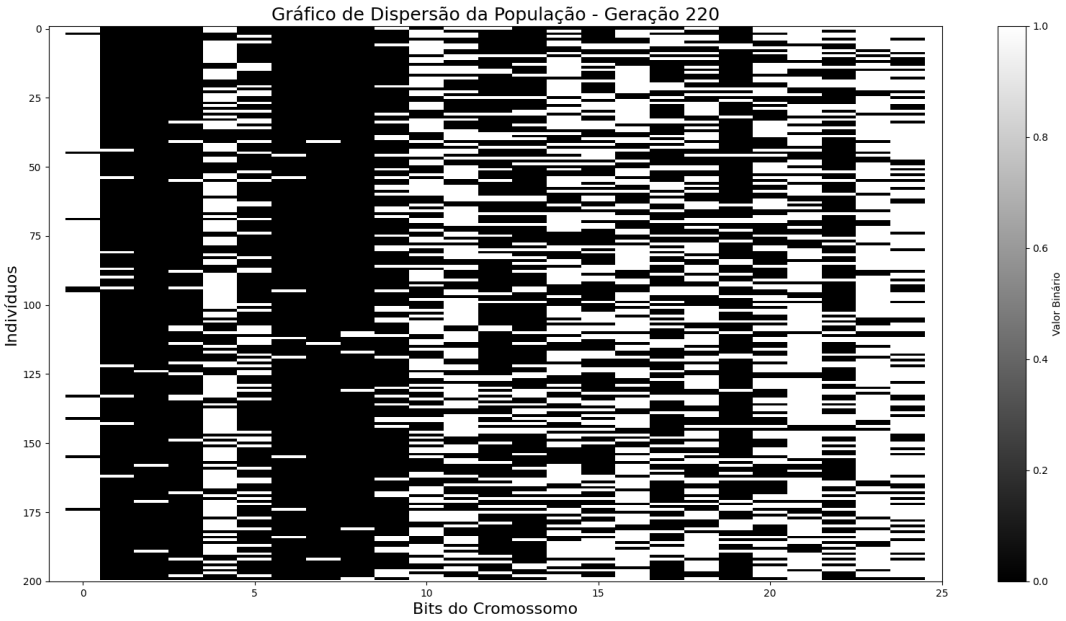

Figure 29: Matrizes 220

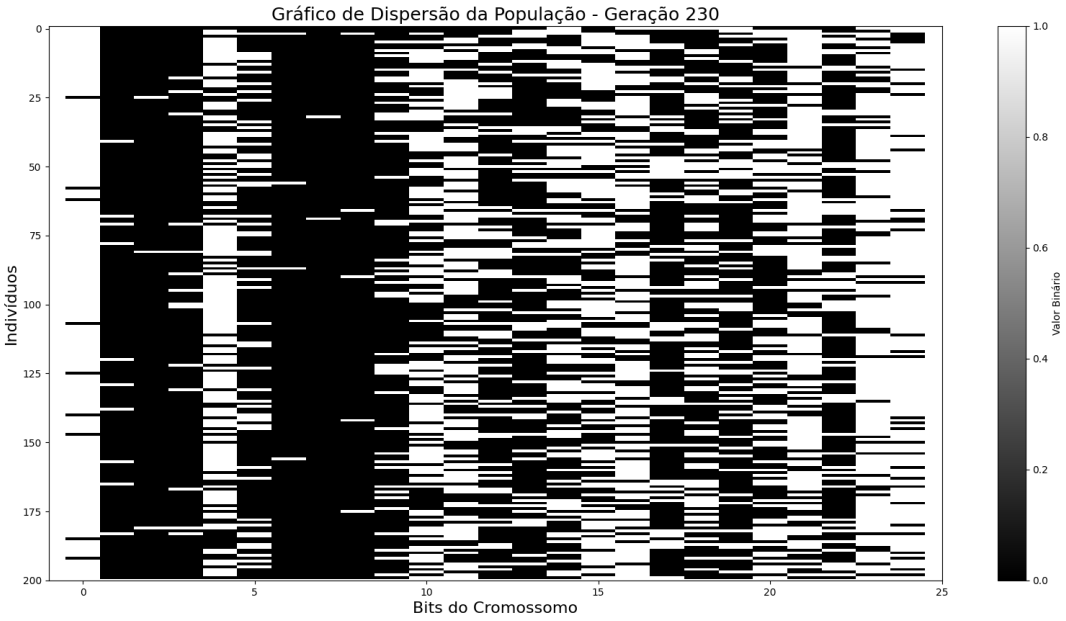

Figure 30: Matrizes 230

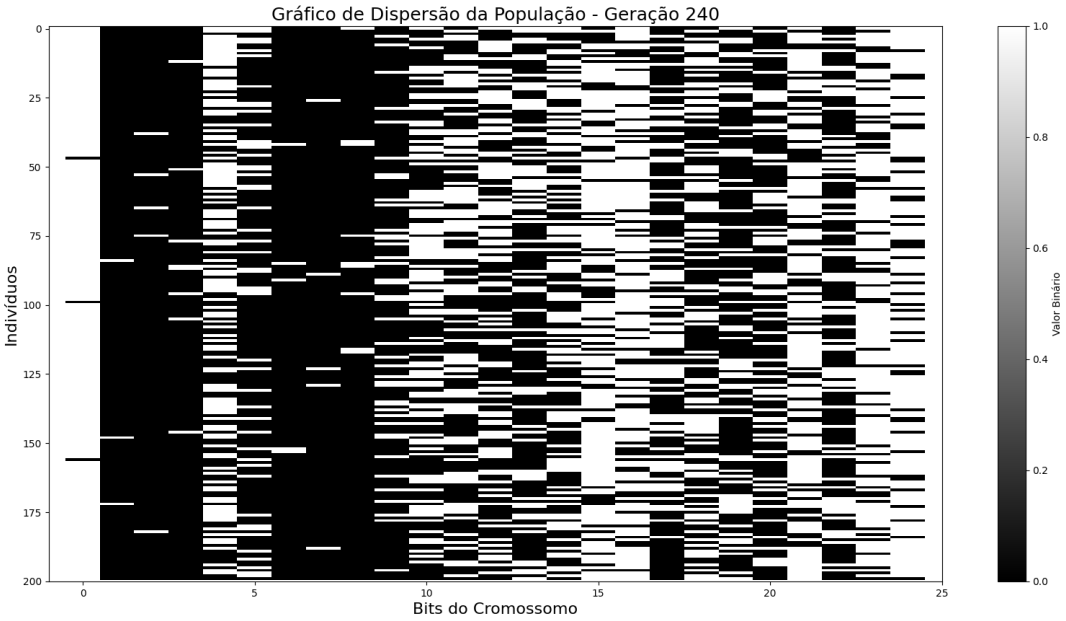

Figure 31: Matrizes 240

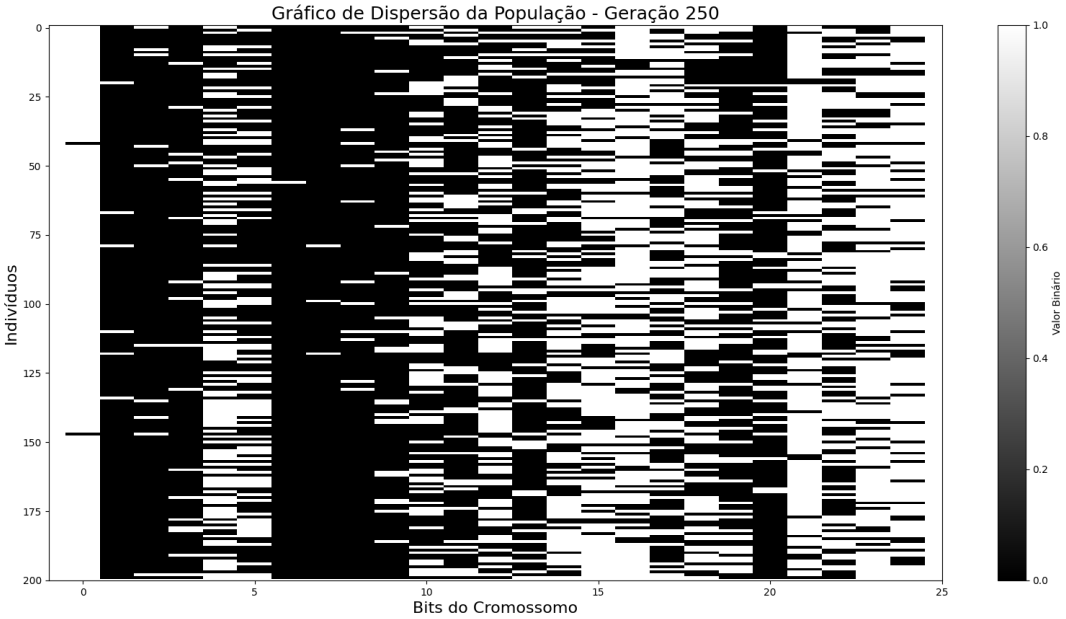

Figure 32: Matrizes 250

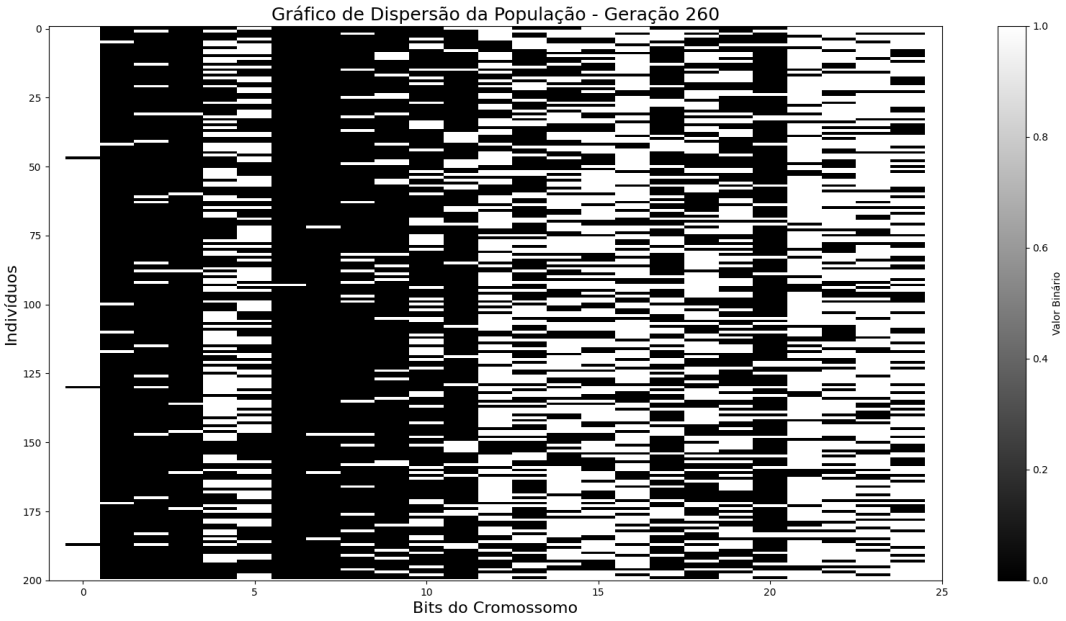

Figure 33: Matrizes 260

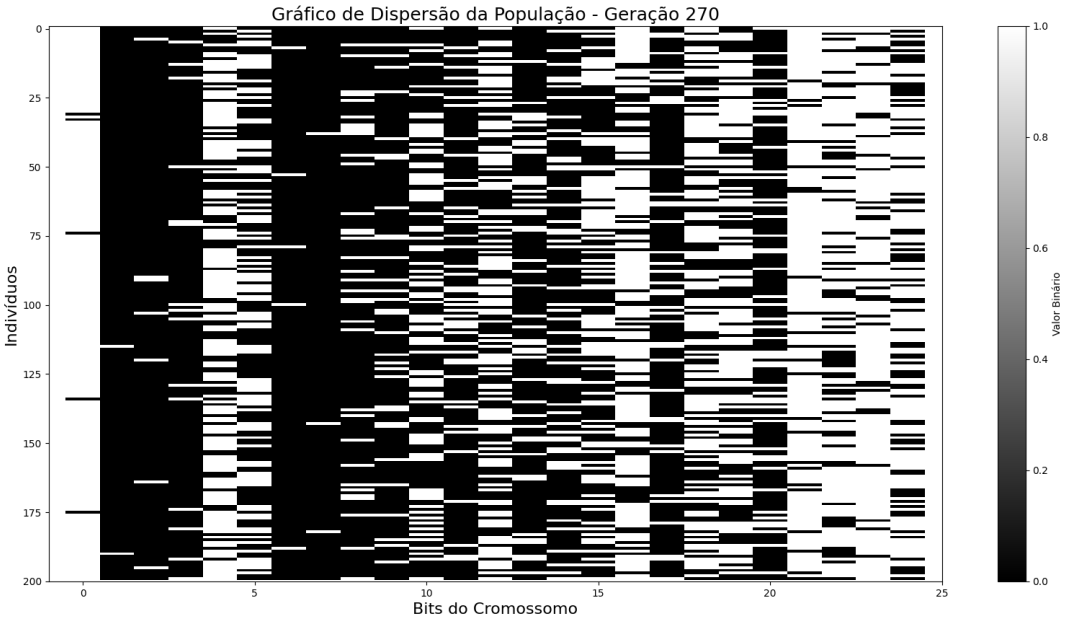

Figure 34: Matrizes 270

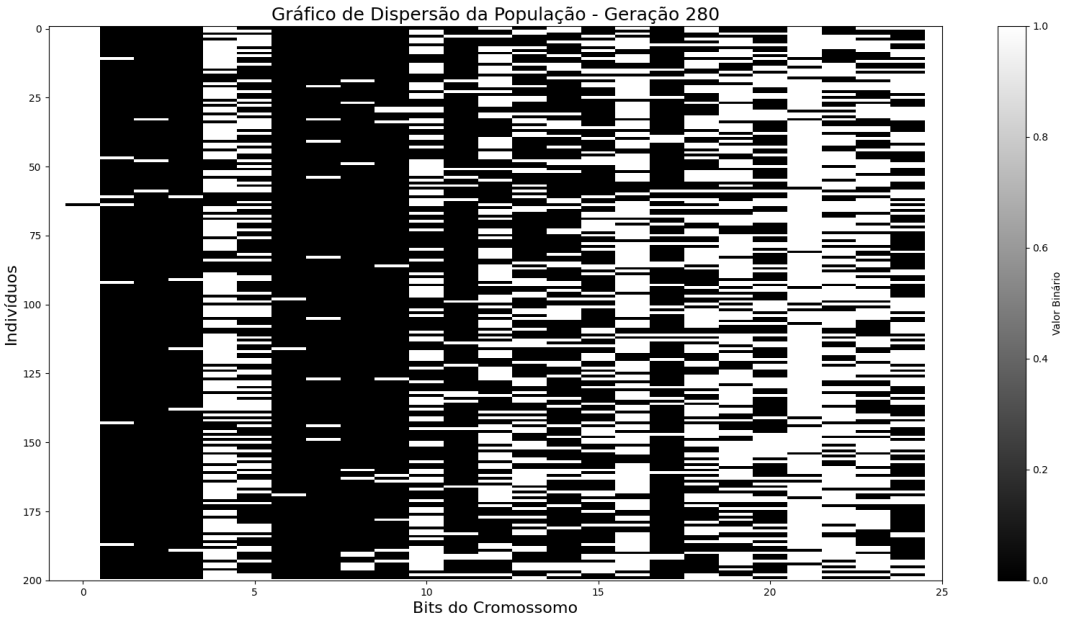

Figure 35: Matrizes 280

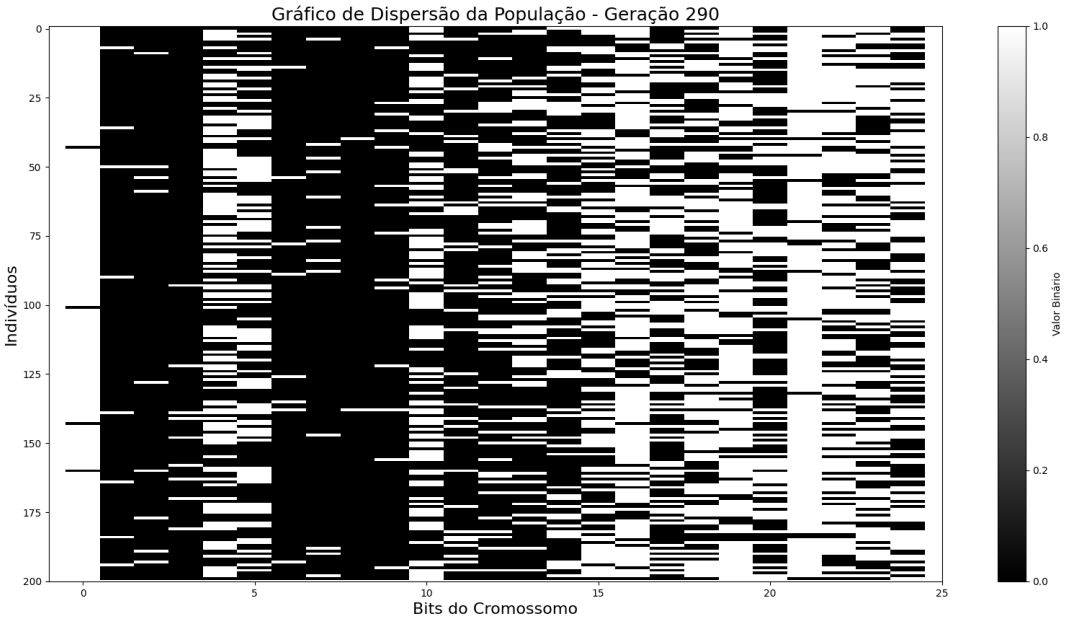

Figure 36: Matrizes 290

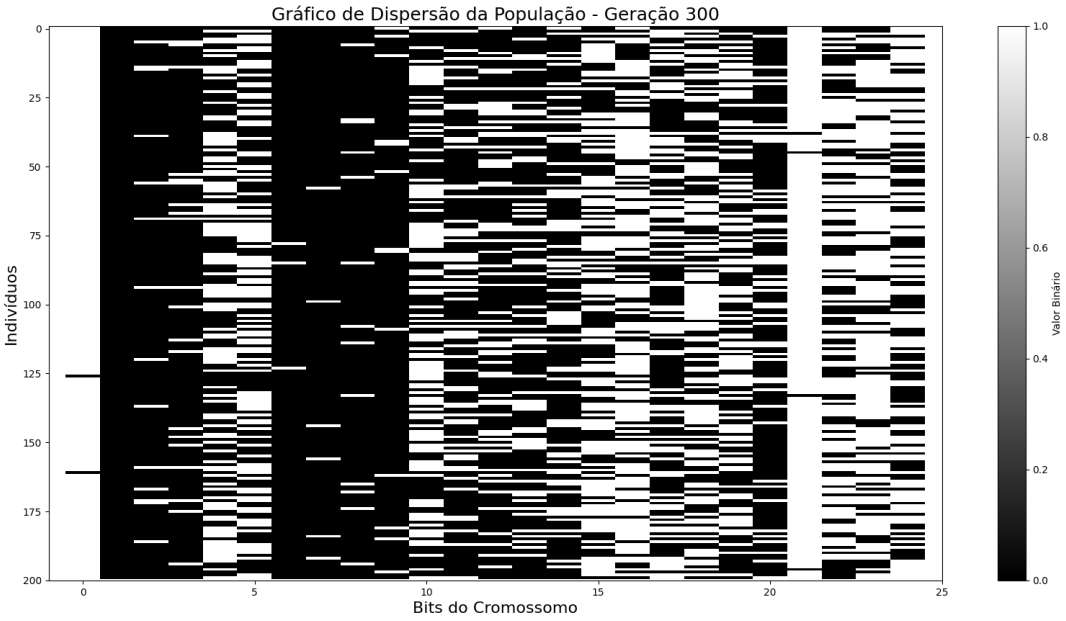

Figure 37: Matrizes 300

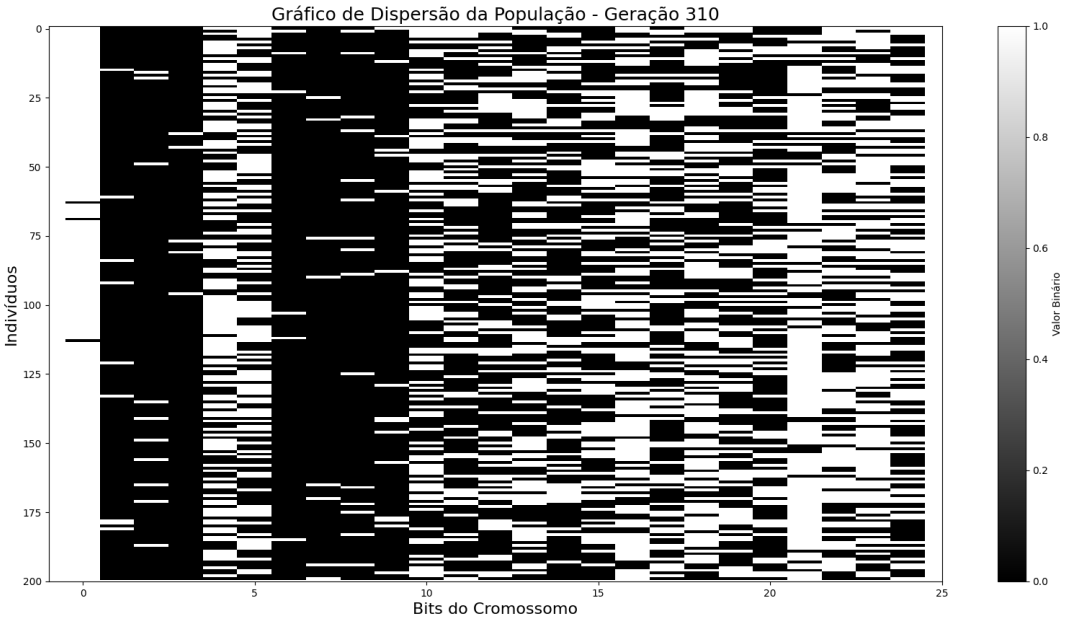

Figure 38: Matrizes 310

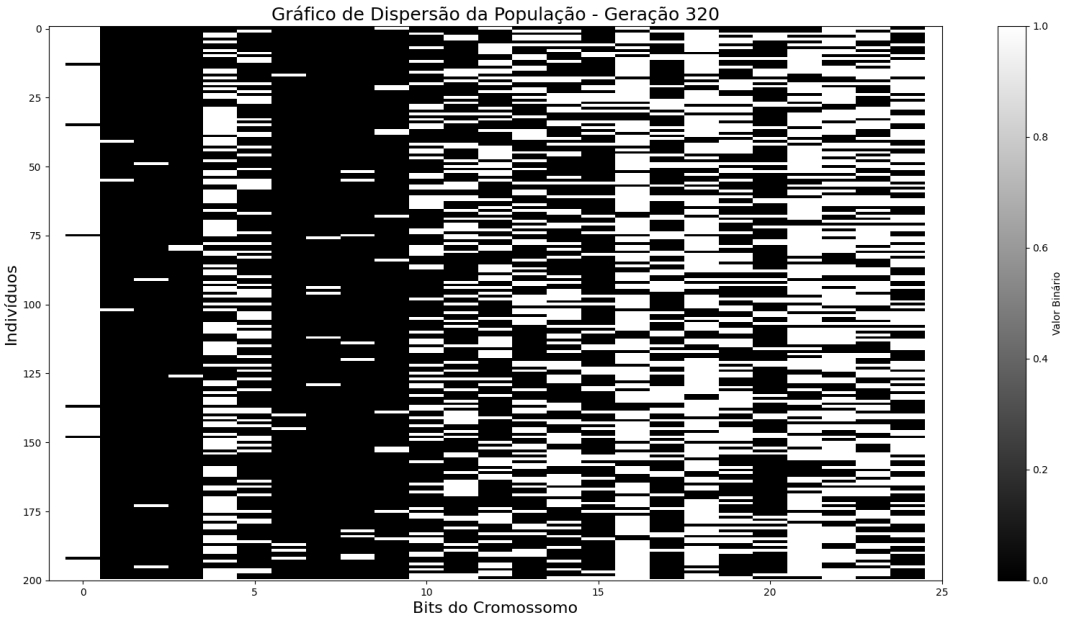

Figure 39: Matrizes 320

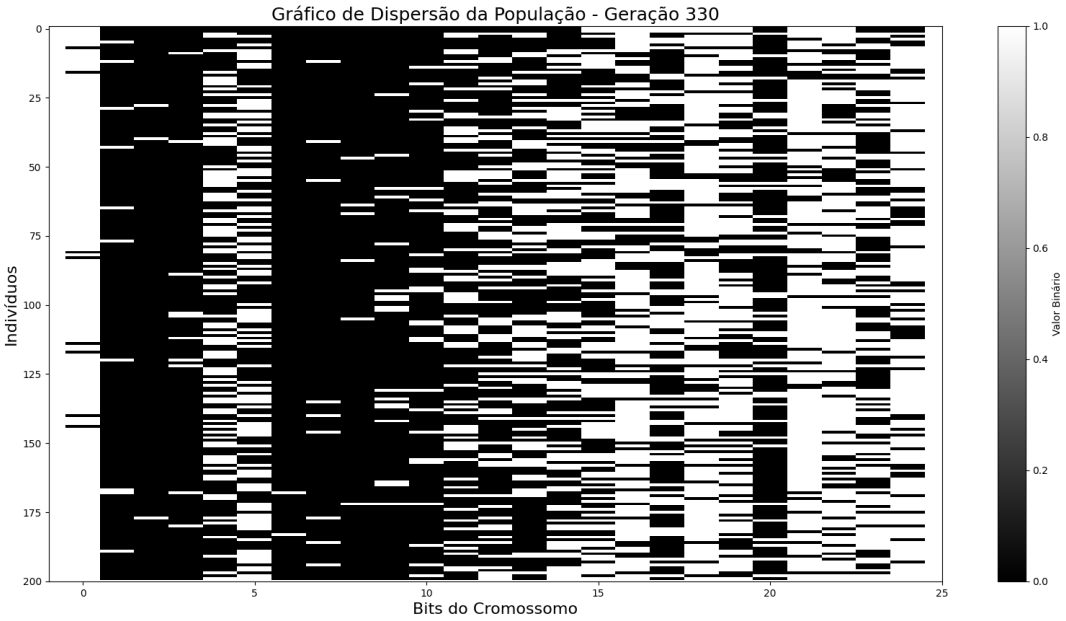

Figure 40: Matrizes 330

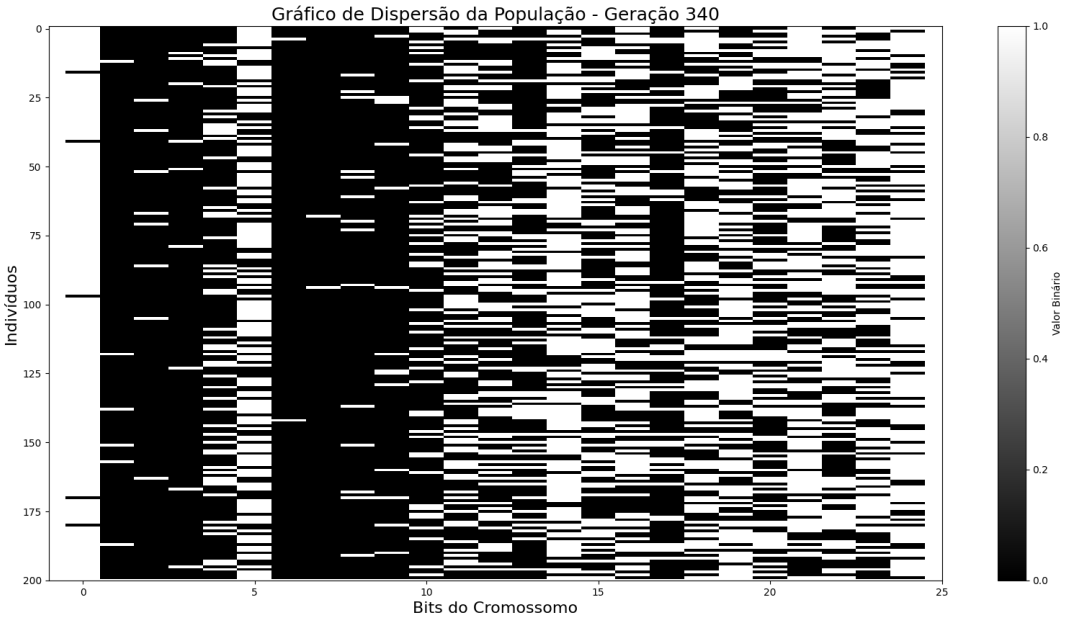

Figure 41: Matrizes 340

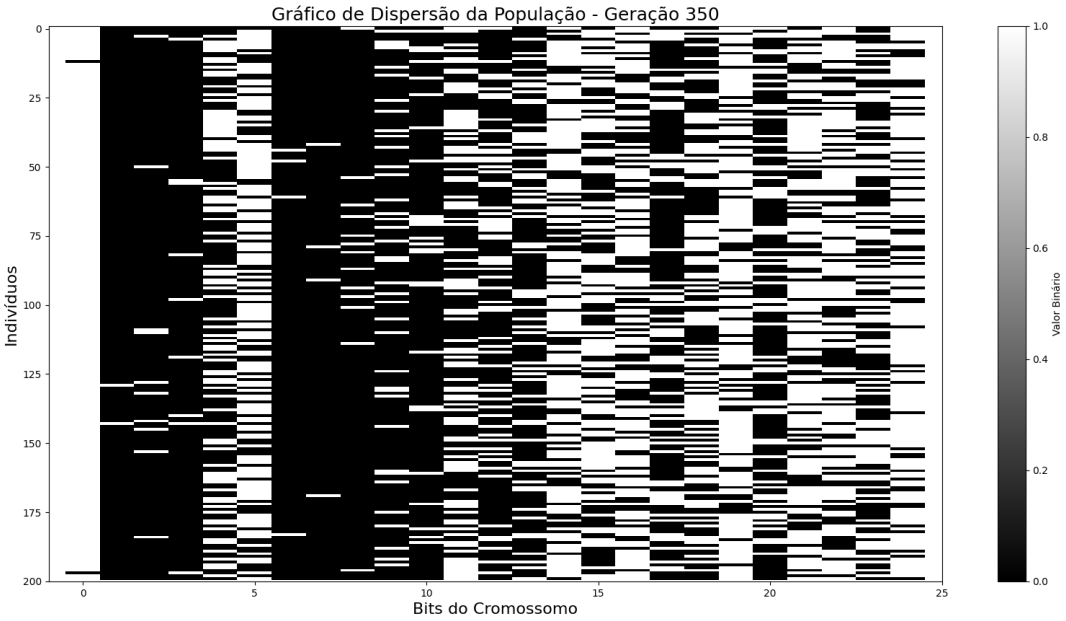

Figure 42: Matrizes 350

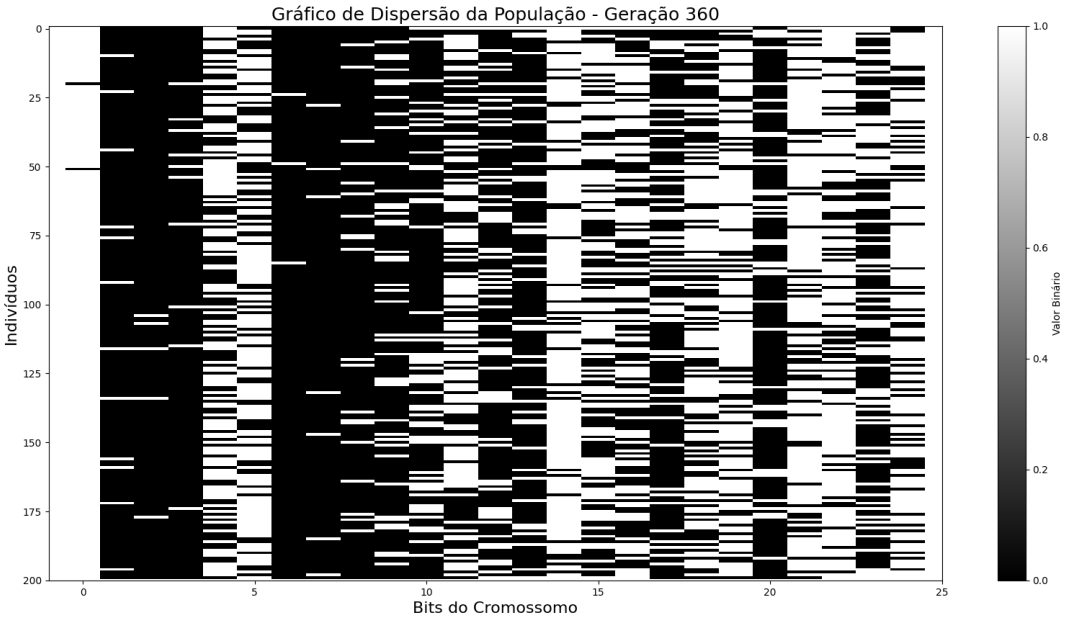

Figure 43: Matrizes 360

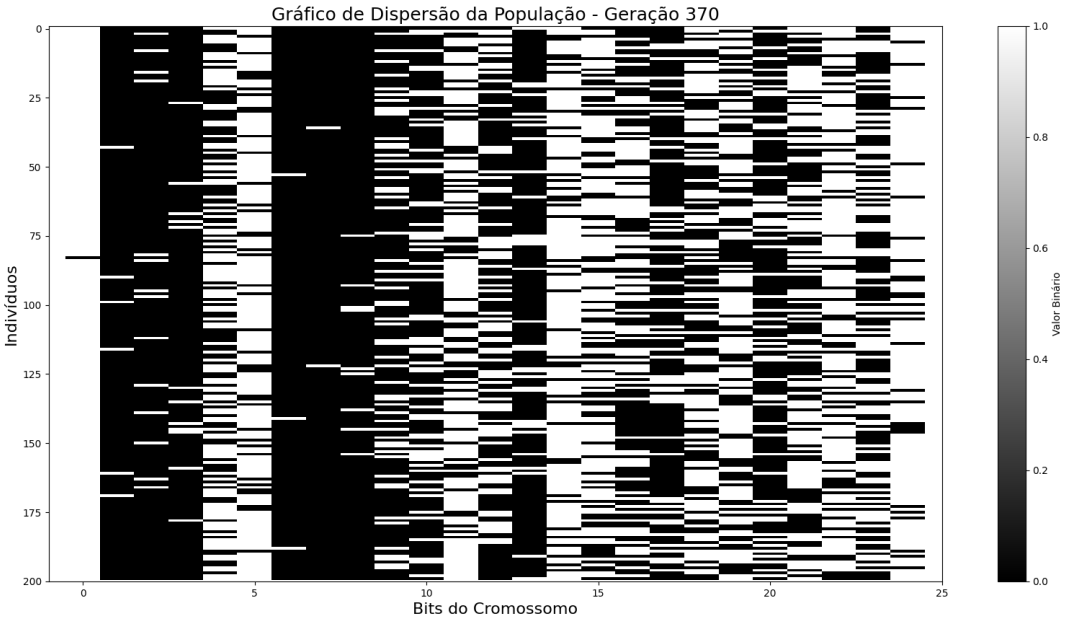

Figure 44: Matrizes 370

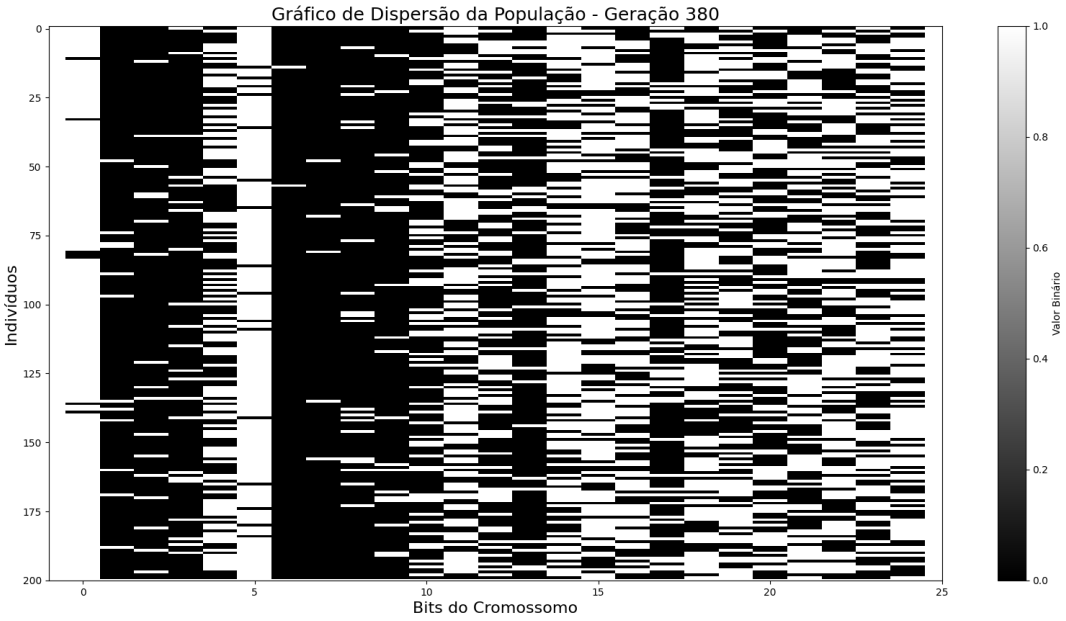

Figure 45: Matrizes 380

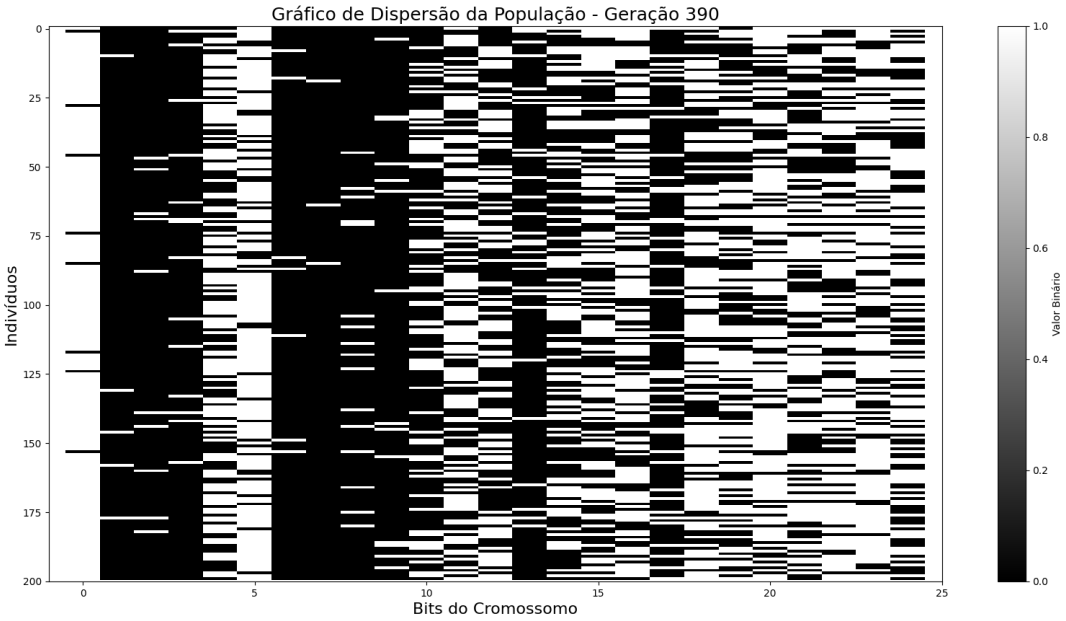

Figure 46: Matrizes 390

