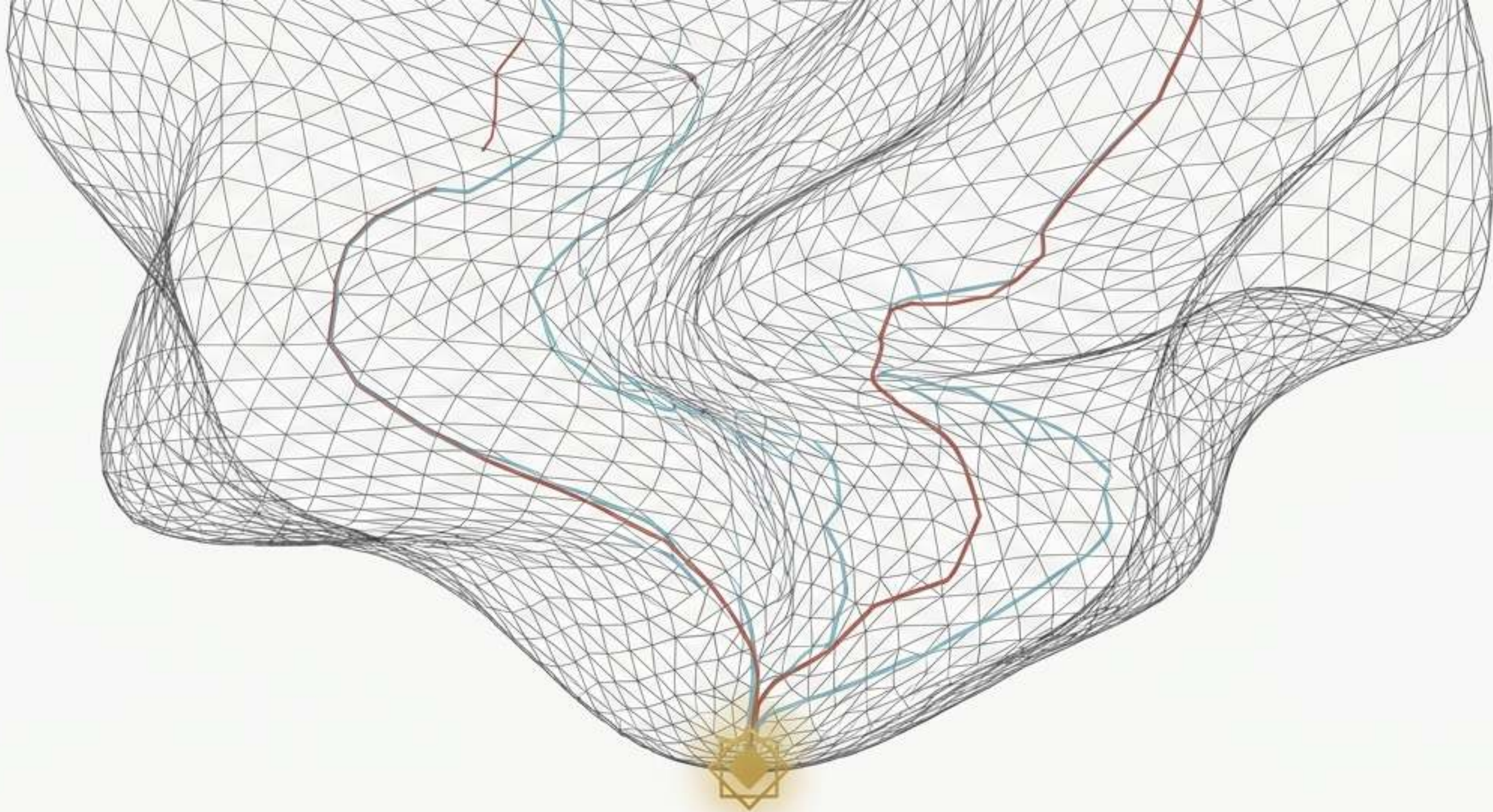




ZO-AdaMM e a Fronteira Populacional

Otimização de Caixa-Preta via Oráculos Adaptativos e
a Evolução para o Consenso de Enxame