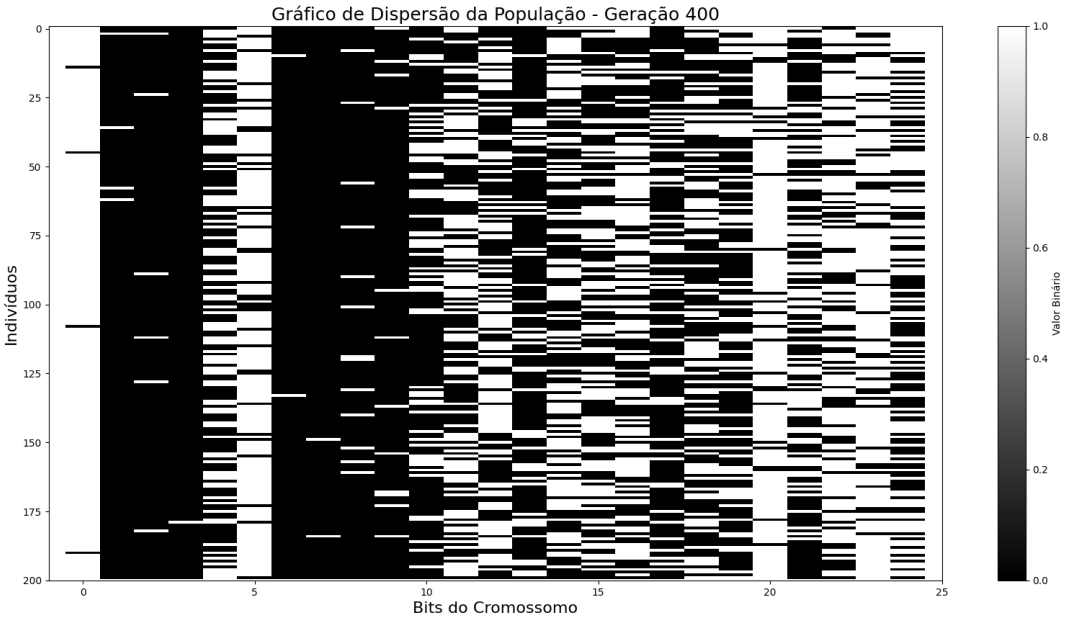

Figure 47: Matrizes 400

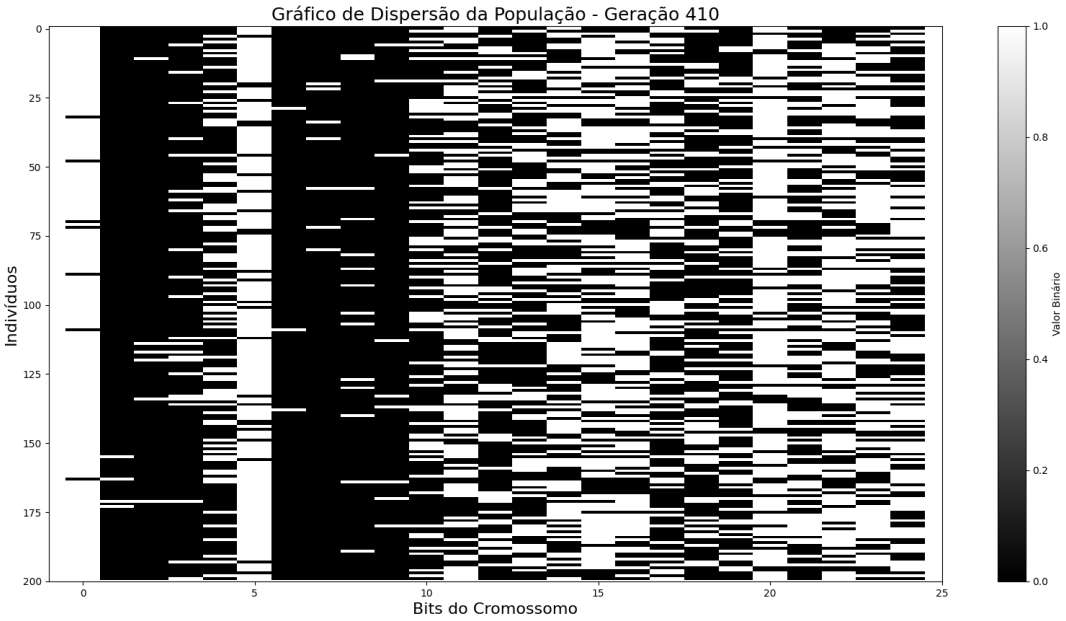

Figure 48: Matrizes 410

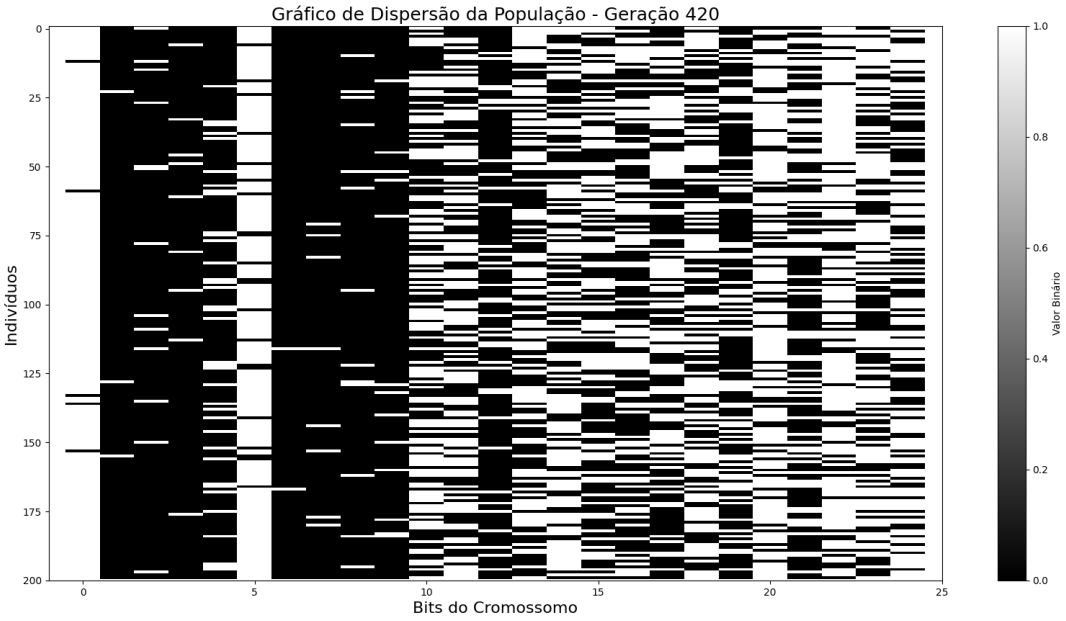

Figure 49: Matrizes 420

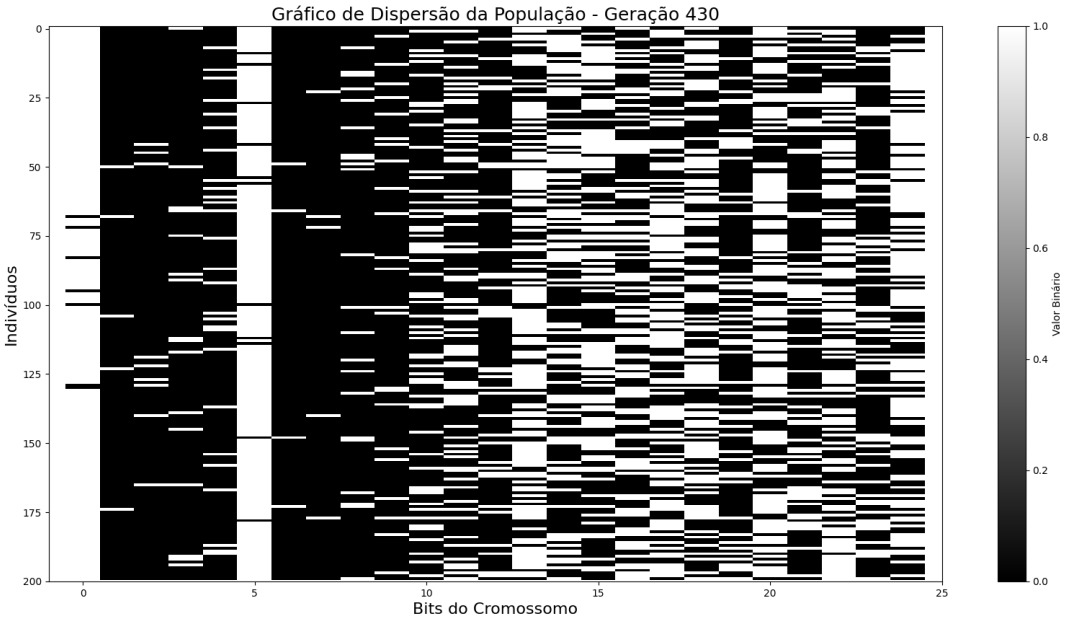

Figure 50: Matrizes 430

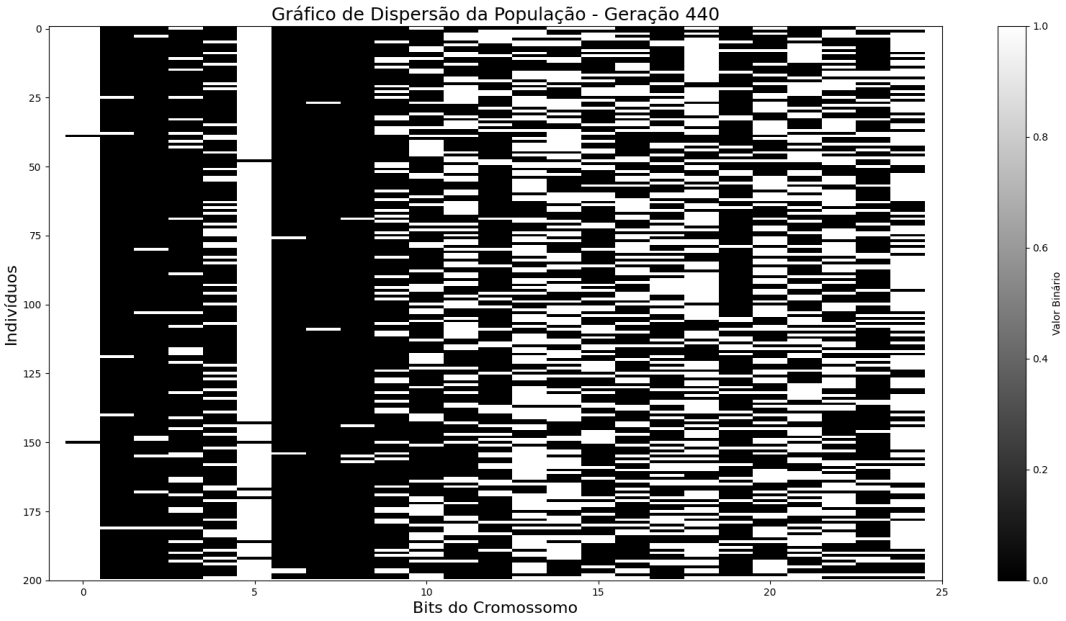

Figure 51: Matrizes 440

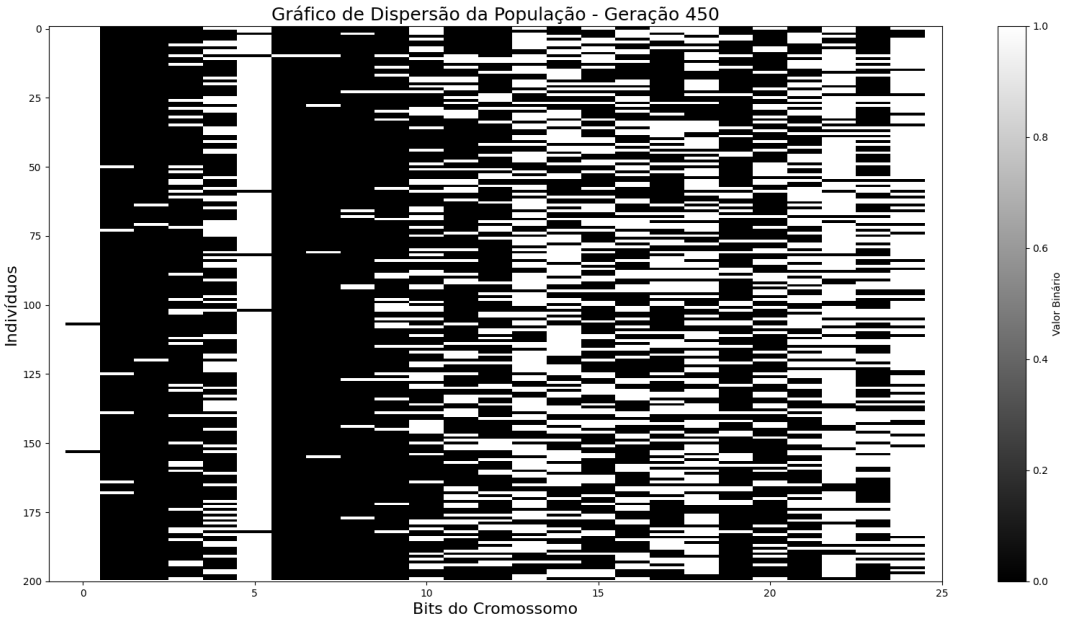

Figure 52: Matrizes 450

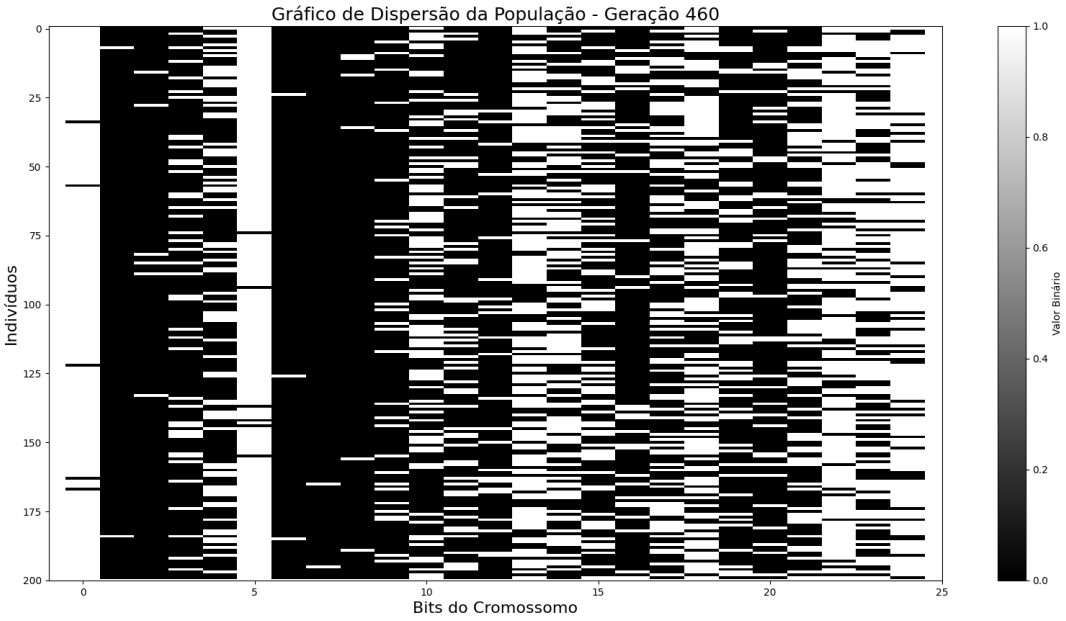

Figure 53: Matrizes 460

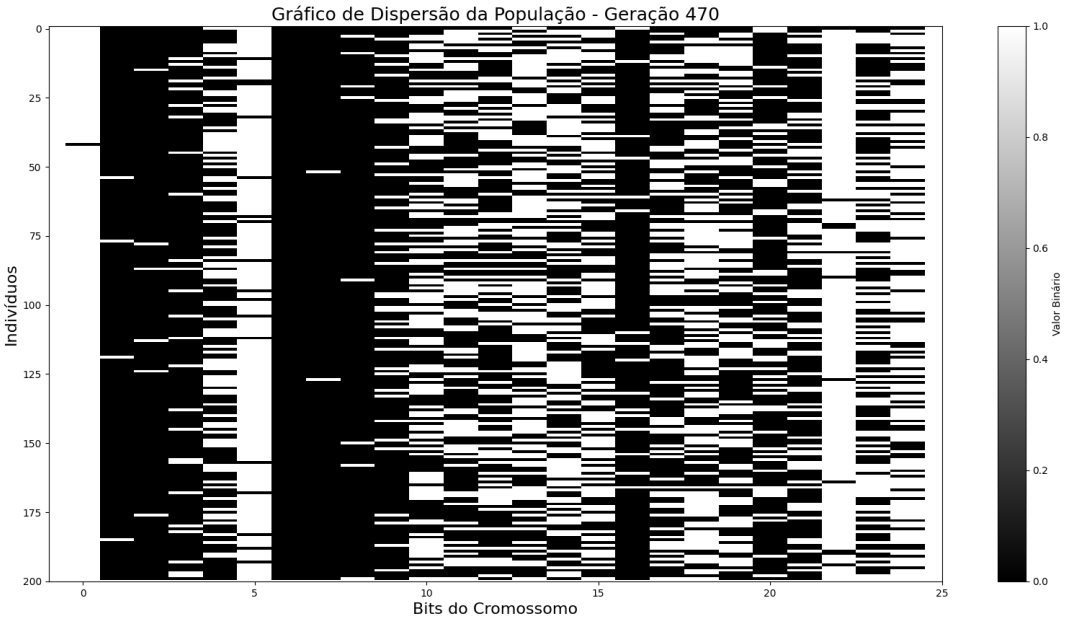

Figure 54: Matrizes 470

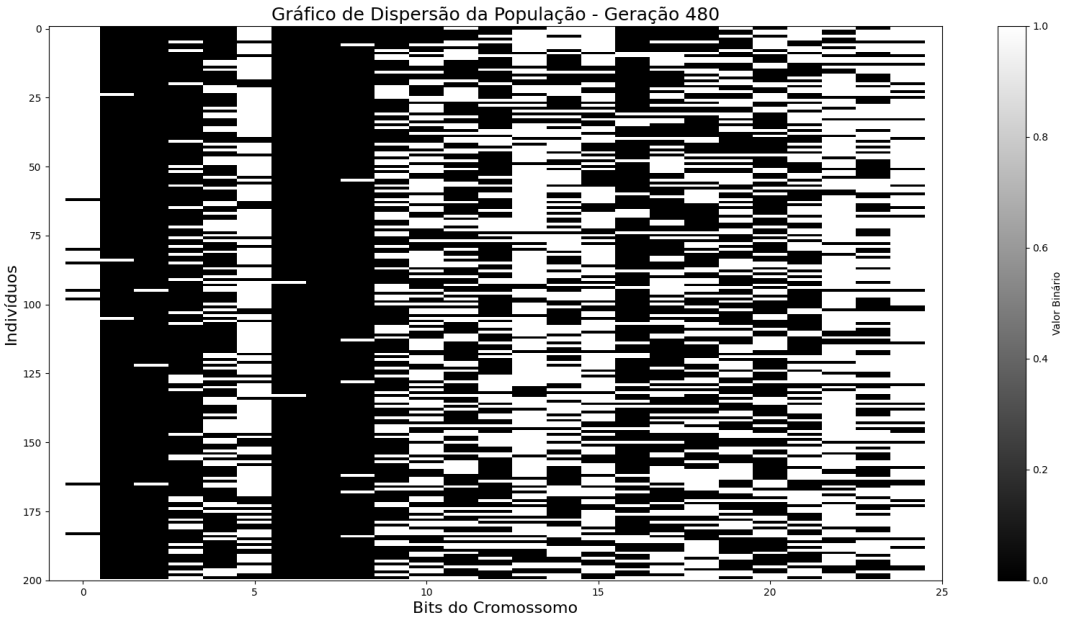

Figure 55: Matrizes 480

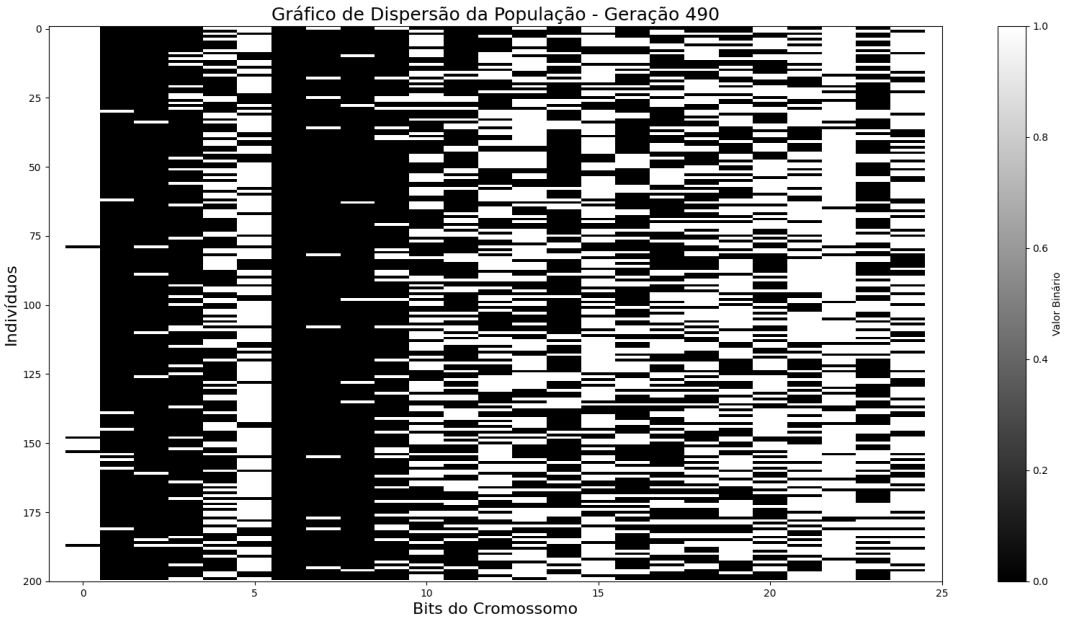

Figure 56: Matrizes 490

B.3 Gráficos 3D da Função F6 com População (Representação Binária)

Aqui, exibimos gráficos tridimensionais que representam a função F6 em relação à população gerada pelo algoritmo genético utilizando a representação binária. Esses gráficos permitem visualizar como os indivíduos se distribuem no espaço de busca e como a qualidade das soluções evolui ao longo das gerações. A análise desses gráficos é crucial para entender o comportamento do algoritmo e a eficácia da representação binária na busca pela solução ótima.

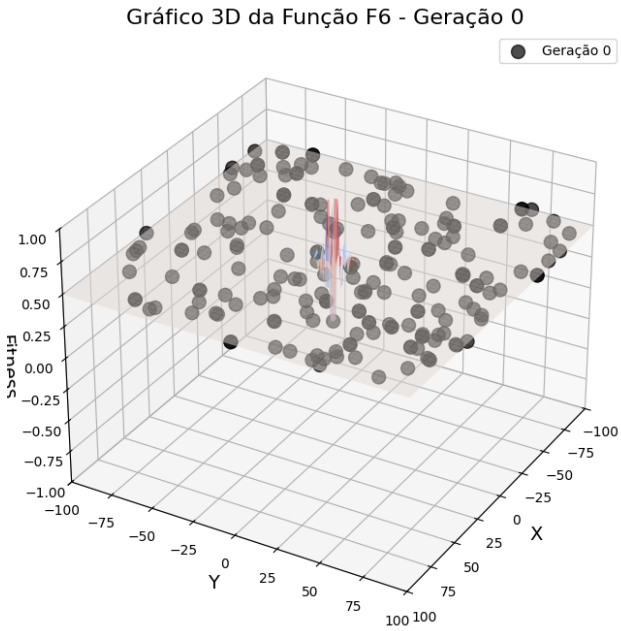

Figure 57: F6 - Binário 0

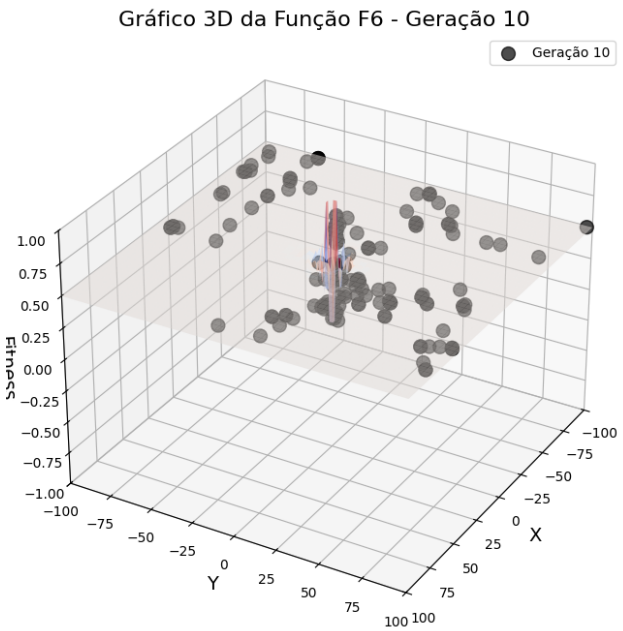

Figure 58: F6 - Binário 10

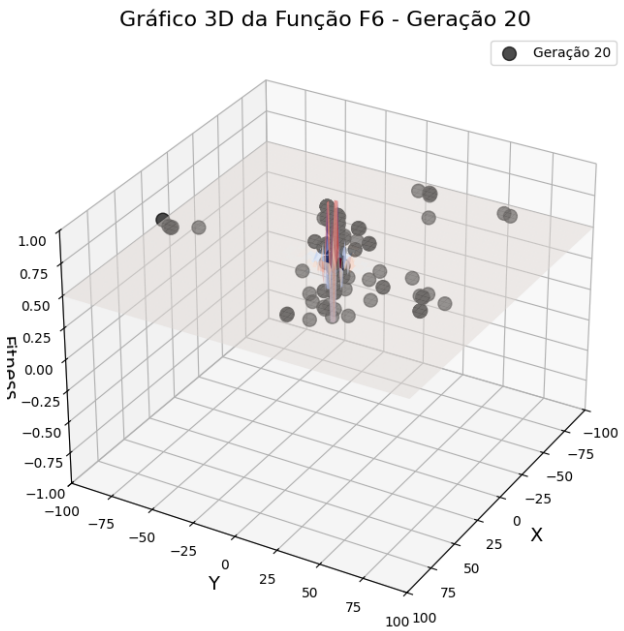

Figure 59: F6 - Binário 20

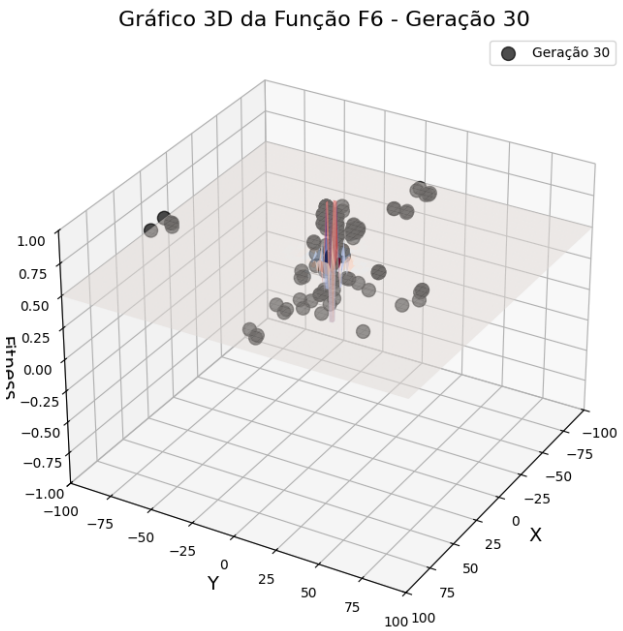

Figure 60: F6 - Binário 30

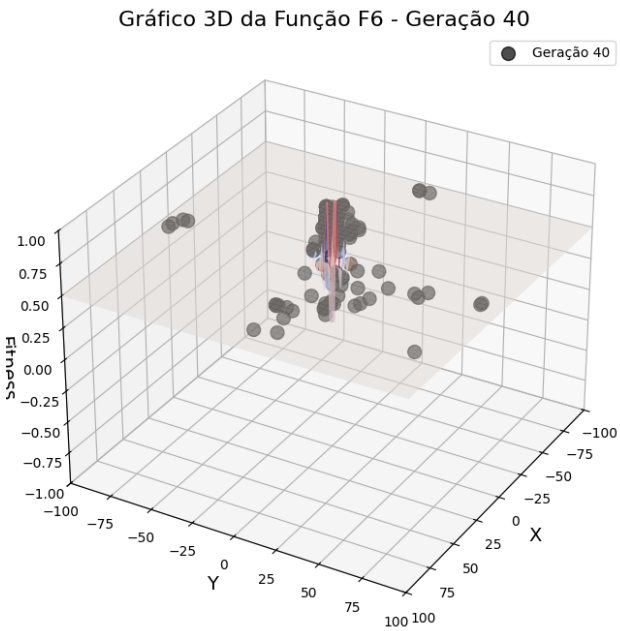

Figure 61: F6 - Binário 40

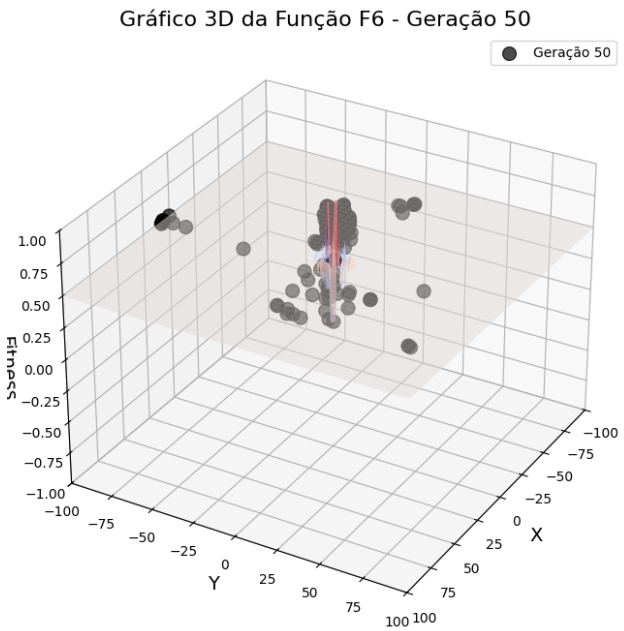

Figure 62: F6 - Binário 50

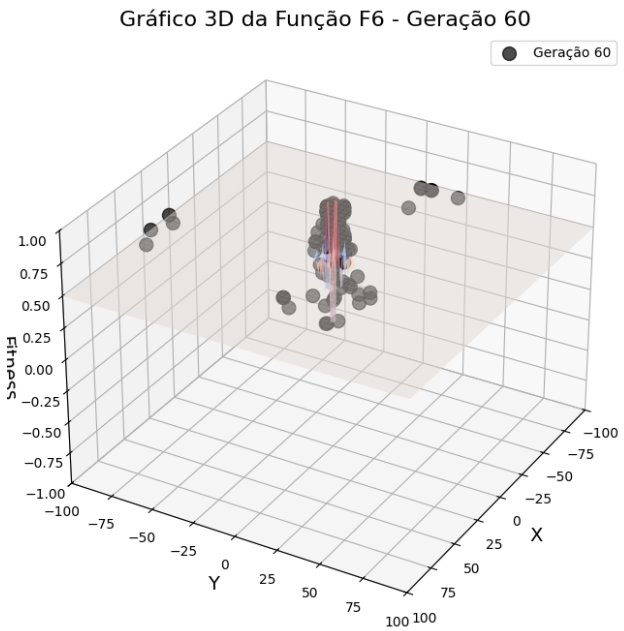

Figure 63: F6 - Binário 60

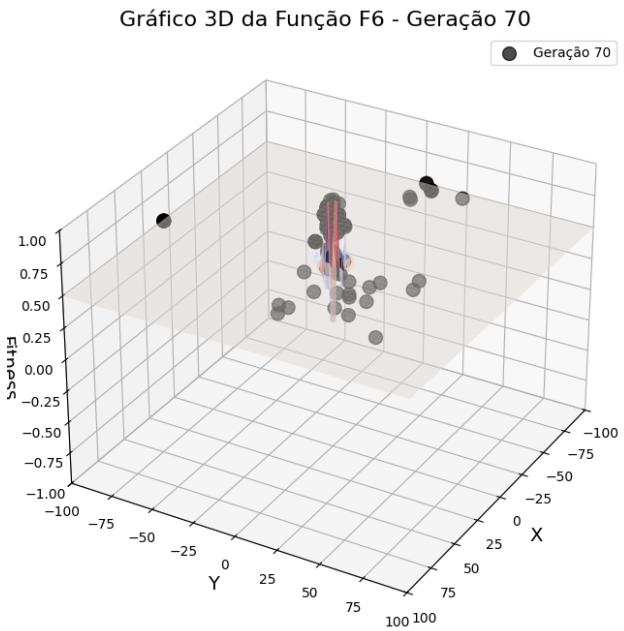

Figure 64: F6 - Binário 70

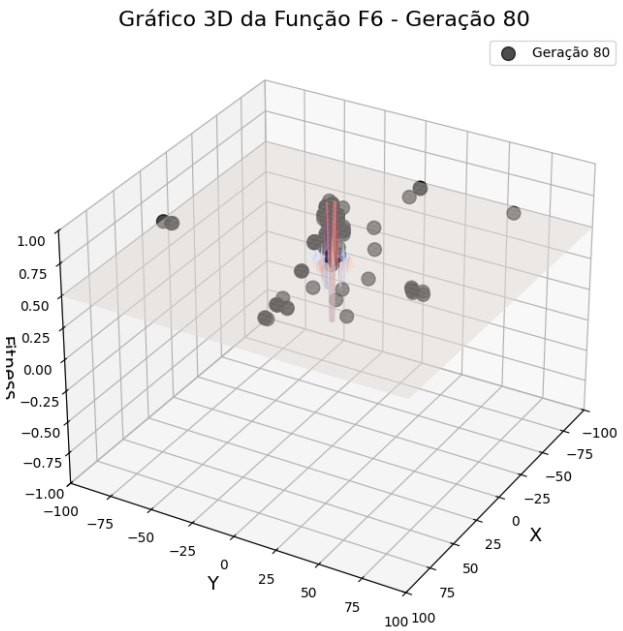

Figure 65: F6 - Binário 80

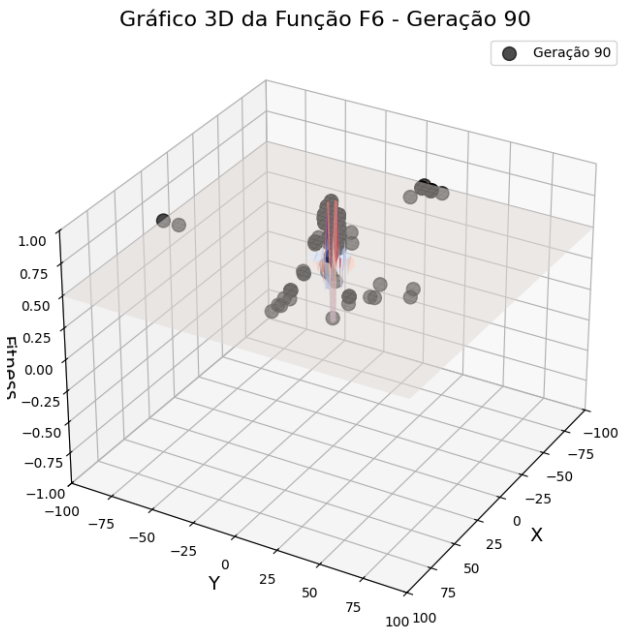

Figure 66: F6 - Binário 90

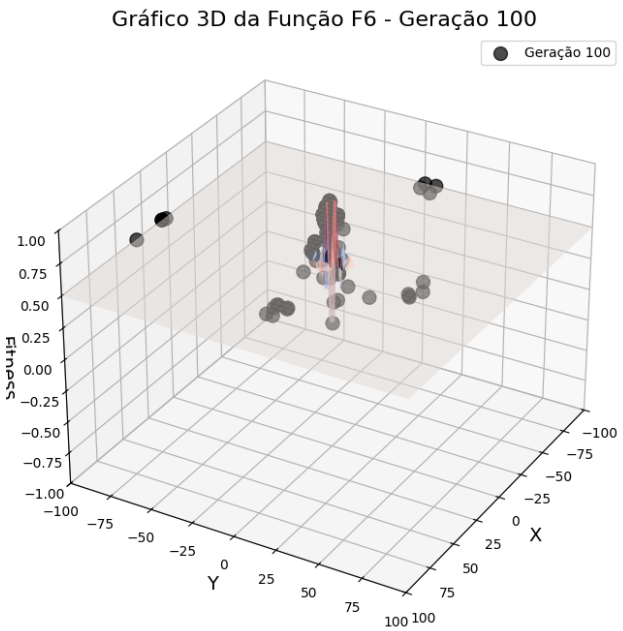

Figure 67: F6 - Binário 100

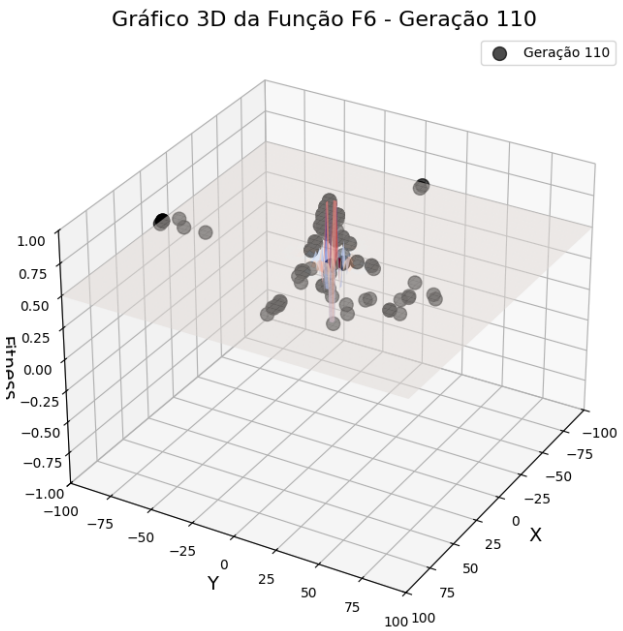

Figure 68: F6 - Binário 110

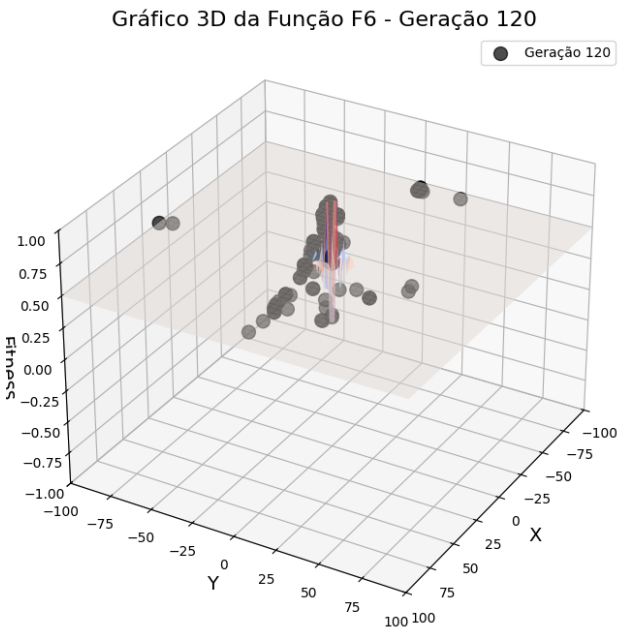

Figure 69: F6 - Binário 120

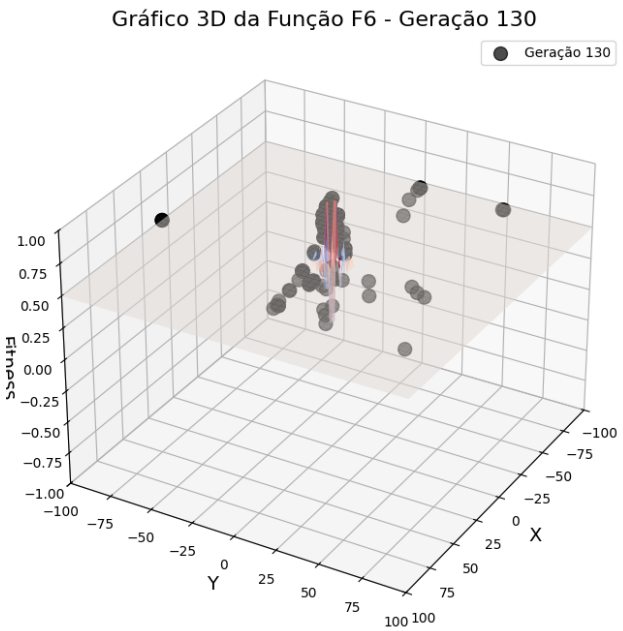

Figure 70: F6 - Binário 130

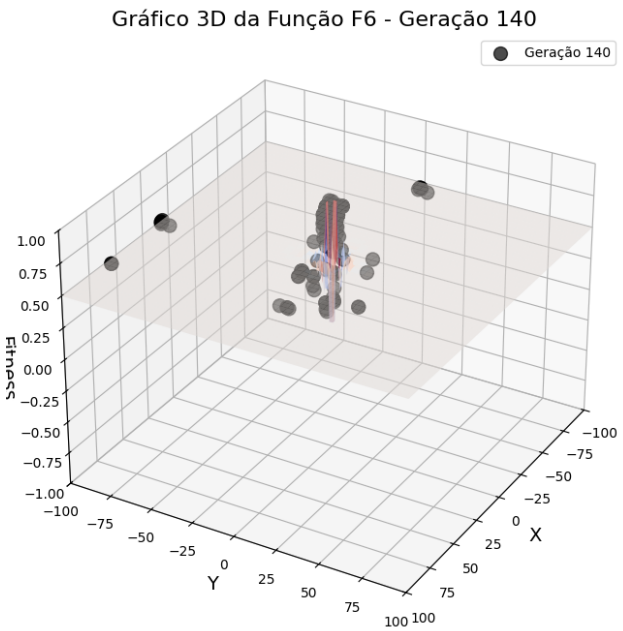

Figure 71: F6 - Binário 140

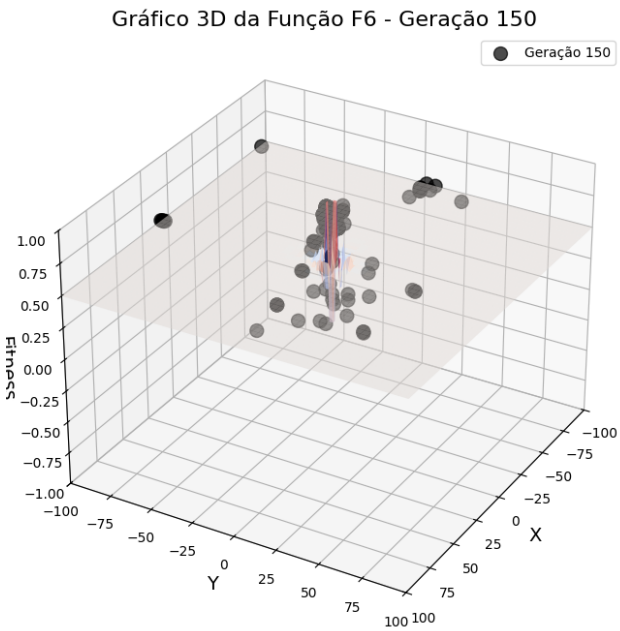

Figure 72: F6 - Binário 150

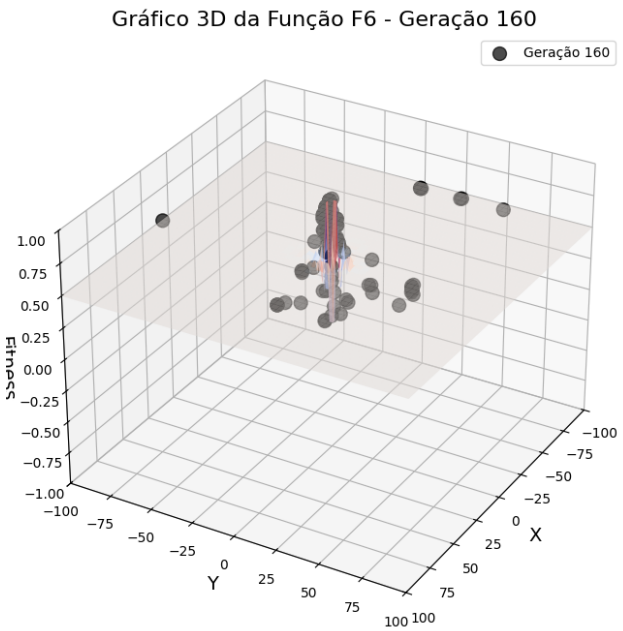

Figure 73: F6 - Binário 160

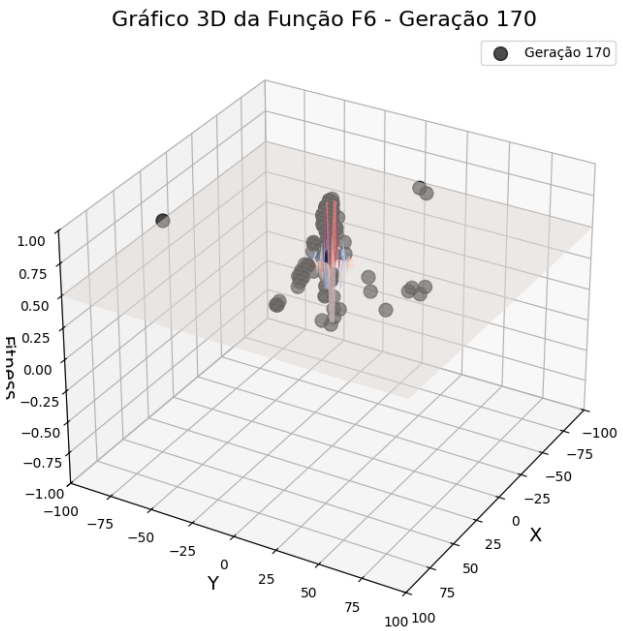

Figure 74: F6 - Binário 170

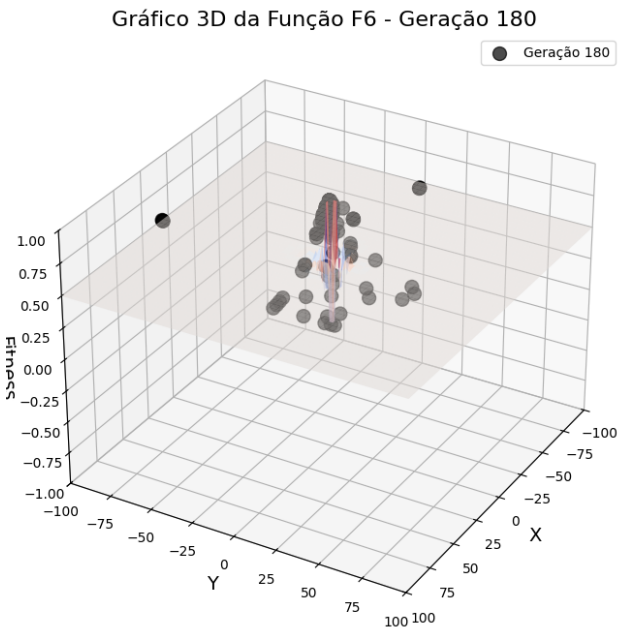