O Desafio da Otimização de Caixa-Preta

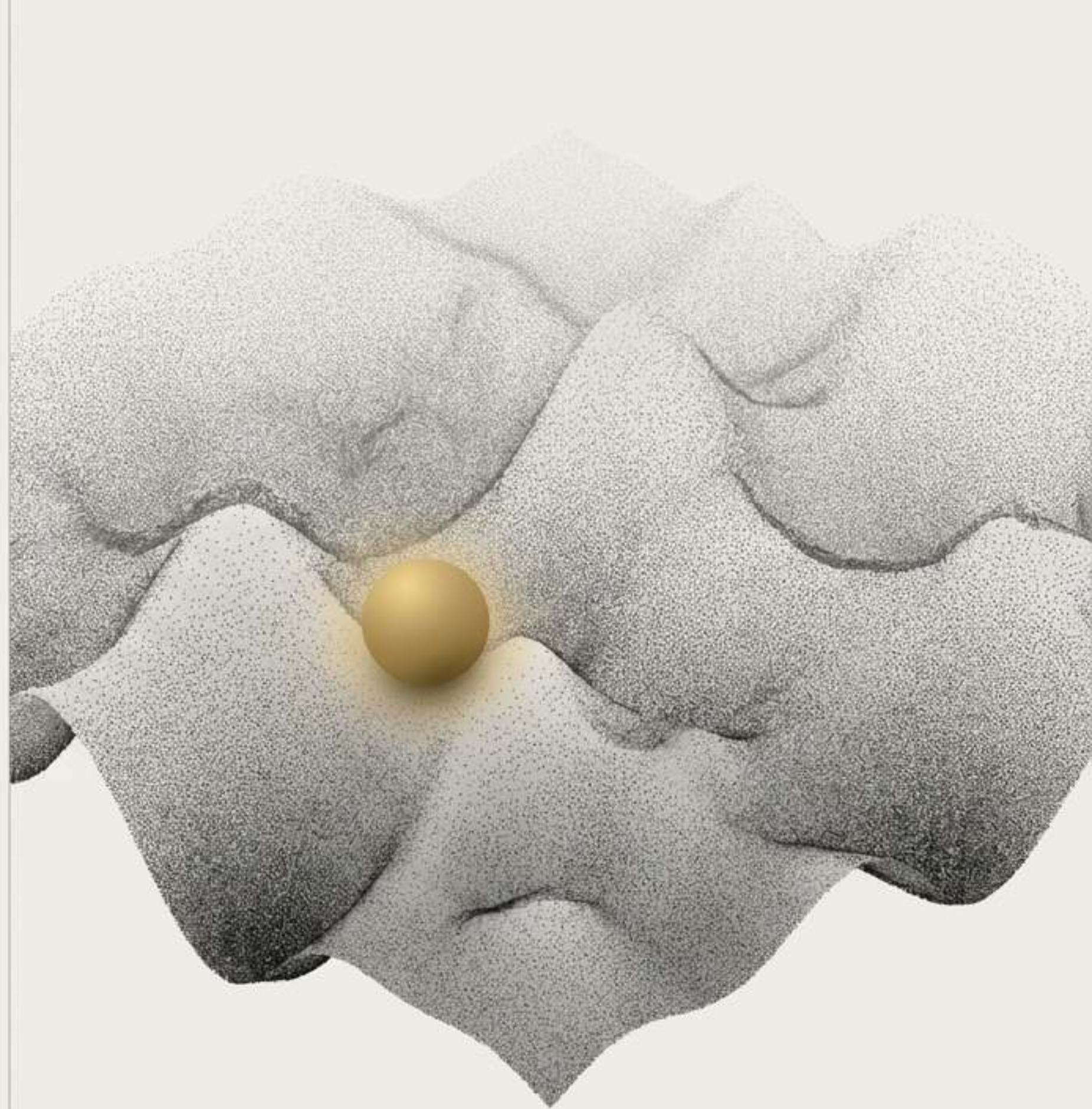
$$\min_{x \in X} f(x) = \mathbb{E}_{\xi}[f(x; \xi)].$$

Gargalo Fundamental

Expressões explícitas de gradiente são custosas ou incomputáveis (ex: ataques a redes neurais).

A Consequência

A amostragem estocástica aliada à estimativa aleatória gera variância extrema, inviabilizando a convergência rápida.