Figure 75: F6 - Binário 180

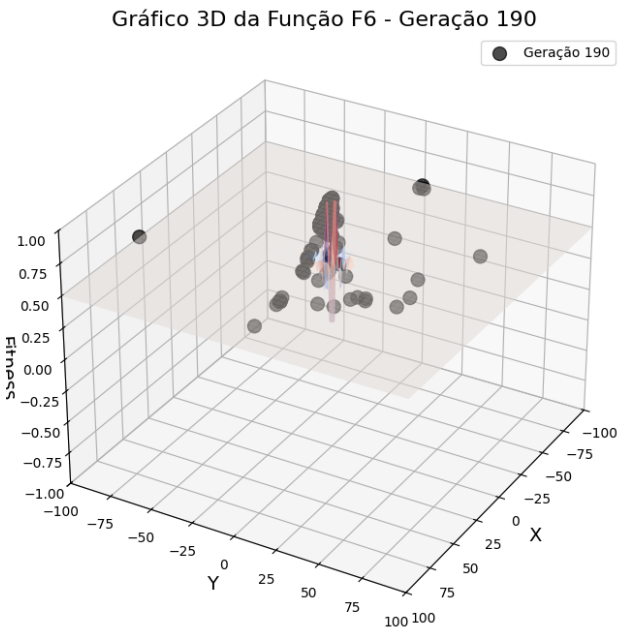

Figure 76: F6 - Binário 190

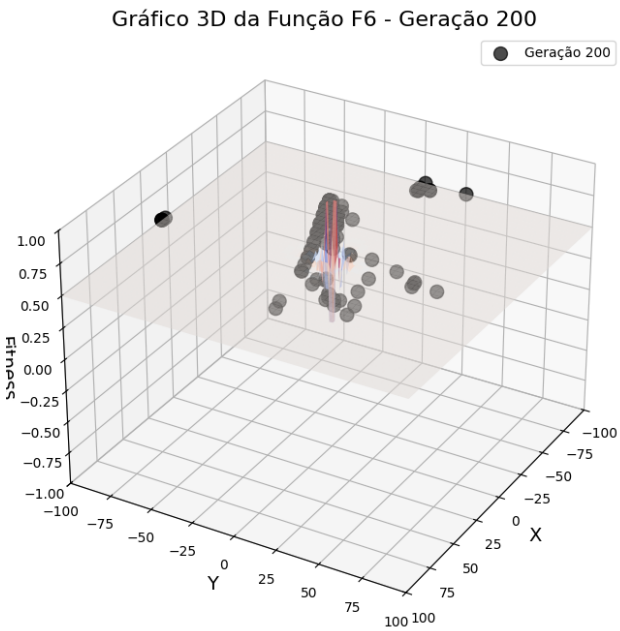

Figure 77: F6 - Binário 200

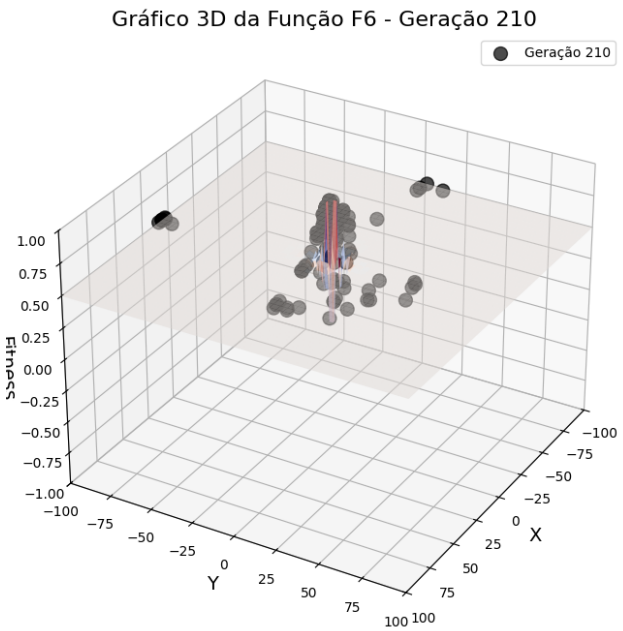

Figure 78: F6 - Binário 210

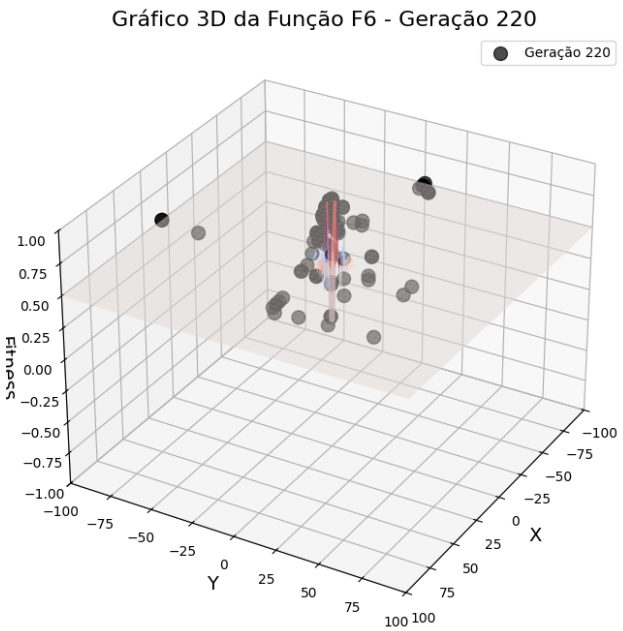

Figure 79: F6 - Binário 220

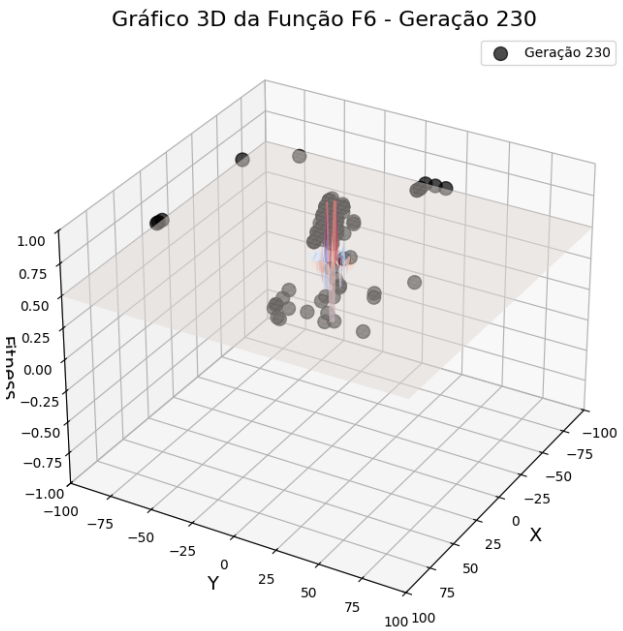

Figure 80: F6 - Binário 230

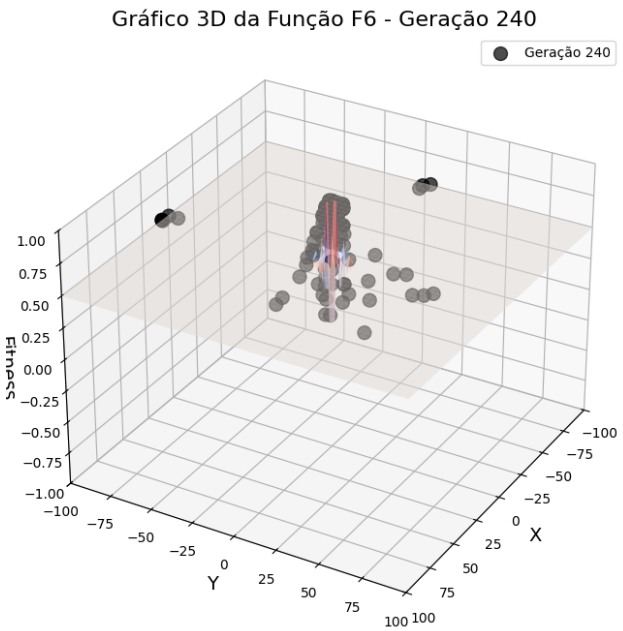

Figure 81: F6 - Binário 240

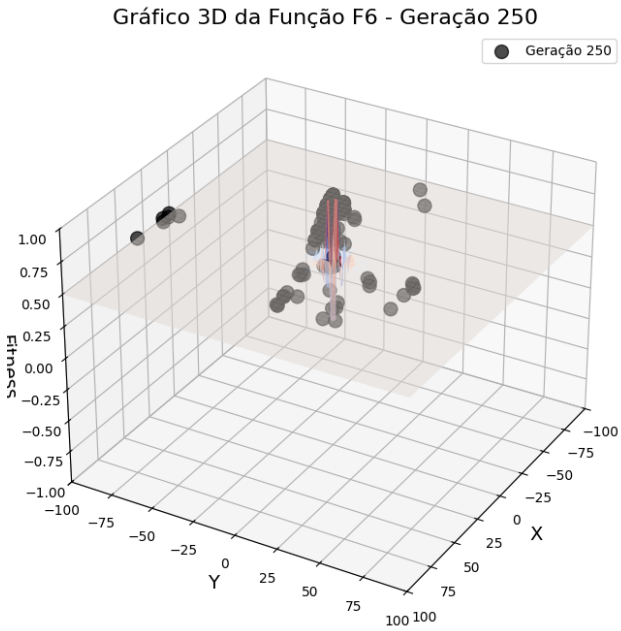

Figure 82: F6 - Binário 250

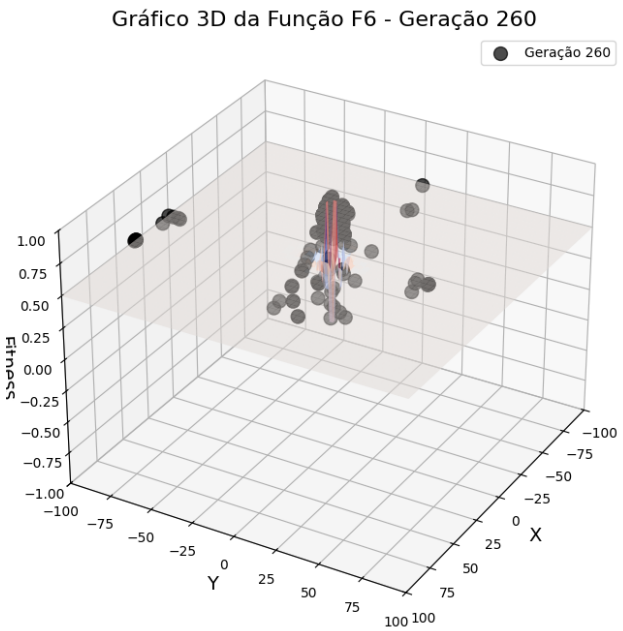

Figure 83: F6 - Binário 260

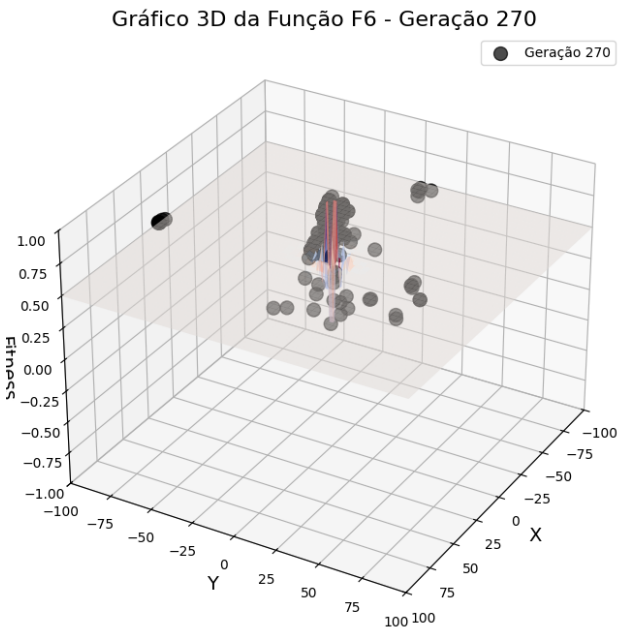

Figure 84: F6 - Binário 270

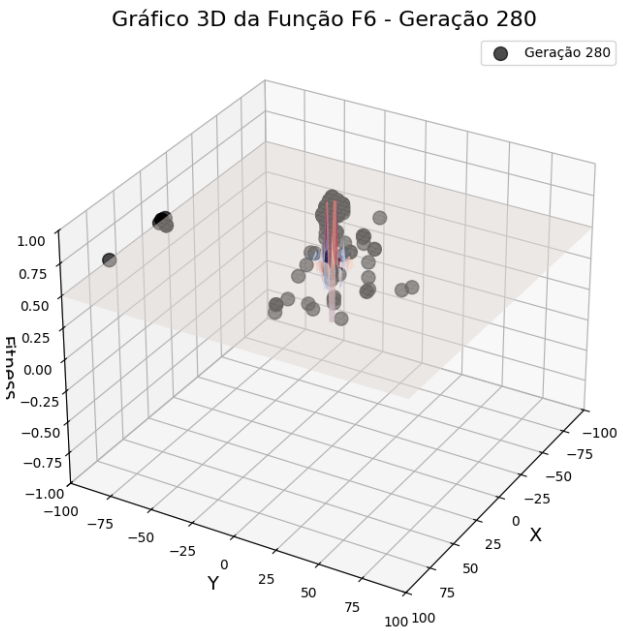

Figure 85: F6 - Binário 280

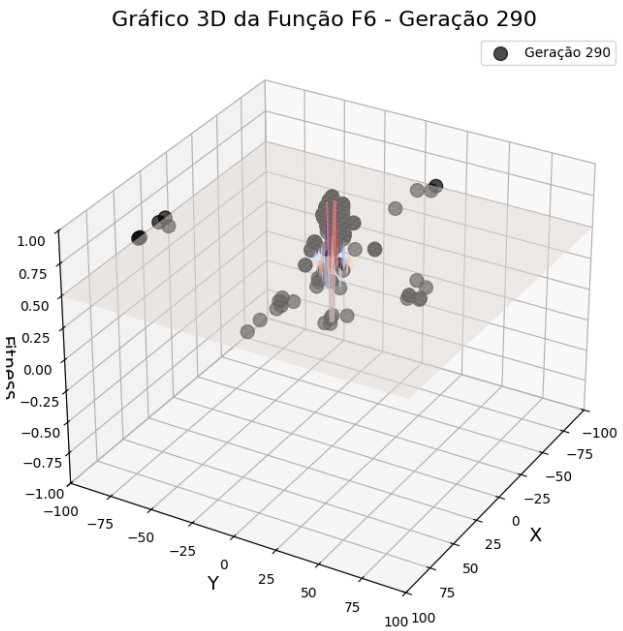

Figure 86: F6 - Binário 290

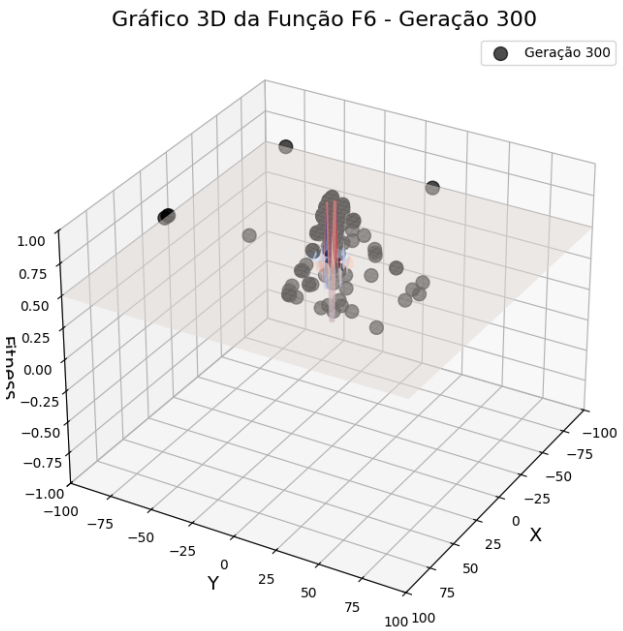

Figure 87: F6 - Binário 300

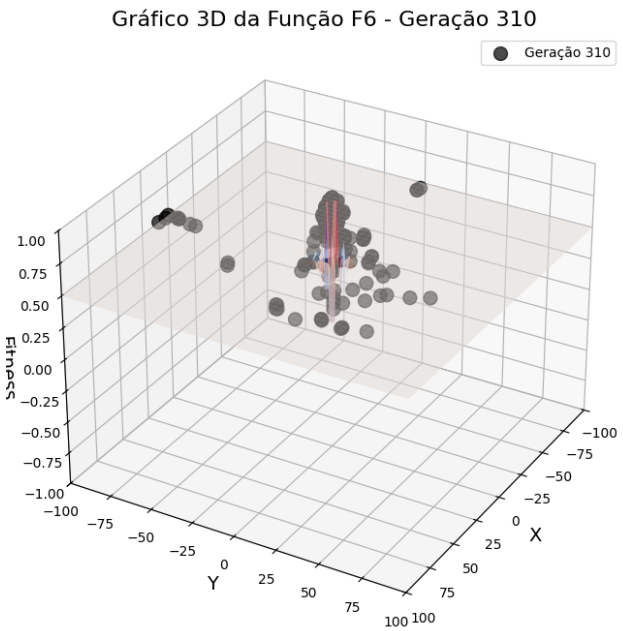

Figure 88: F6 - Binário 310

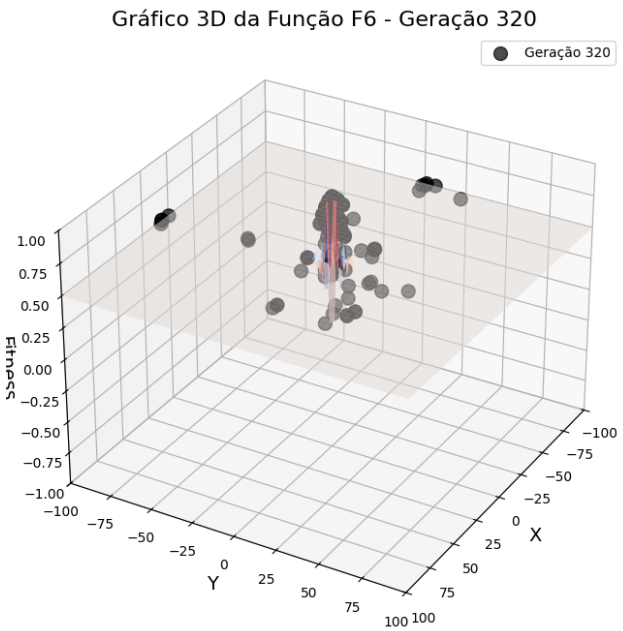

Figure 89: F6 - Binário 320

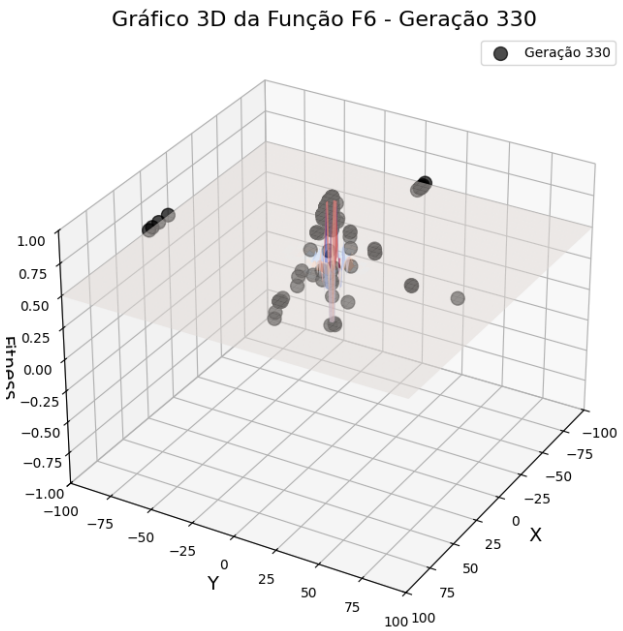

Figure 90: F6 - Binário 330

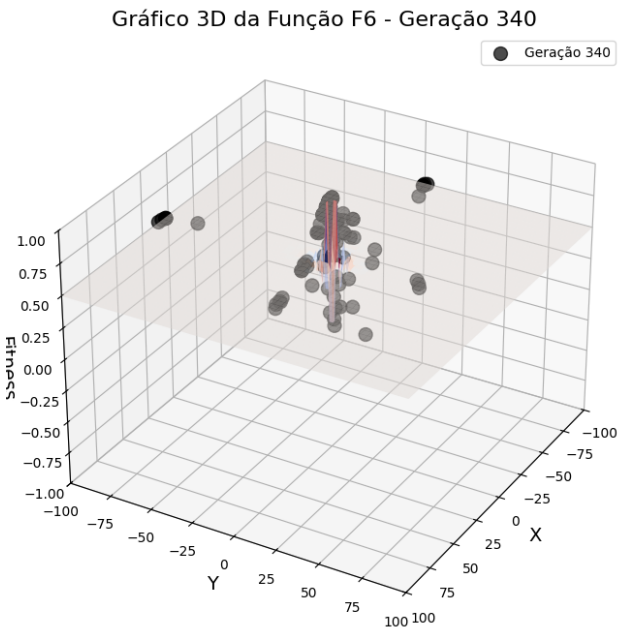

Figure 91: F6 - Binário 340

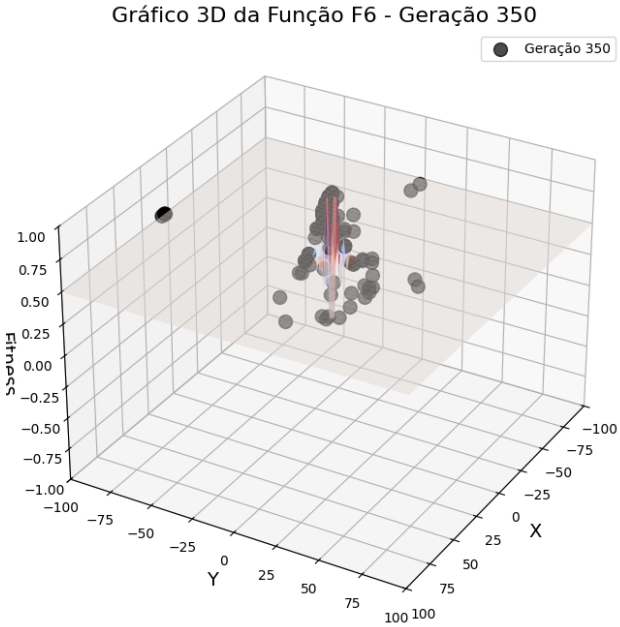

Figure 92: F6 - Binário 350

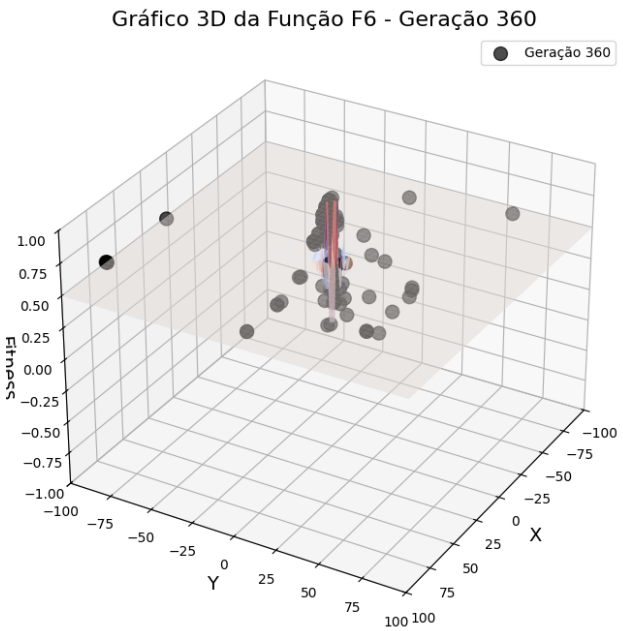

Figure 93: F6 - Binário 360