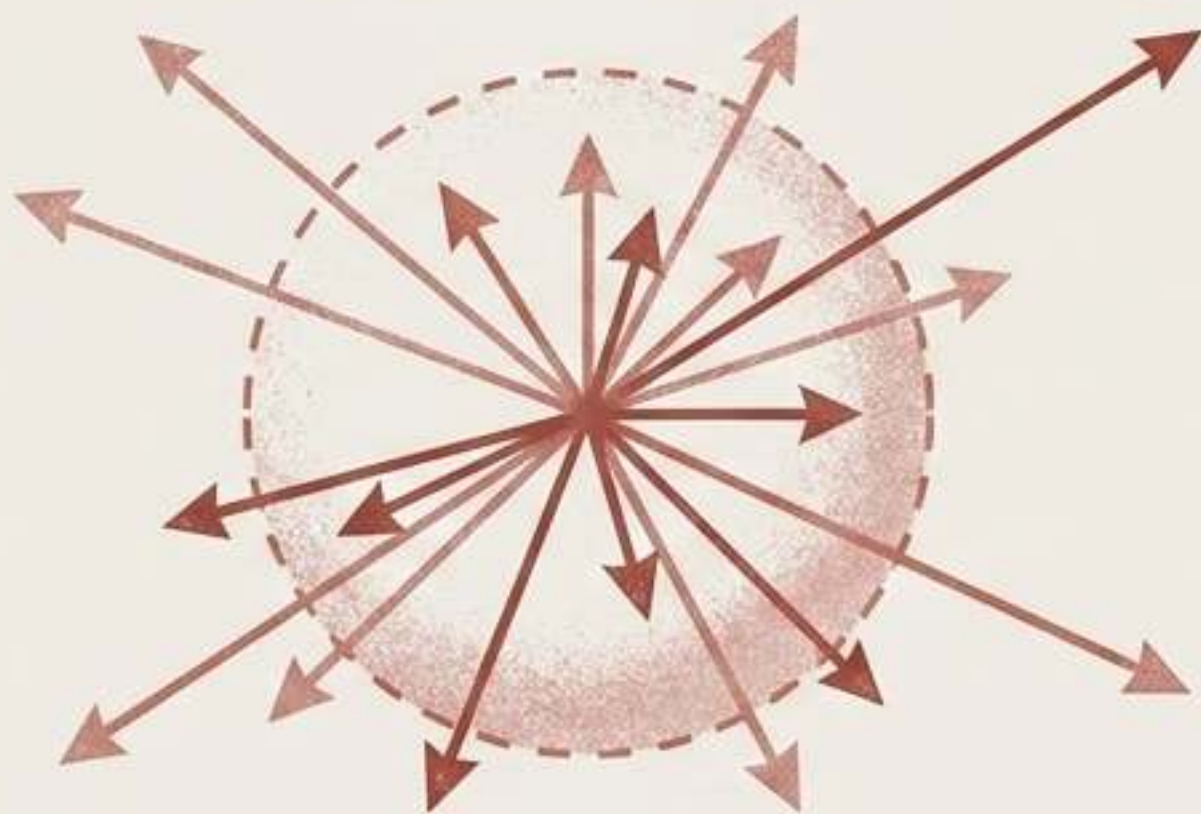


A Bússola Cega: Estimadores de Ordem Zero

$$\hat{\nabla} f(x) = \frac{d}{\mu} [f(x + \mu \mathbf{u}) - f(x)] \mathbf{u}$$

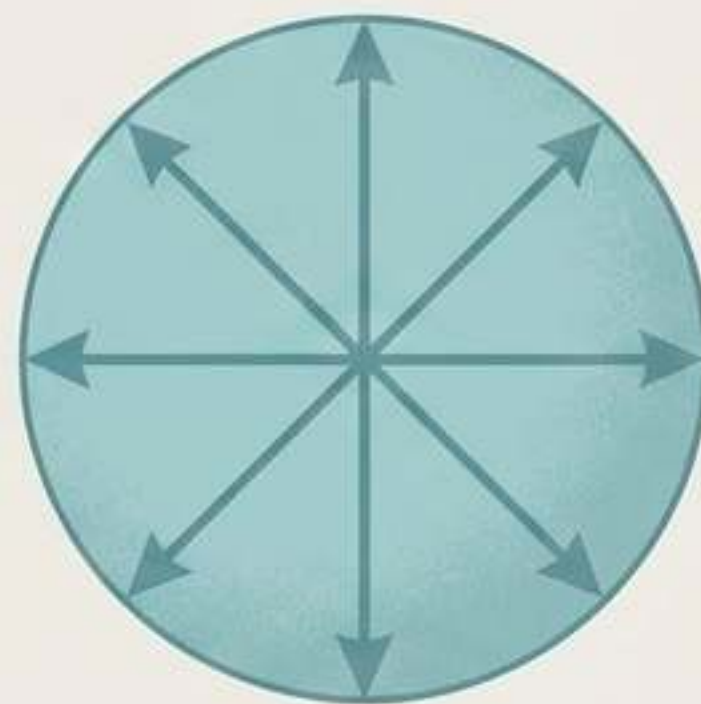
Diferença finita em uma direção aleatória unitária $\mathbf{u}$, suavizada por $\mu > 0$.

Oráculo Gaussiano



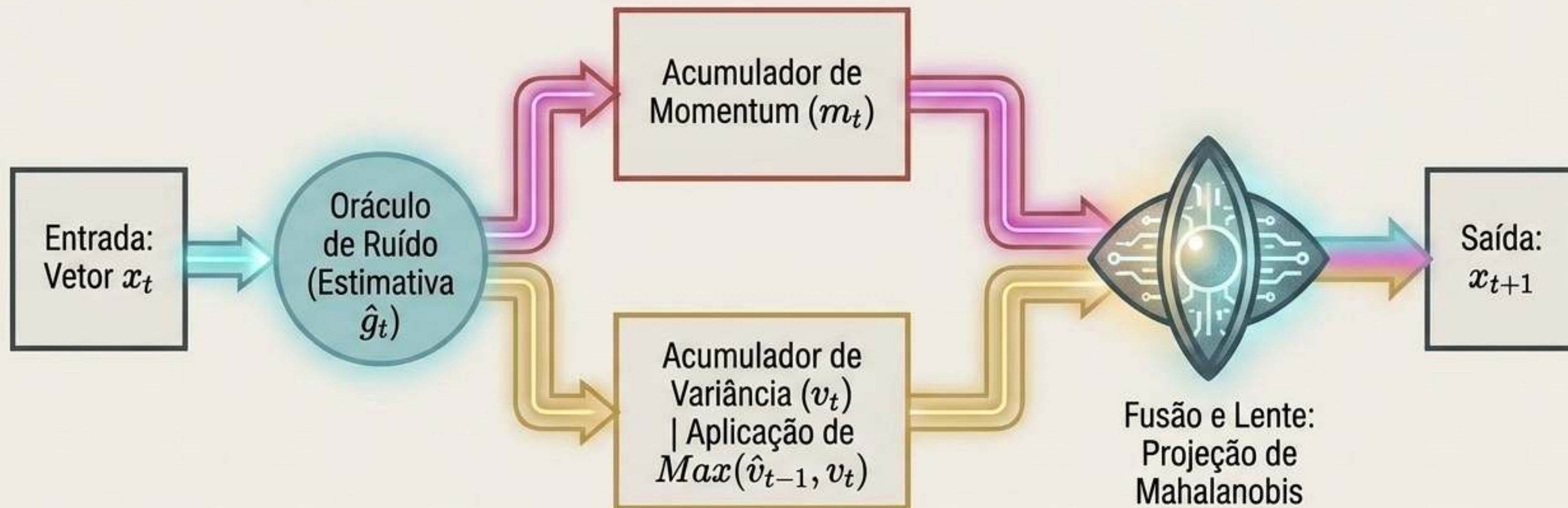
Subtext: $u \sim \mathcal{N}(0, I)$. Vetores ilimitados introduzem instabilidade letal na convergência com taxas adaptativas.

Oráculo Uniforme (ZO-AdaMM)



Subtext: $u \sim U_B$. A distribuição restrita garante a delimitação do espaço, controlando a norma L-infinito.