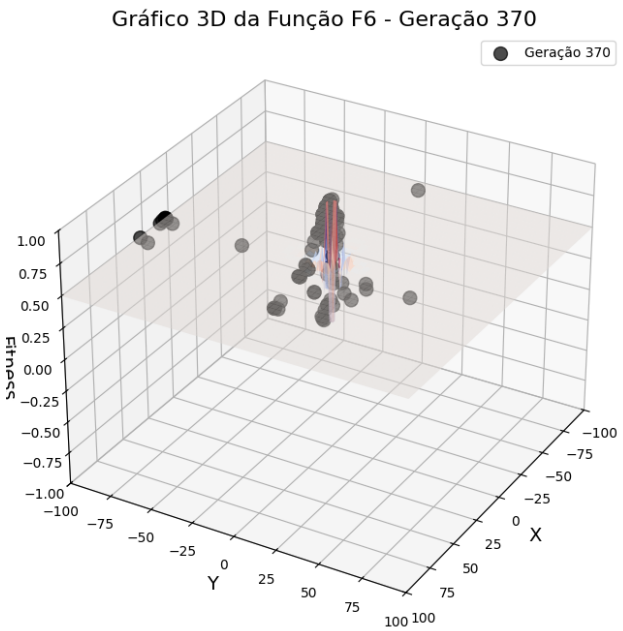

Figure 94: F6 - Binário 370

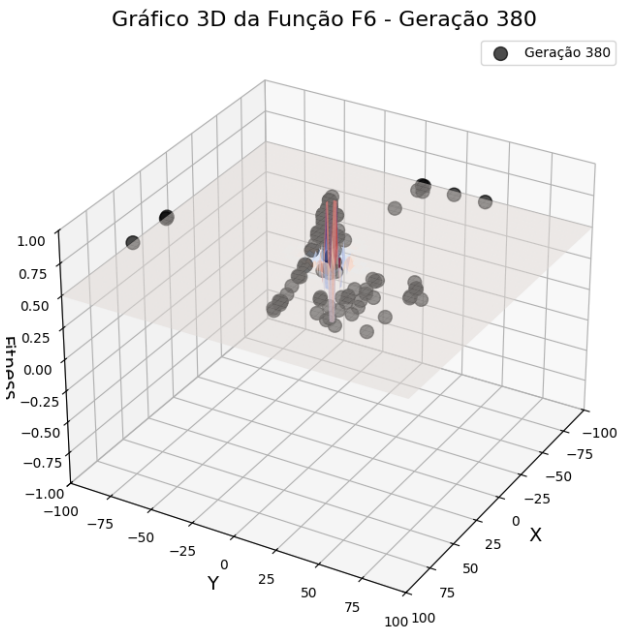

Figure 95: F6 - Binário 380

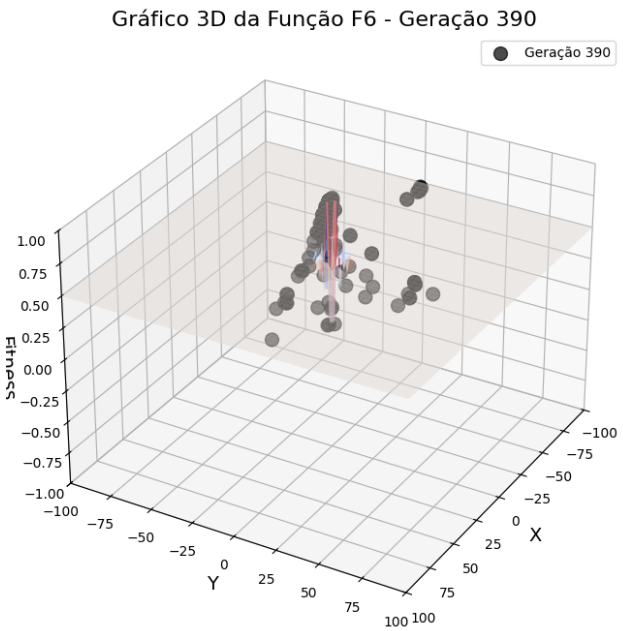

Figure 96: F6 - Binário 390

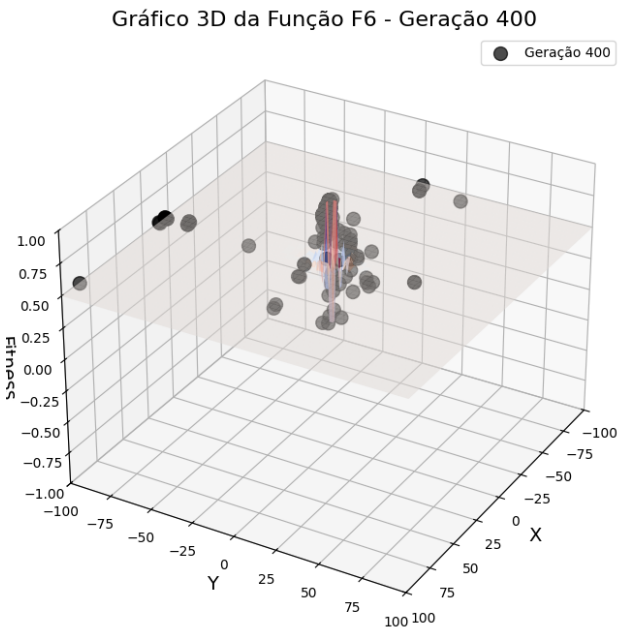

Figure 97: F6 - Binário 400

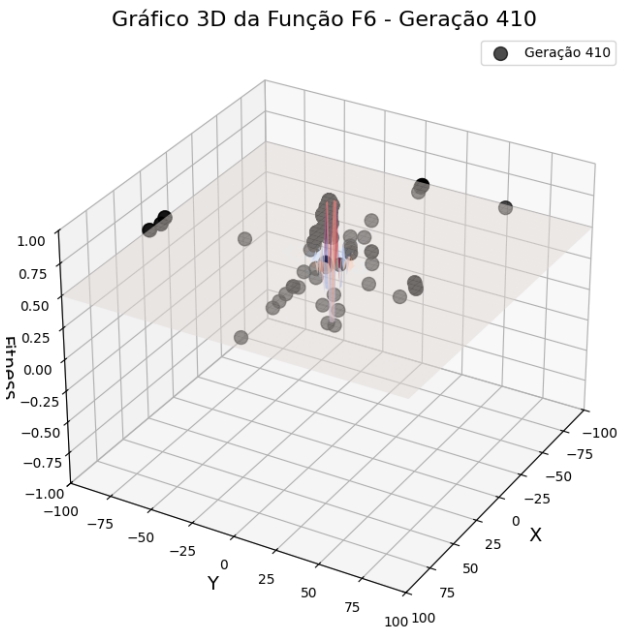

Figure 98: F6 - Binário 410

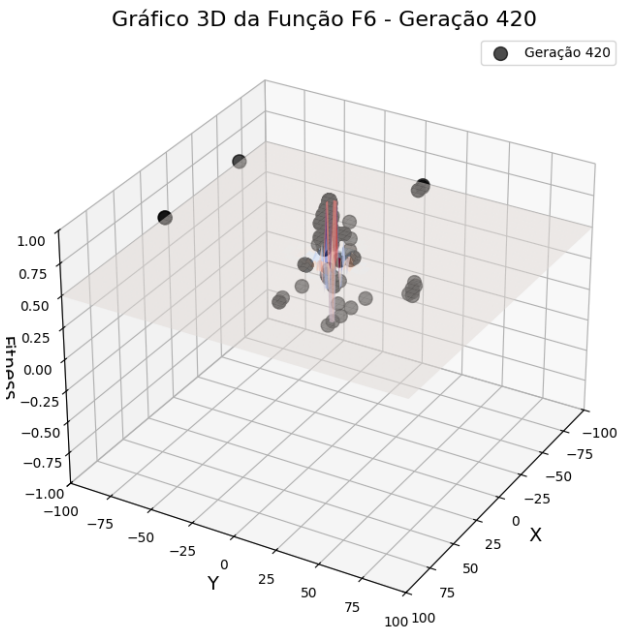

Figure 99: F6 - Binário 420

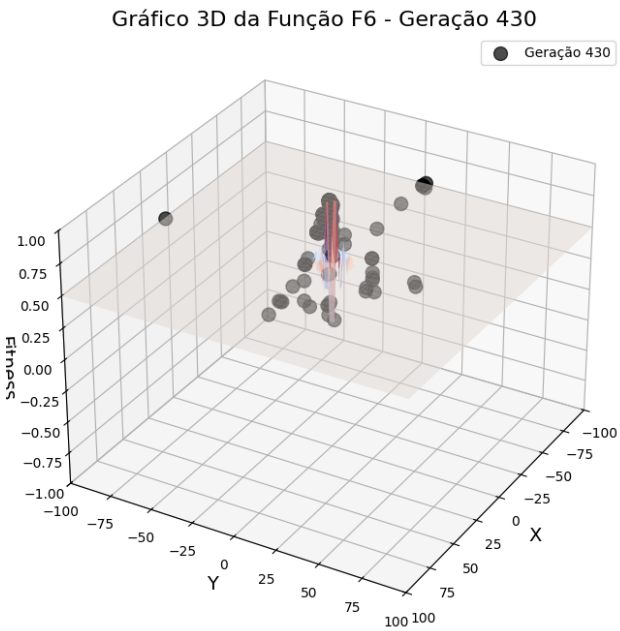

Figure 100: F6 - Binário 430

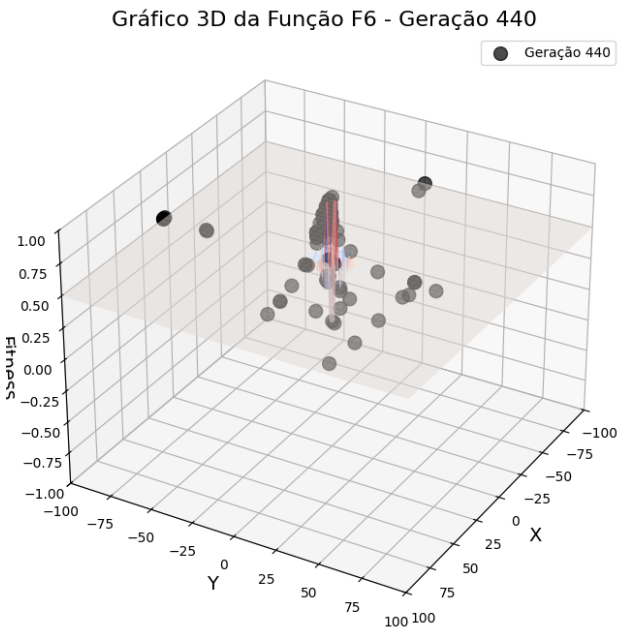

Figure 101: F6 - Binário 440

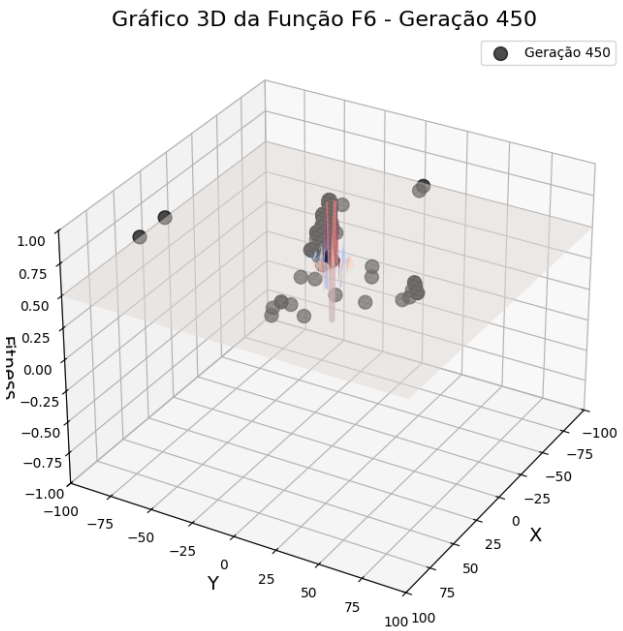

Figure 102: F6 - Binário 450

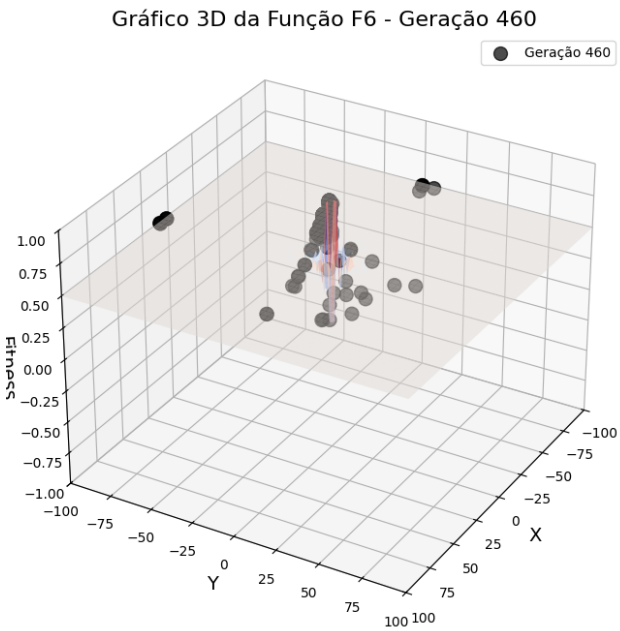

Figure 103: F6 - Binário 460

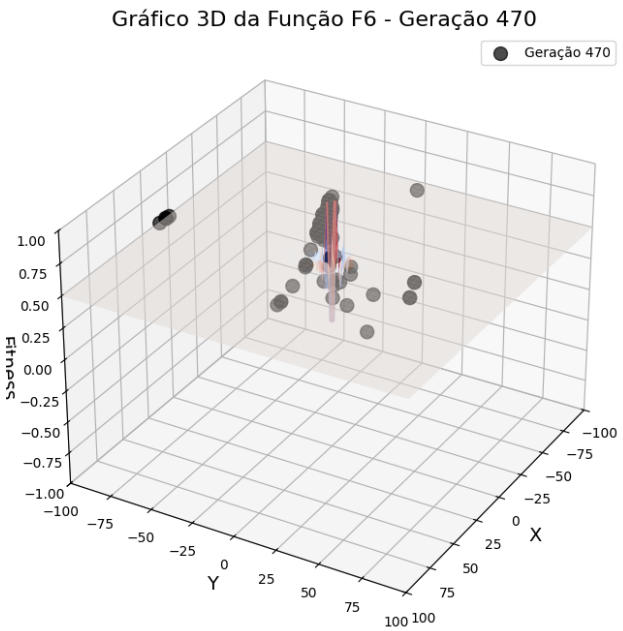

Figure 104: F6 - Binário 470

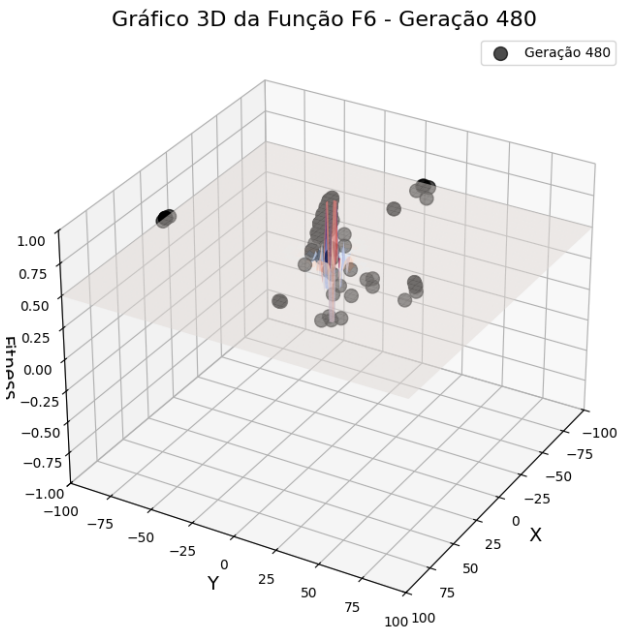

Figure 105: F6 - Binário 480

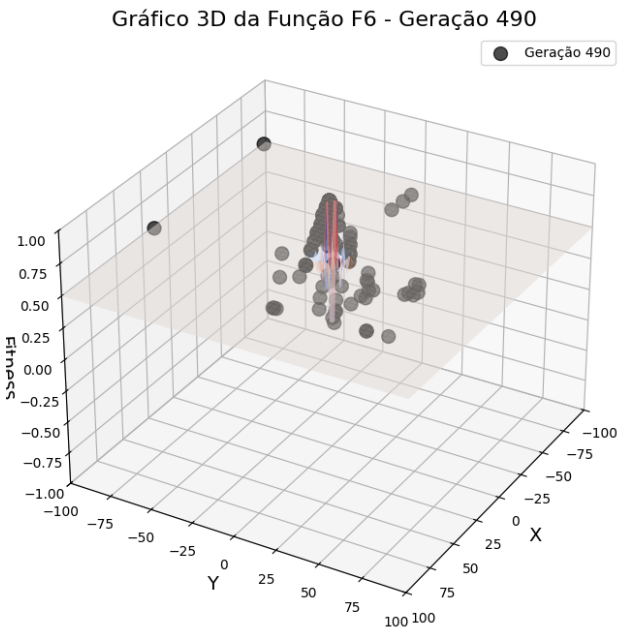

Figure 106: F6 - Binário 490

B.4 Resultados da Representação Real - Curva dos Individuos

Melhoria Inicial

Nos primeiros ciclos, observamos uma melhoria significativa na qualidade da solução. A fitness aumenta rapidamente, passando de 0.942724 na geração 0 para 0.990220 na geração 5. Essa rápida melhoria inicial sugere que o algoritmo está explorando efetivamente o espaço de soluções e encontrando soluções promissoras.

Estabilização e Oscilações

Após a geração 5, a taxa de melhoria começa a estabilizar, com a fitness oscilando em torno de valores próximos a 0.96 a 0.99. Essa estabilização indica que o algoritmo pode estar se aproximando de um ótimo local, onde as soluções não estão mais se diversificando significativamente, mas sim se agrupando em torno de um valor de fitness relativamente alto.

Convergência para um Ótimo Local

A partir da geração 10, a fitness atinge valores muito próximos de 1, com algumas oscilações menores. A presença de valores como 0.999538 na geração 10 e 0.962775 nas gerações subsequentes sugere que o algoritmo pode ter encontrado um ótimo local, onde as soluções não estão mais se movendo significativamente em direção a uma solução melhor.

Estabilidade das Variáveis

As colunas que representam os valores de x e y mostram um comportamento mais volátil nas primeiras gerações, mas tendem a se estabilizar em torno de certos valores à medida que o algoritmo avança. Isso é um indicativo de que, enquanto a qualidade da solução (fitness) se estabiliza, as variáveis que compõem essa solução também estão se ajustando para valores que maximizam a função objetivo.

Exploração e Exploração

A oscilação nos valores de x e y sugere que o algoritmo está equilibrando a exploração de novas soluções e a exploração de soluções já conhecidas. Essa dinâmica é crucial em algoritmos genéticos, pois permite que o algoritmo não fique preso em ótimos locais, mas também não perca a qualidade das soluções.

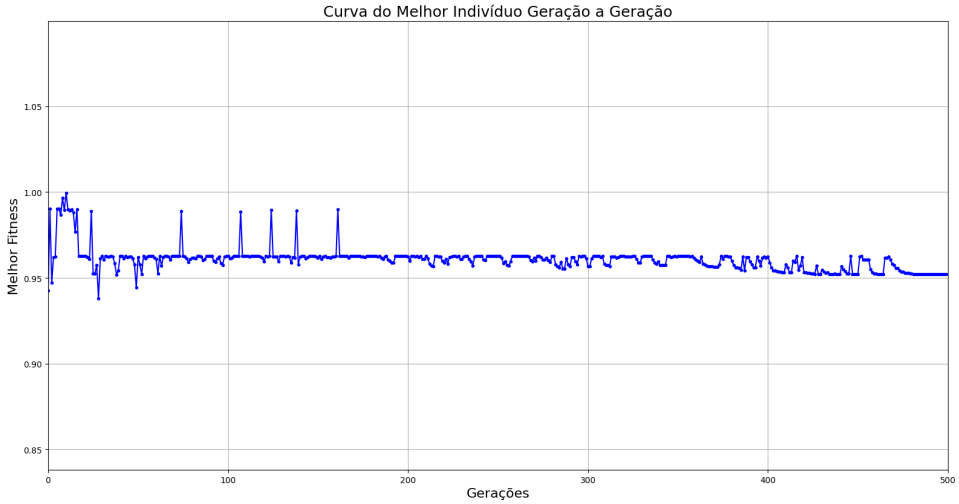

Figure 101: Individuos - Real 0

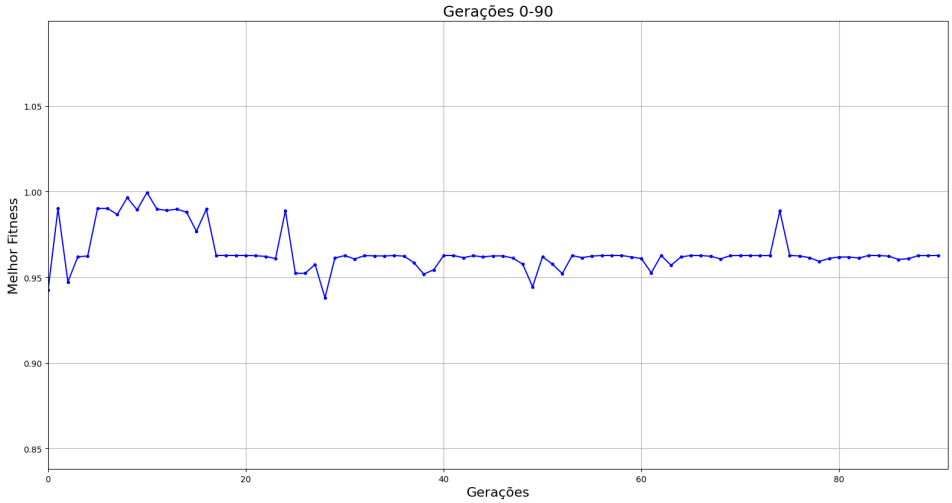

Figure 102: Individuos - Real 1

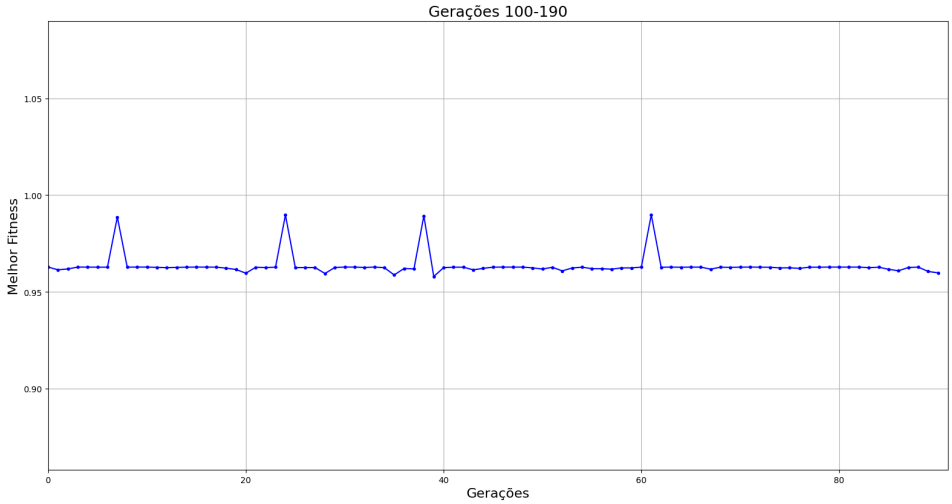

Figure 103: Individuos - Real 2

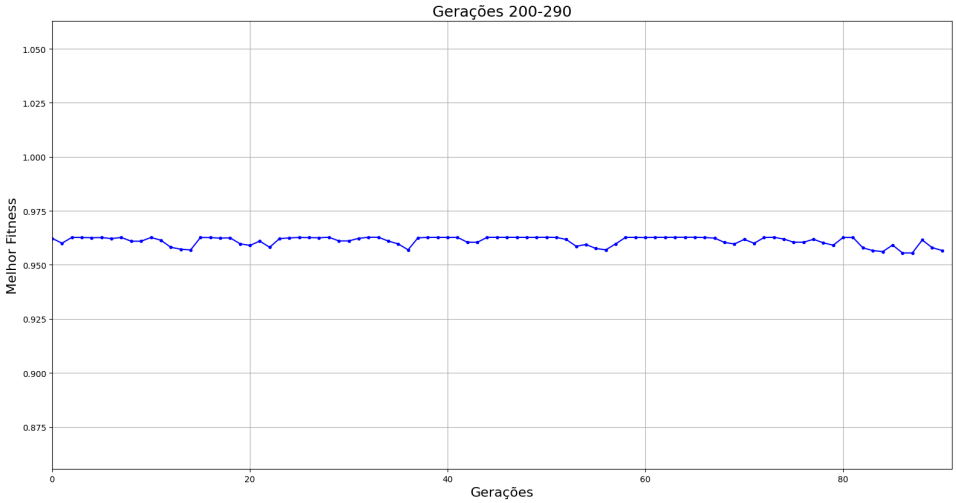

Figure 104: Individuos - Real 3

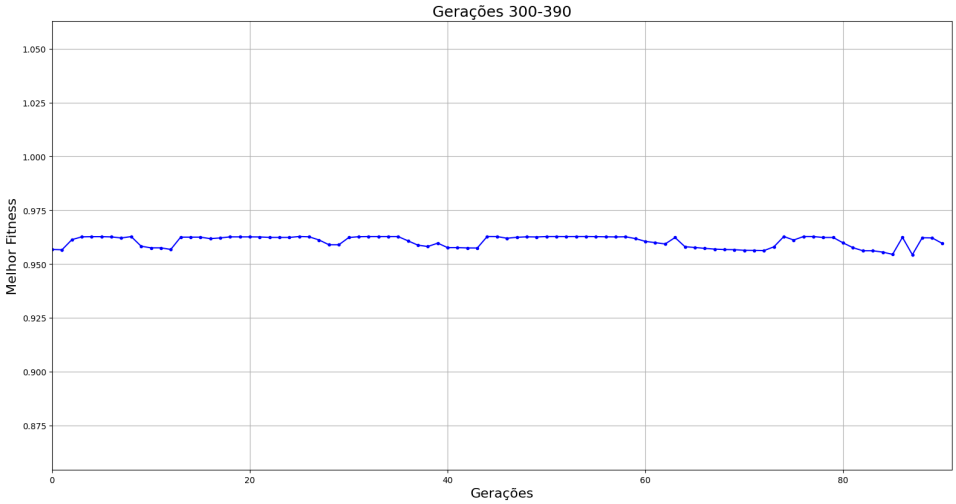

Figure 105: Individuos - Real 4

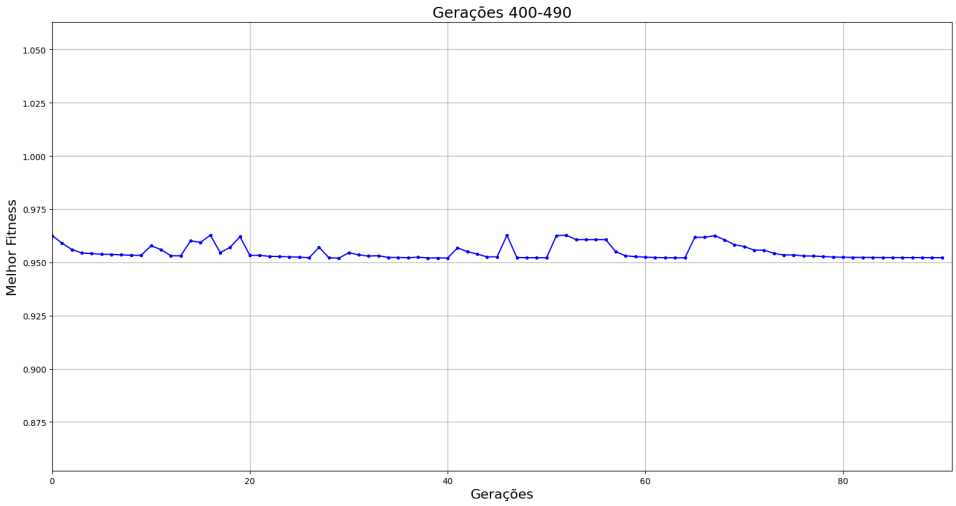

Figure 106: Individuos - Real 5

B.5 Gráficos 3D da Função F6 com População (Representação Real)

Nesta subseção, apresentamos gráficos tridimensionais semelhantes, mas focados na população gerada com a representação real. Esses gráficos ilustram a interação dos indivíduos com a função F6, permitindo uma comparação direta com os resultados obtidos pela representação binária. A visualização dos dados em 3D ajuda a identificar padrões de convergência e a eficácia da representação real na otimização do problema em questão.

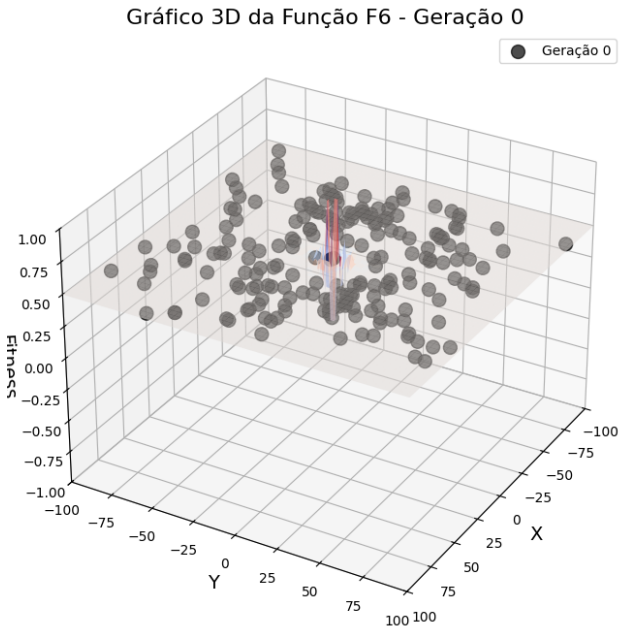

Figure 107: F6 - Real 0

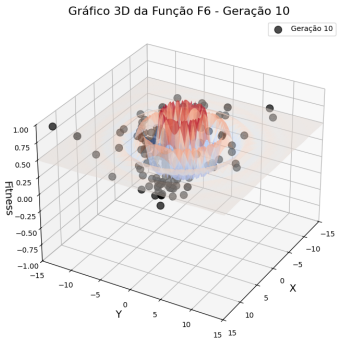

Figure 108: F6 - Real 10

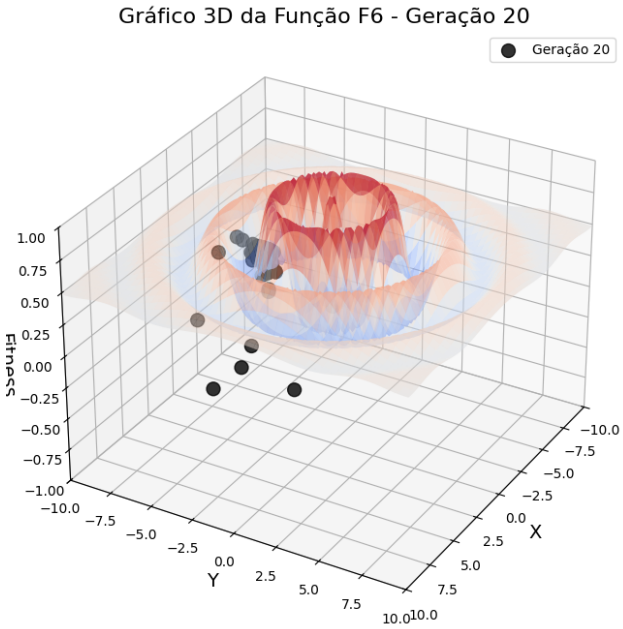

Figure 109: F6 - Real 20

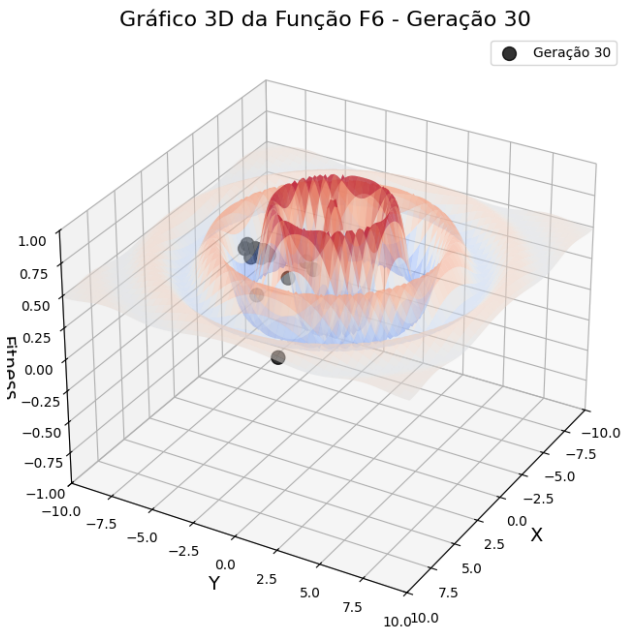

Figure 110: F6 - Real 30

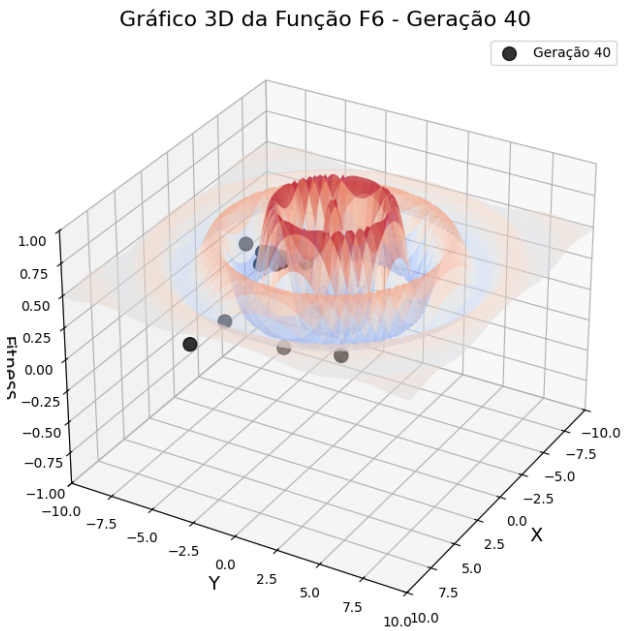

Figure 111: F6 - Real 40

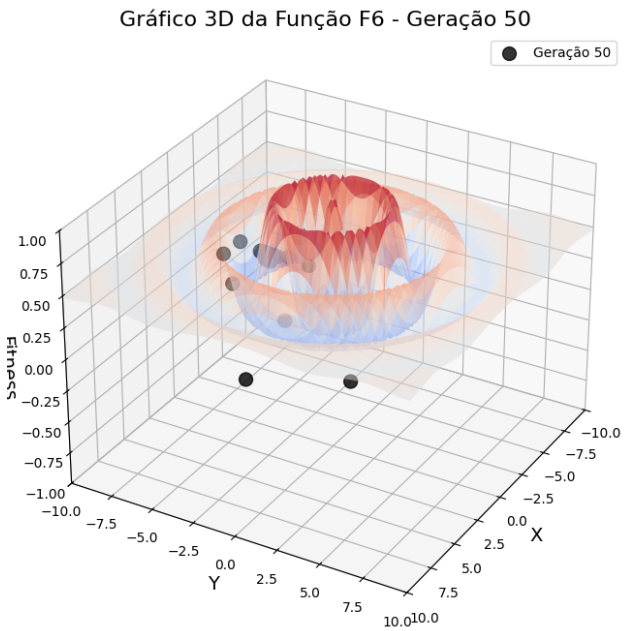

Figure 112: F6 - Real 50

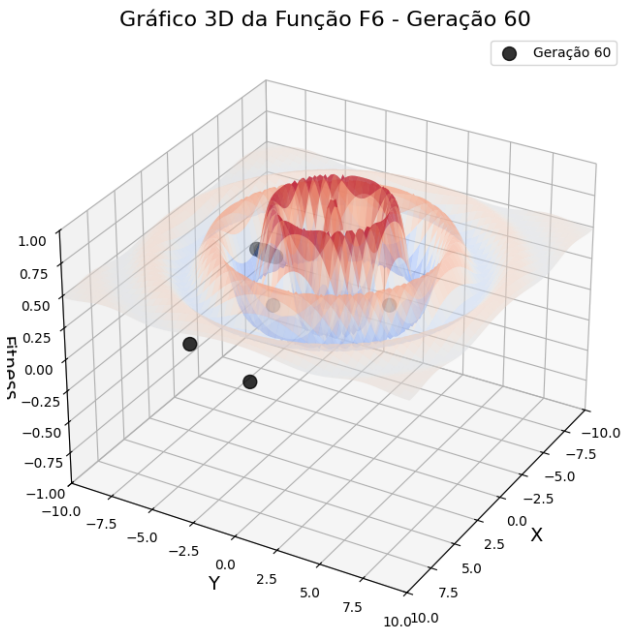

Figure 113: F6 - Real 60

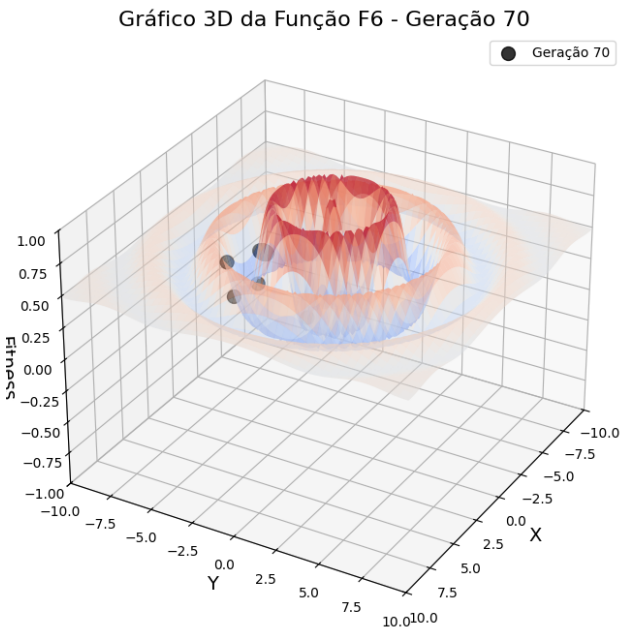

Figure 114: F6 - Real 70

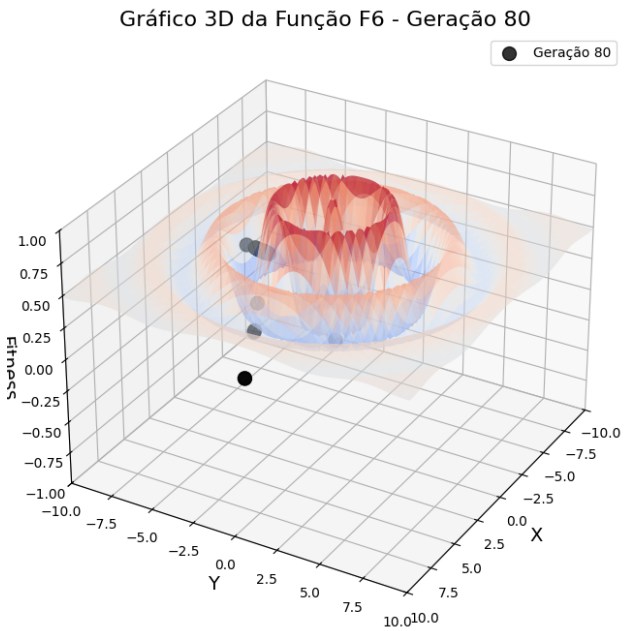

Figure 115: F6 - Real 80

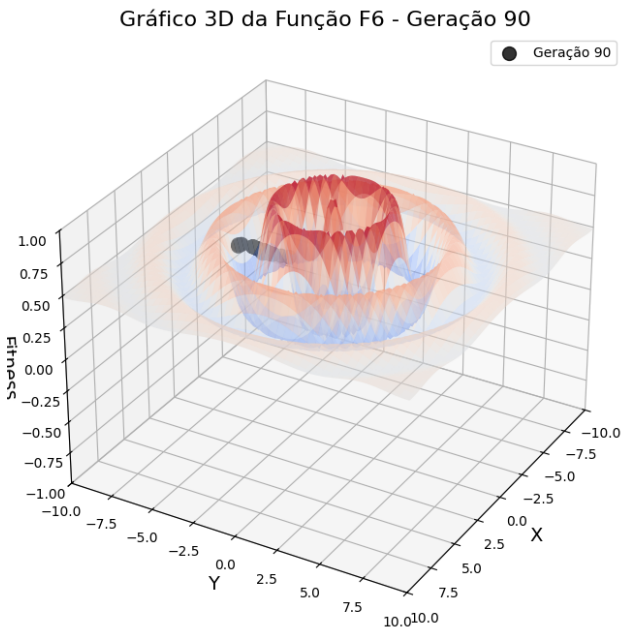

Figure 116: F6 - Real 90

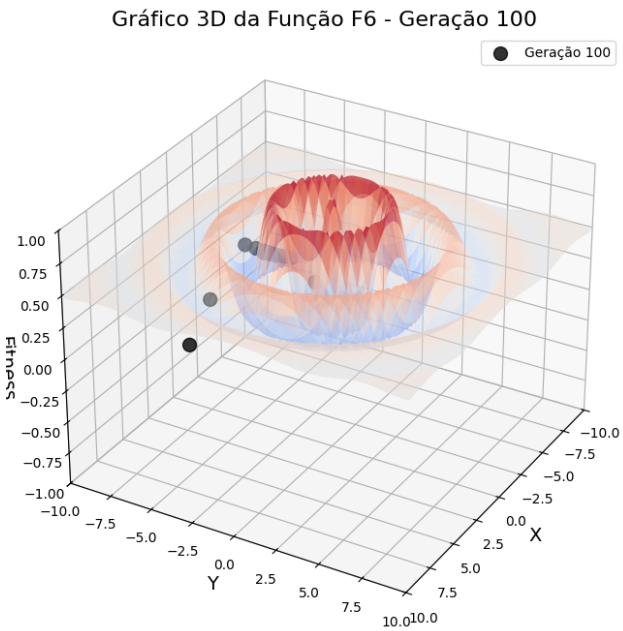

Figure 117: F6 - Real 100

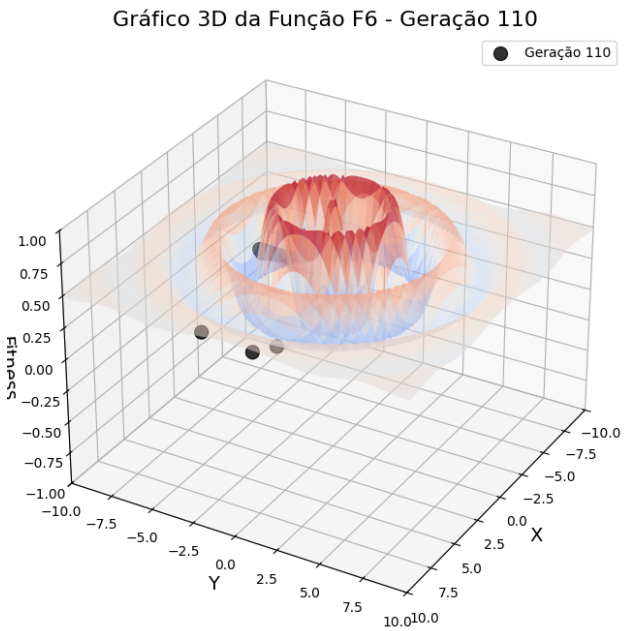

Figure 118: F6 - Real 110

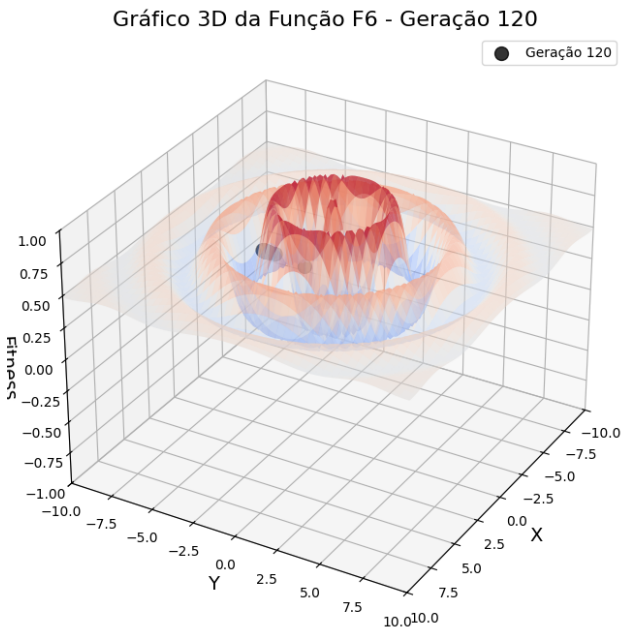

Figure 119: F6 - Real 120

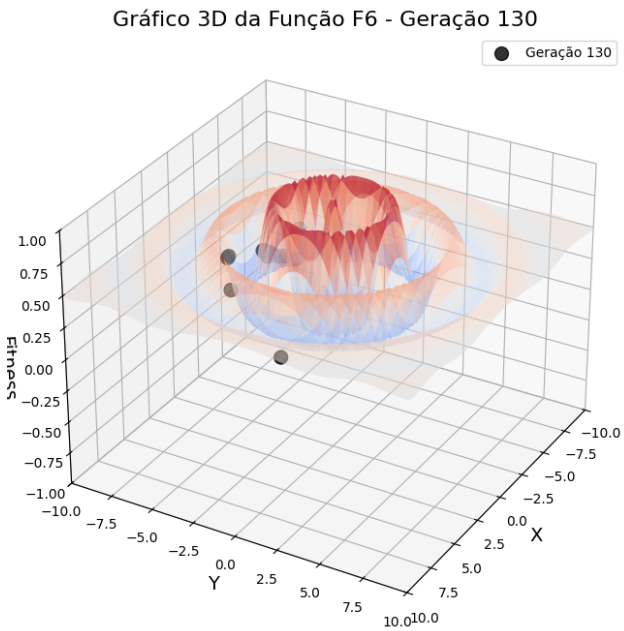

Figure 120: F6 - Real 130

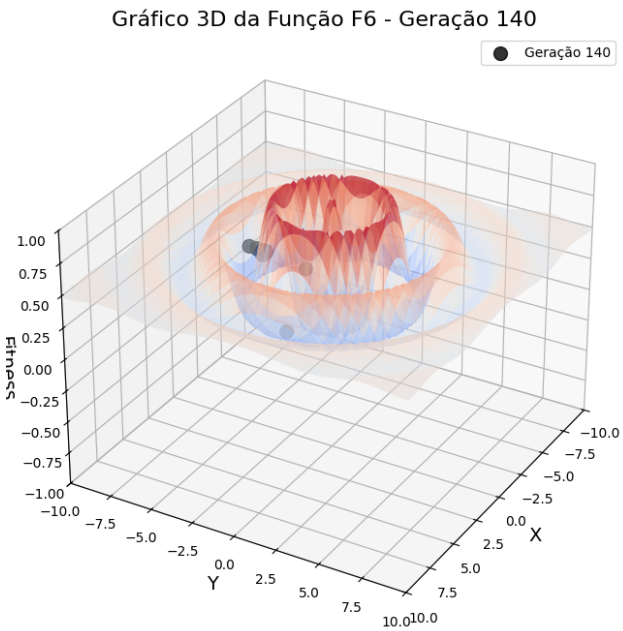

Figure 121: F6 - Real 140

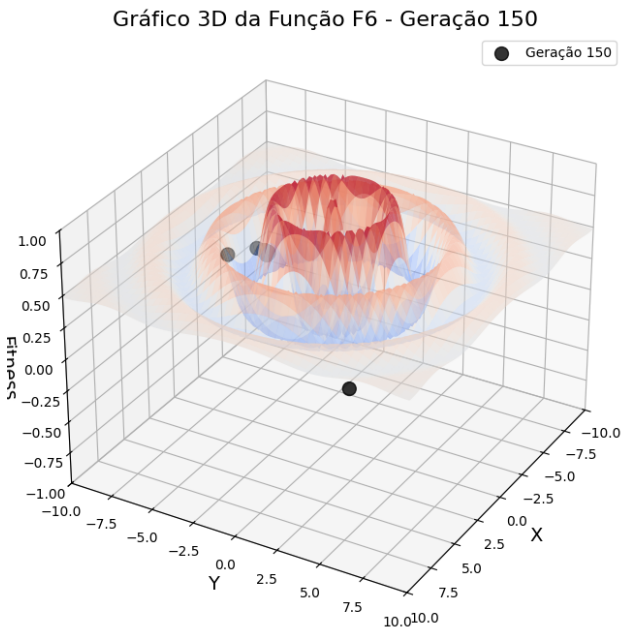

Figure 122: F6 - Real 150

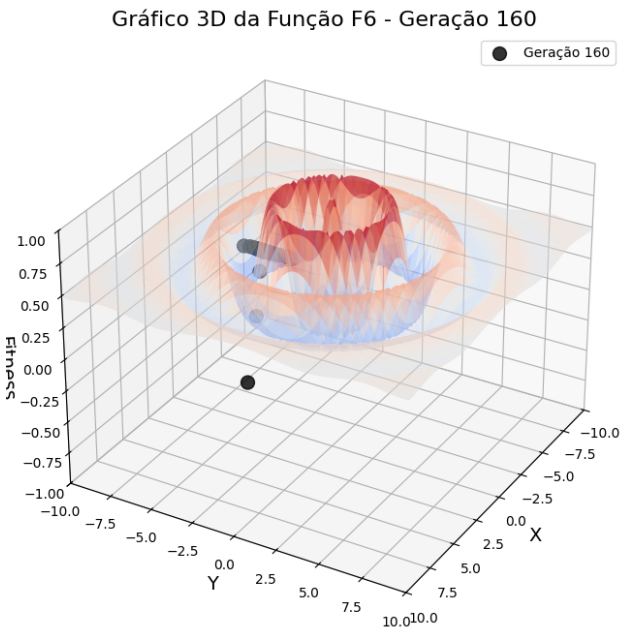

Figure 123: F6 - Real 160

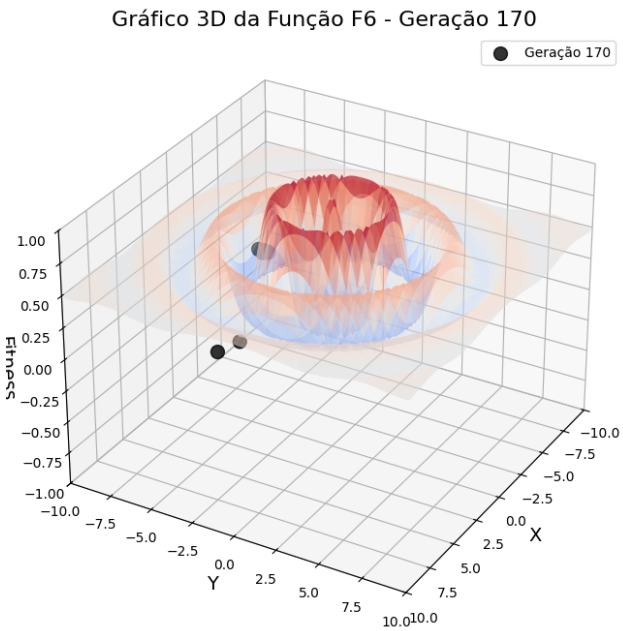

Figure 124: F6 - Real 170

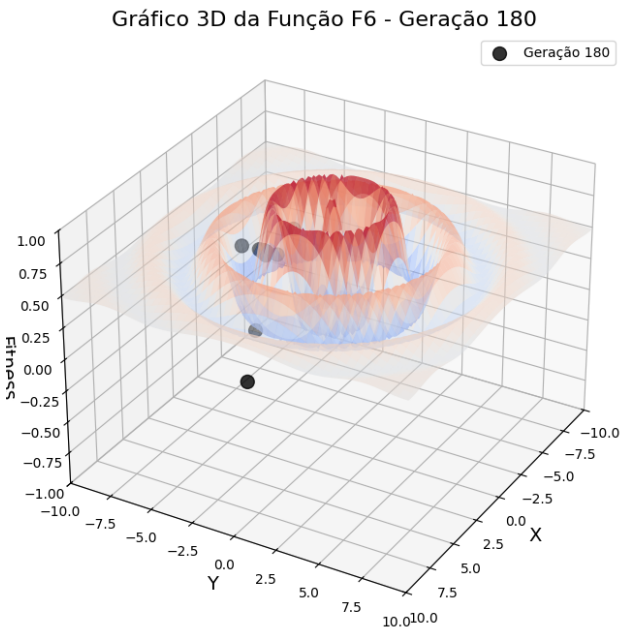

Figure 125: F6 - Real 180

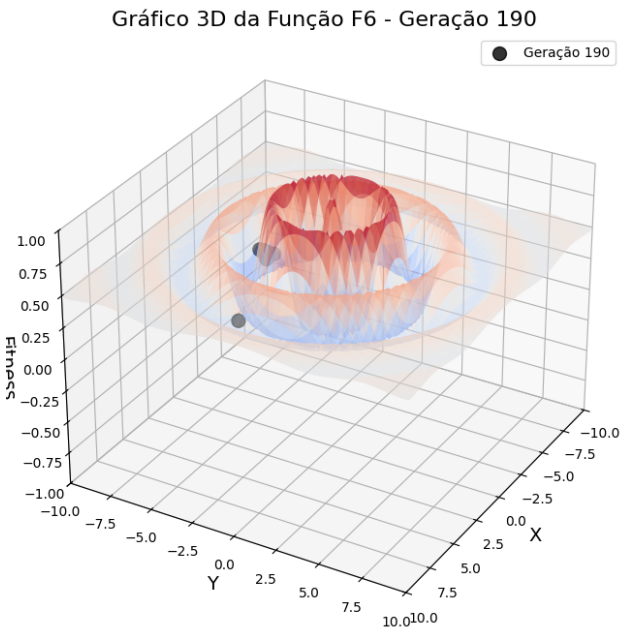

Figure 126: F6 - Real 190

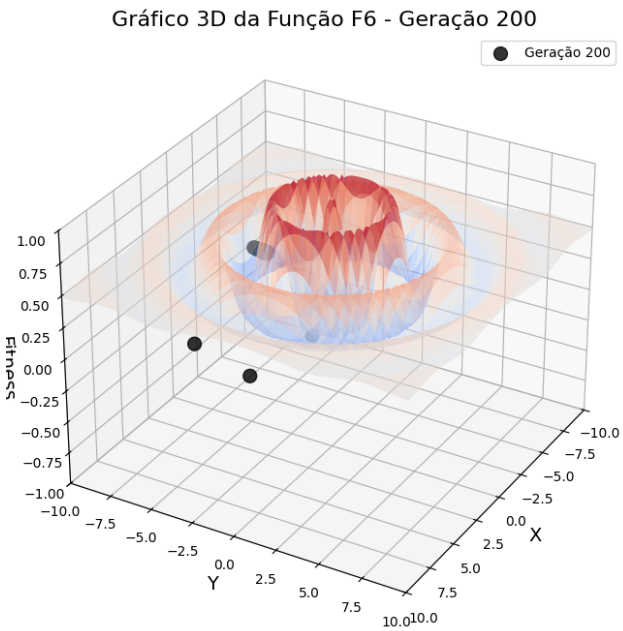

Figure 127: F6 - Real 200

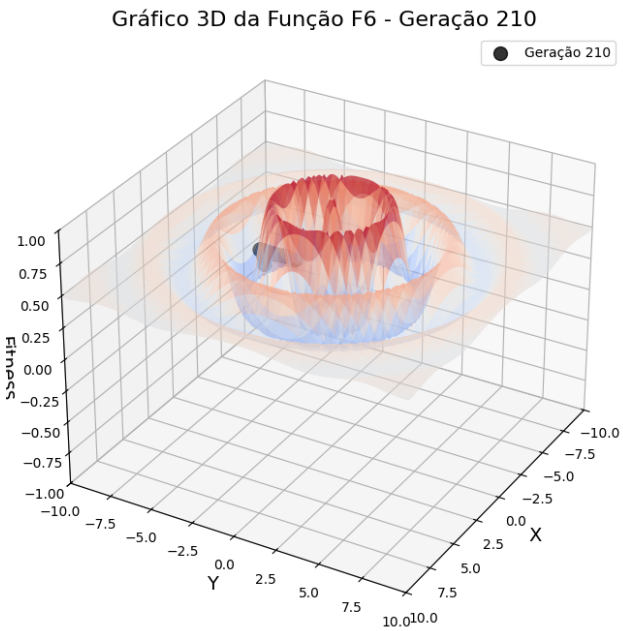

Figure 128: F6 - Real 210

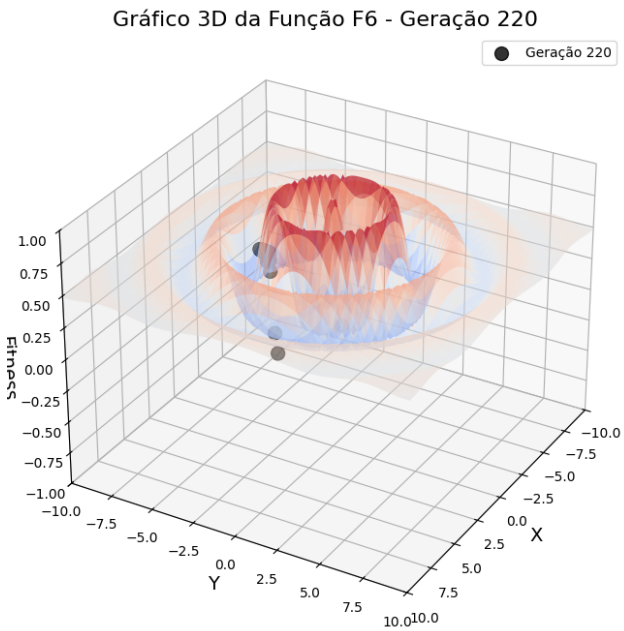

Figure 129: F6 - Real 220

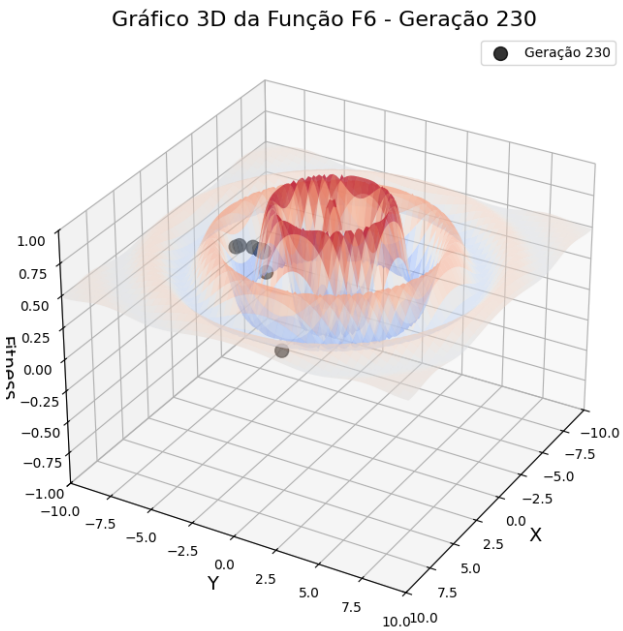

Figure 130: F6 - Real 230

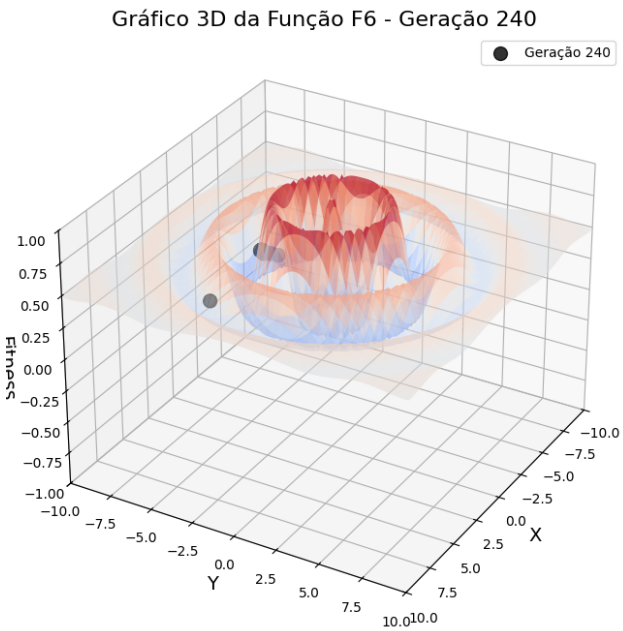

Figure 131: F6 - Real 240

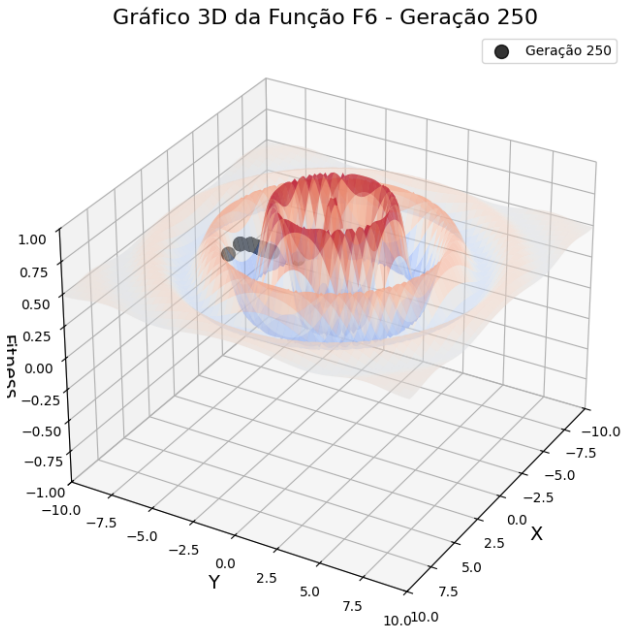

Figure 132: F6 - Real 250

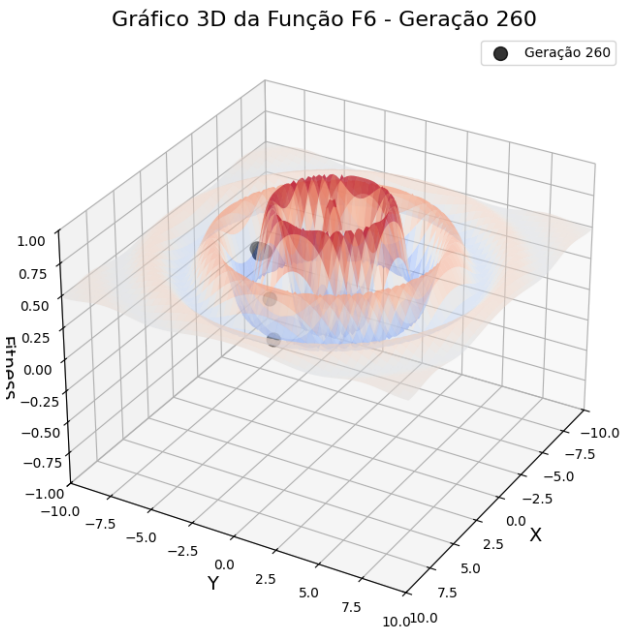

Figure 133: F6 - Real 260

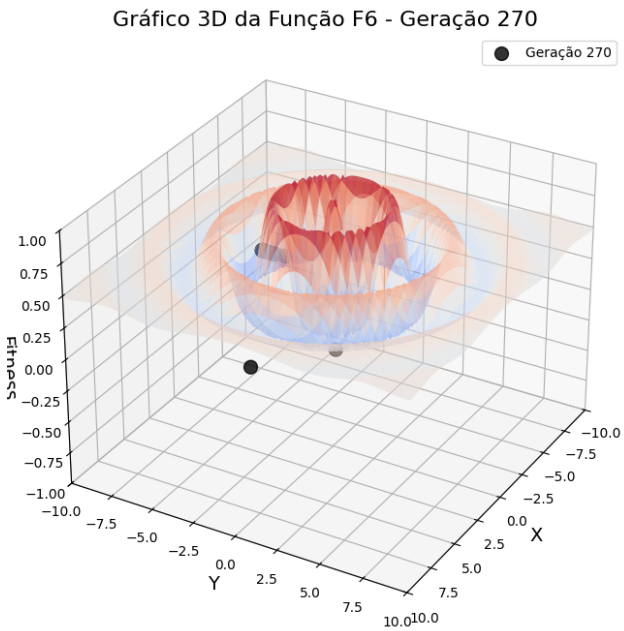

Figure 134: F6 - Real 270

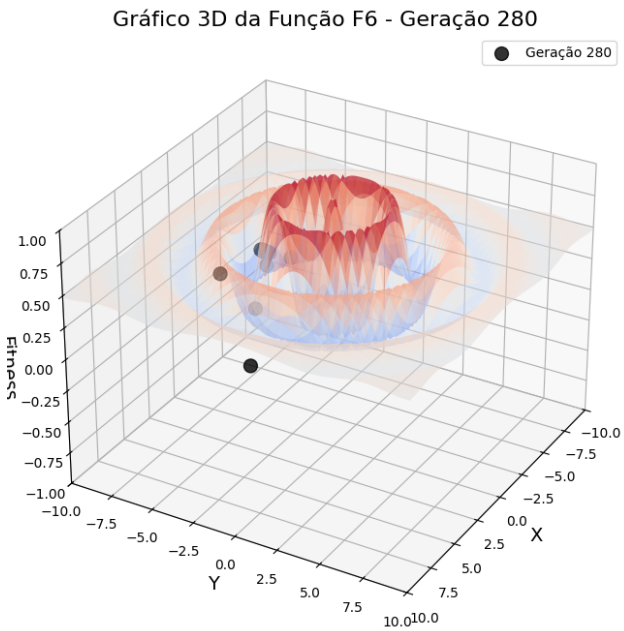

Figure 135: F6 - Real 280

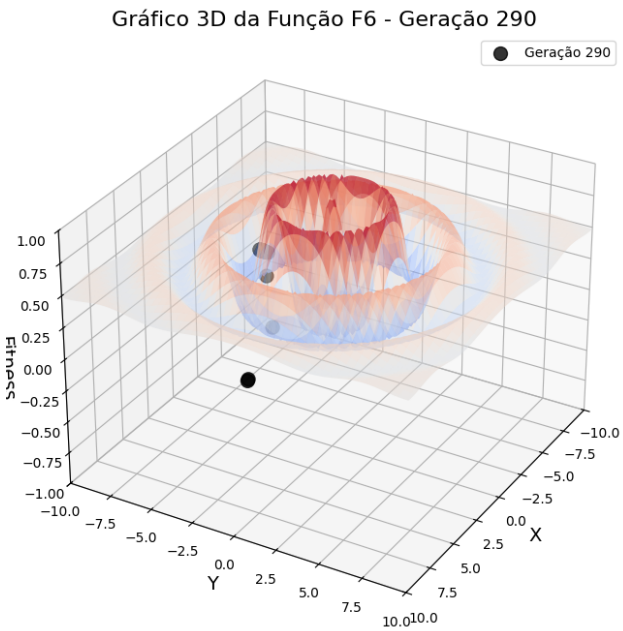

Figure 136: F6 - Real 290

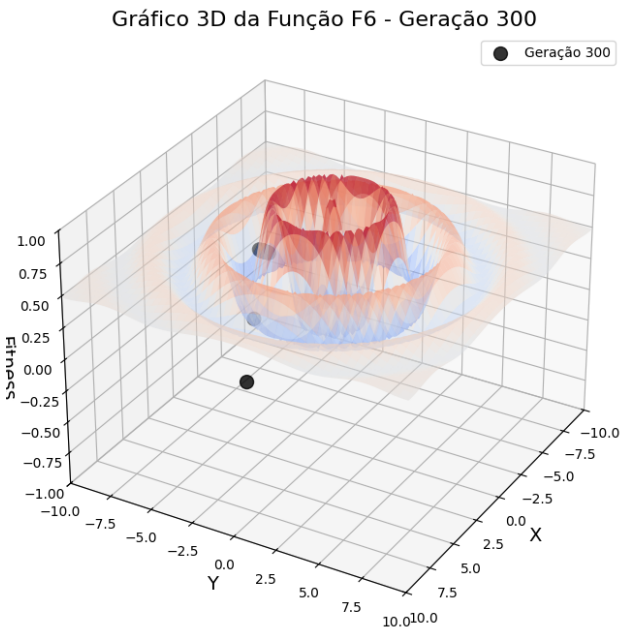

Figure 137: F6 - Real 300

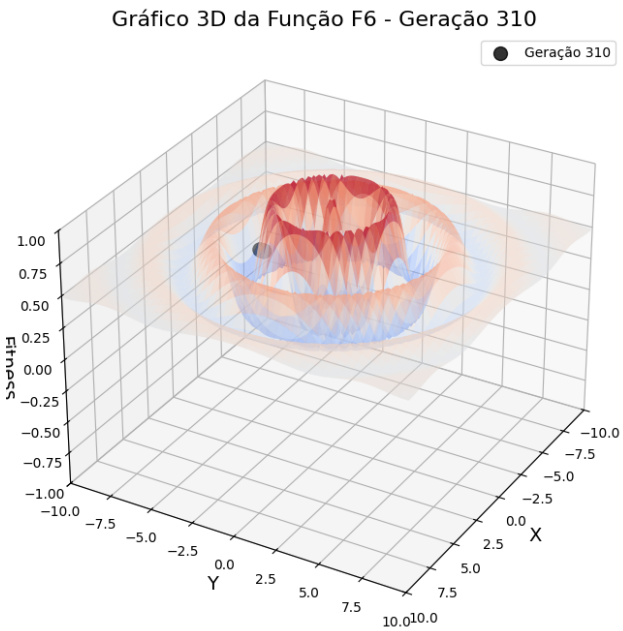

Figure 138: F6 - Real 310

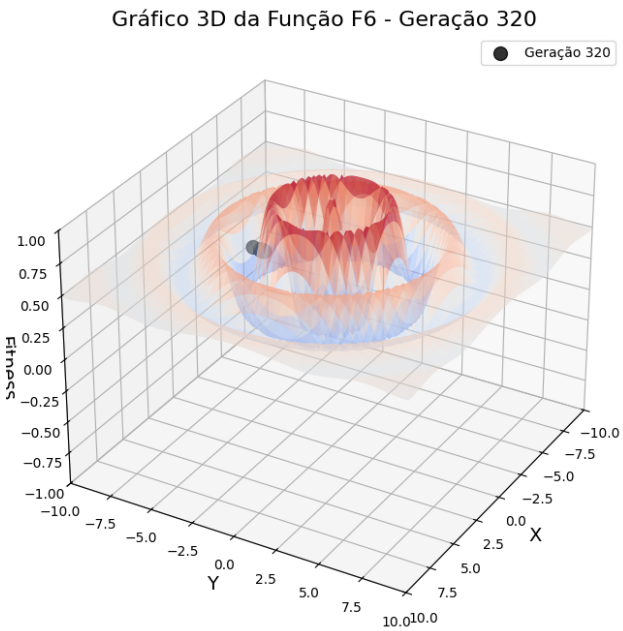

Figure 139: F6 - Real 320

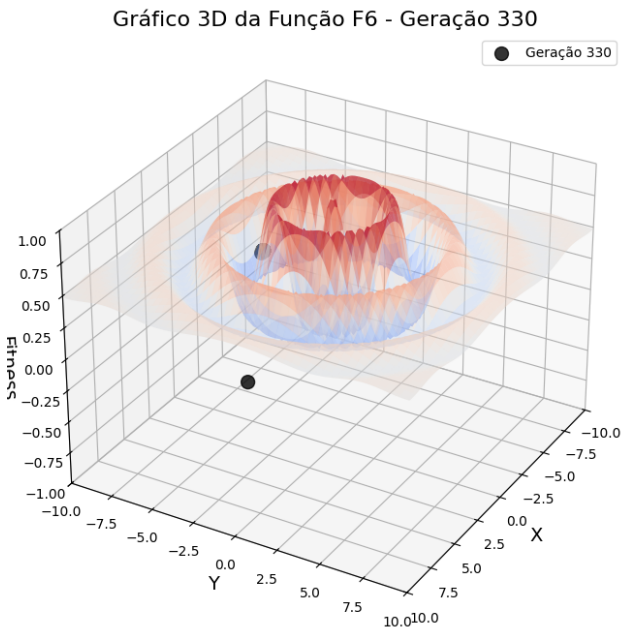

Figure 140: F6 - Real 330

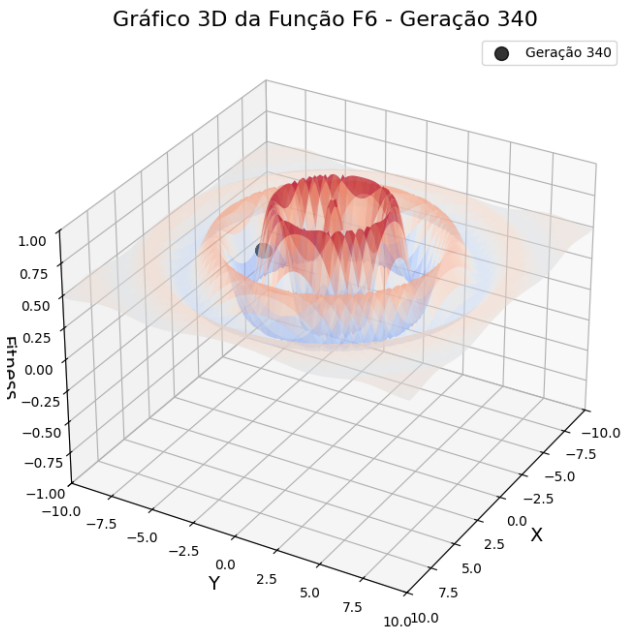

Figure 141: F6 - Real 340

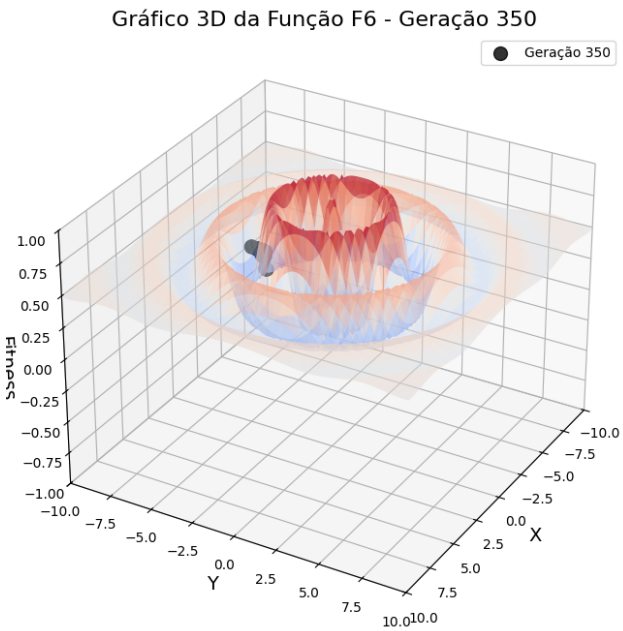

Figure 142: F6 - Real 350

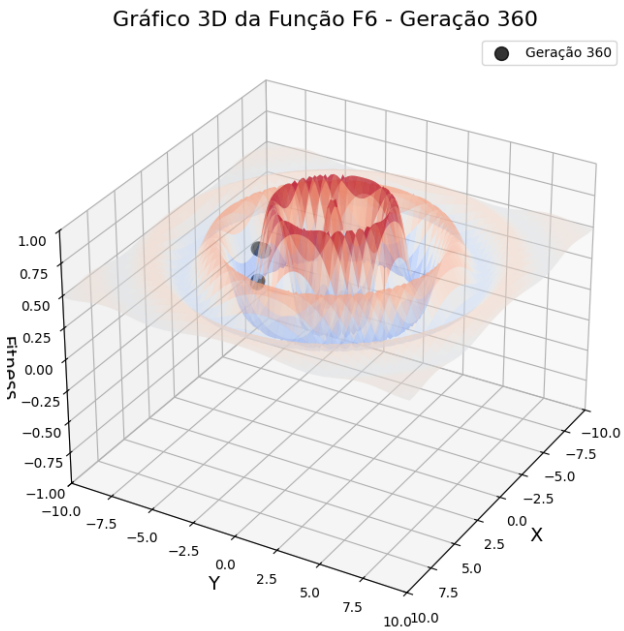

Figure 143: F6 - Real 360

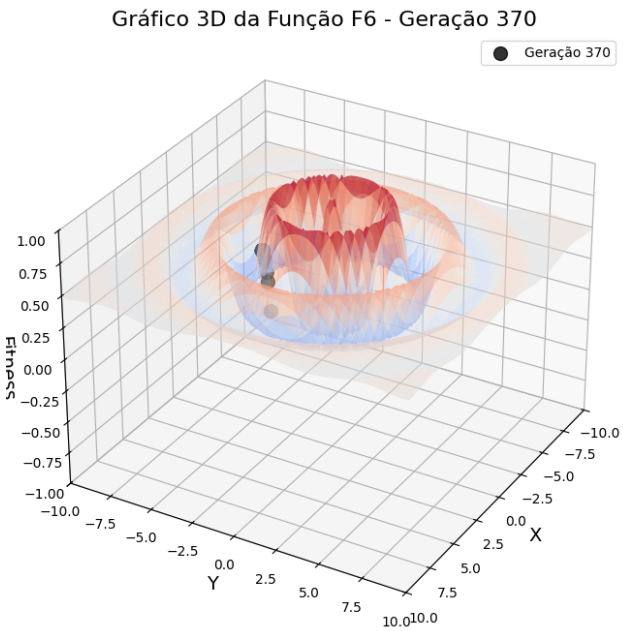

Figure 144: F6 - Real 370

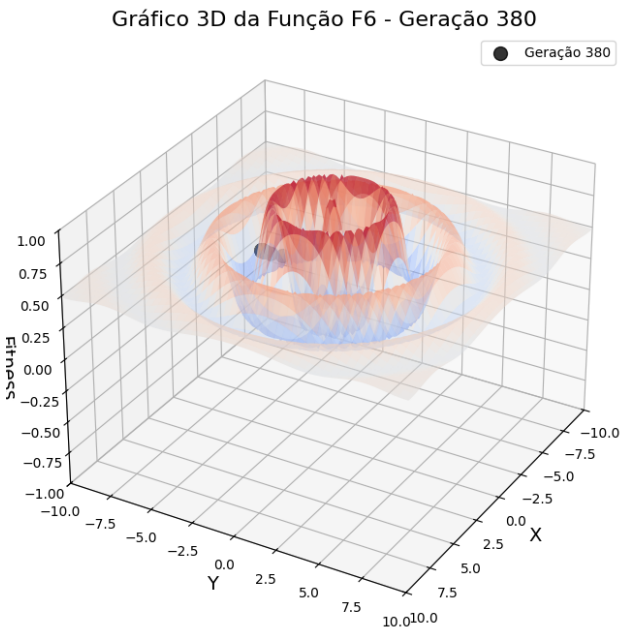

Figure 145: F6 - Real 380

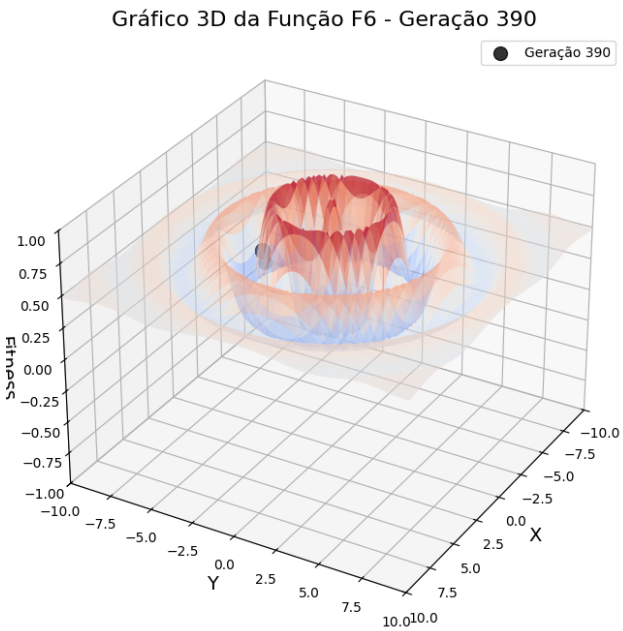

Figure 146: F6 - Real 390

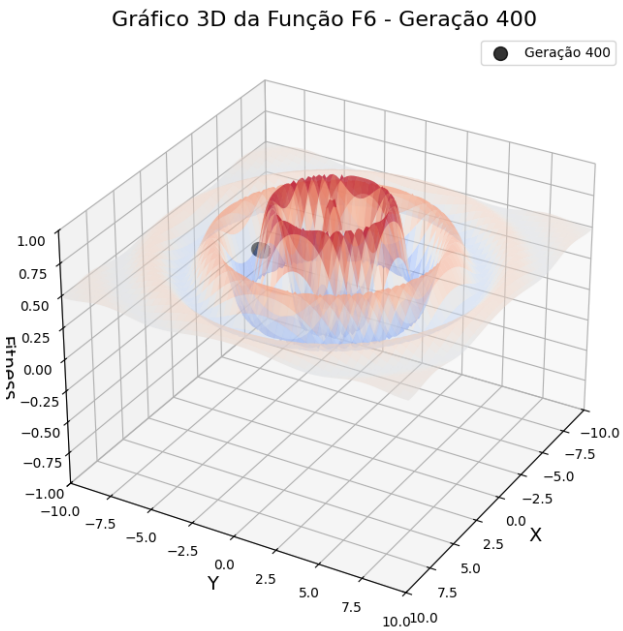

Figure 147: F6 - Real 400

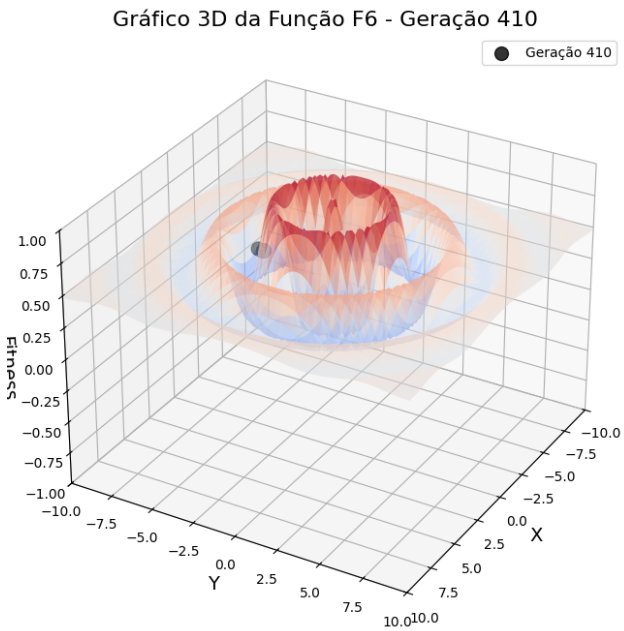

Figure 148: F6 - Real 410

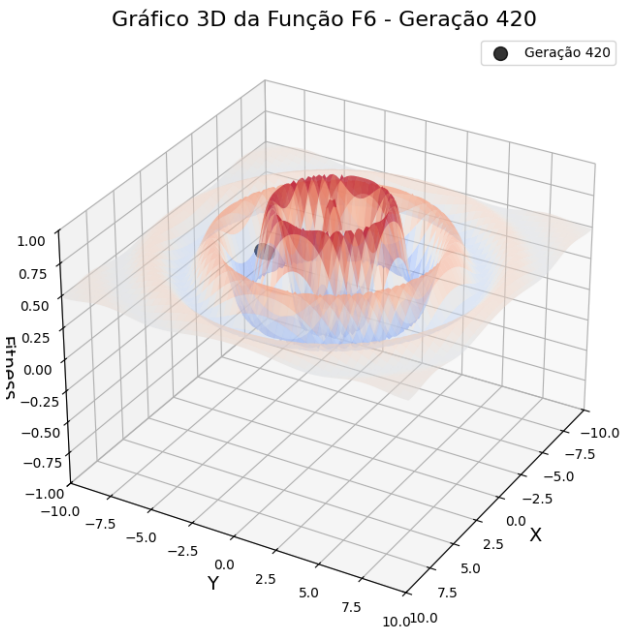

Figure 149: F6 - Real 420

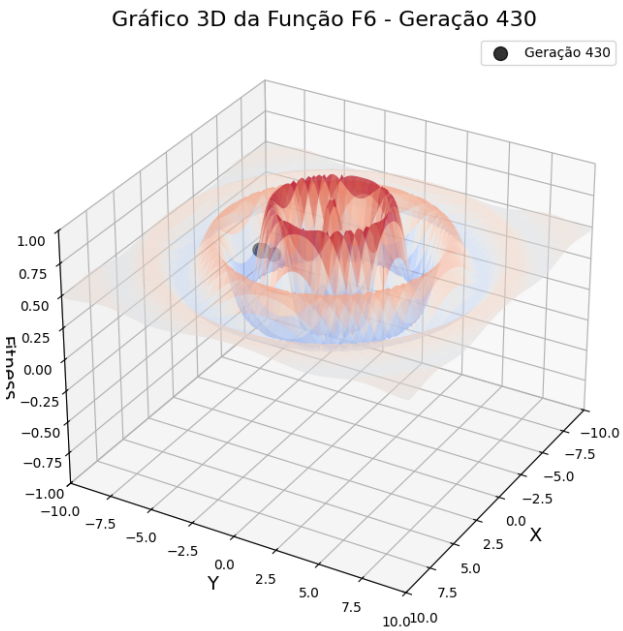

Figure 150: F6 - Real 430

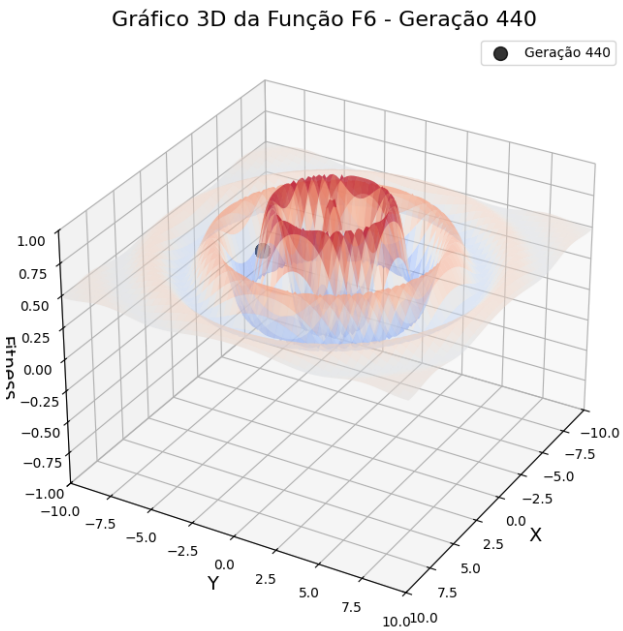

Figure 151: F6 - Real 440

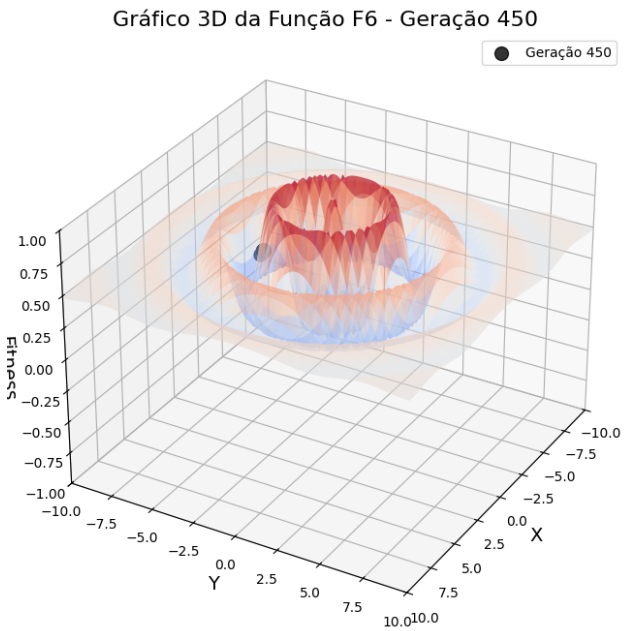

Figure 152: F6 - Real 450

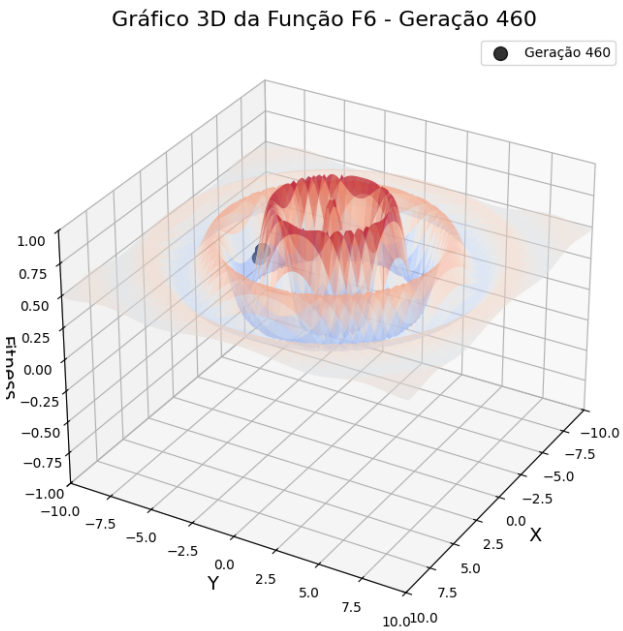

Figure 153: F6 - Real 460

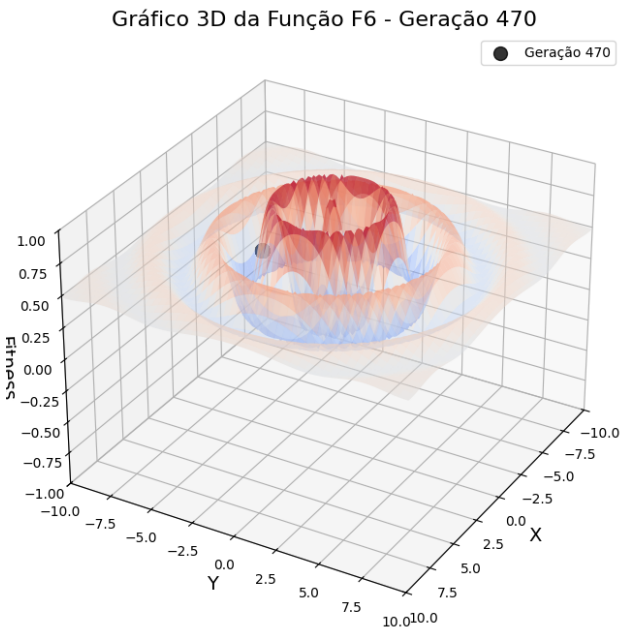

Figure 154: F6 - Real 470

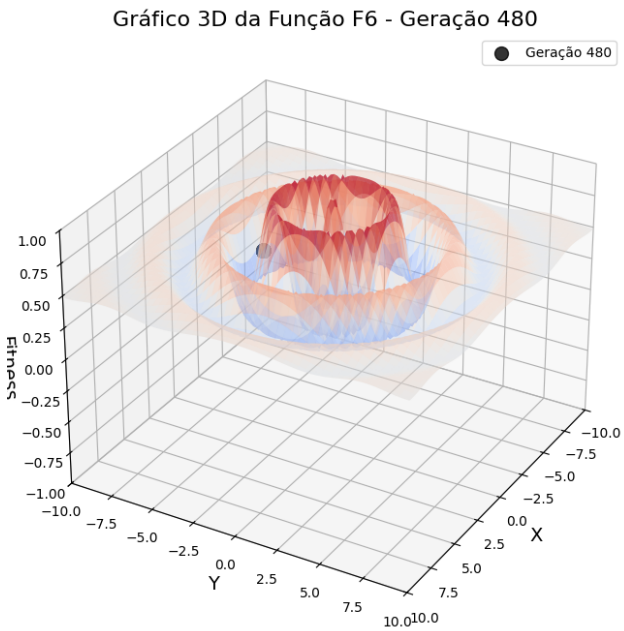

Figure 155: F6 - Real 480

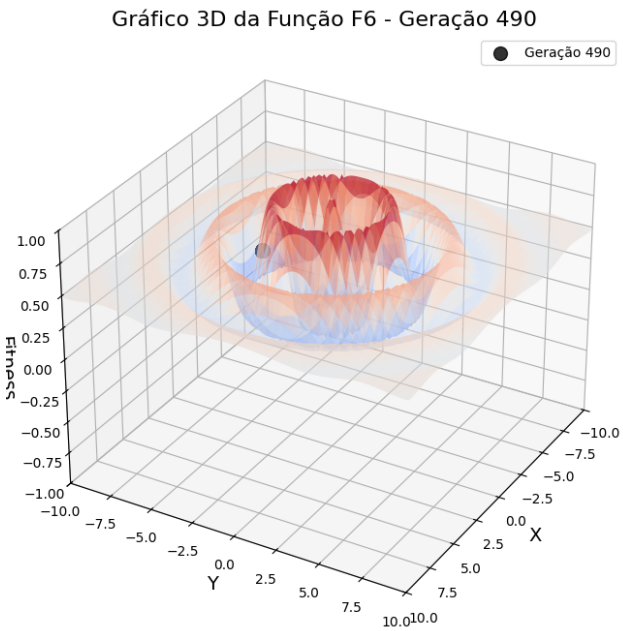

Figure 156: F6 - Real 490