A Dinâmica Interna do ZO-AdaMM

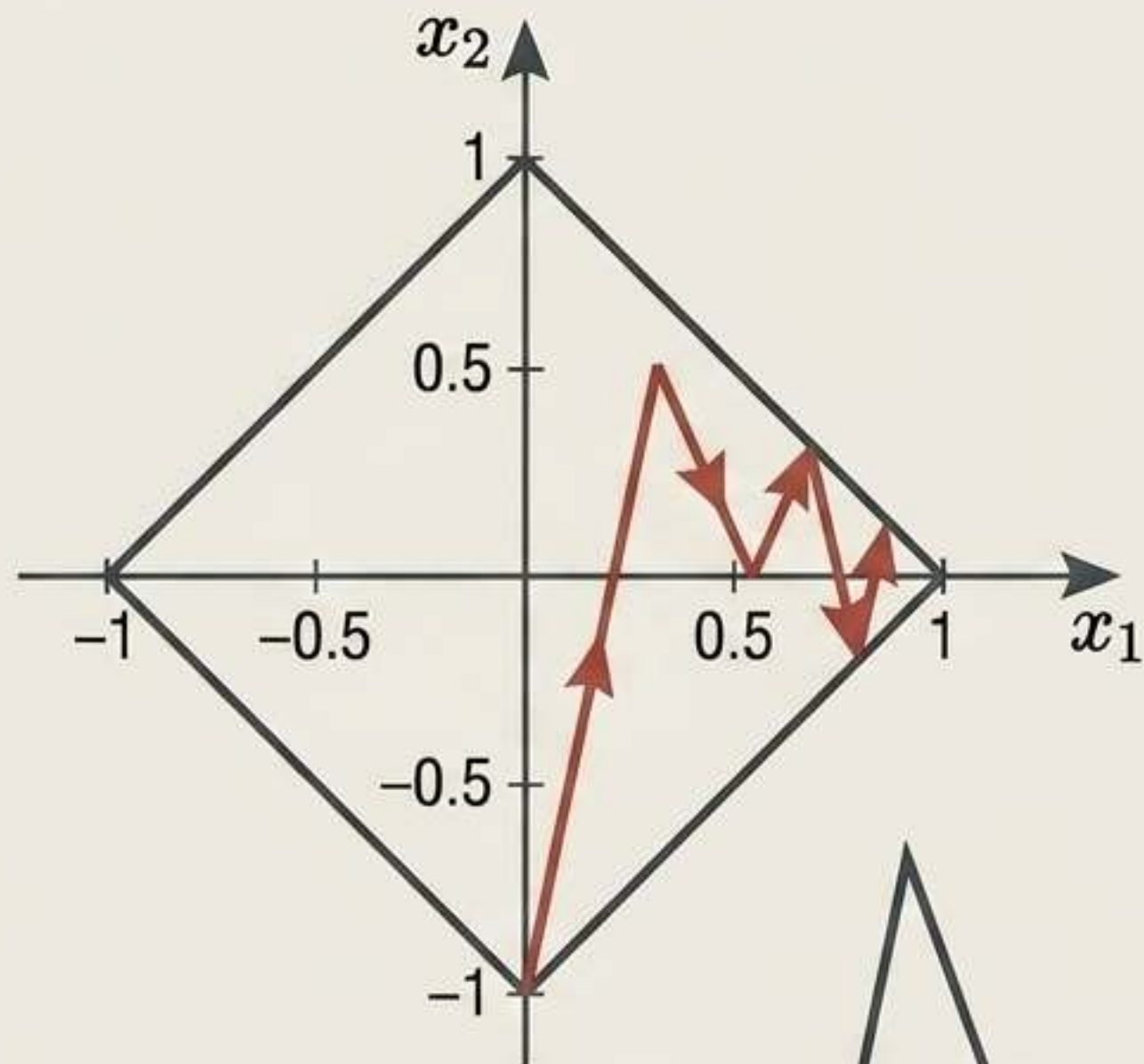


O Paradoxo da Projeção Euclidiana

Problema: Minimizar $-2x_1 - x_2$

Restrição: Norma L1 ($|x_1 + x_2| \leq 1$)

Inicialização: $x = [0.5, 0.5]^T$

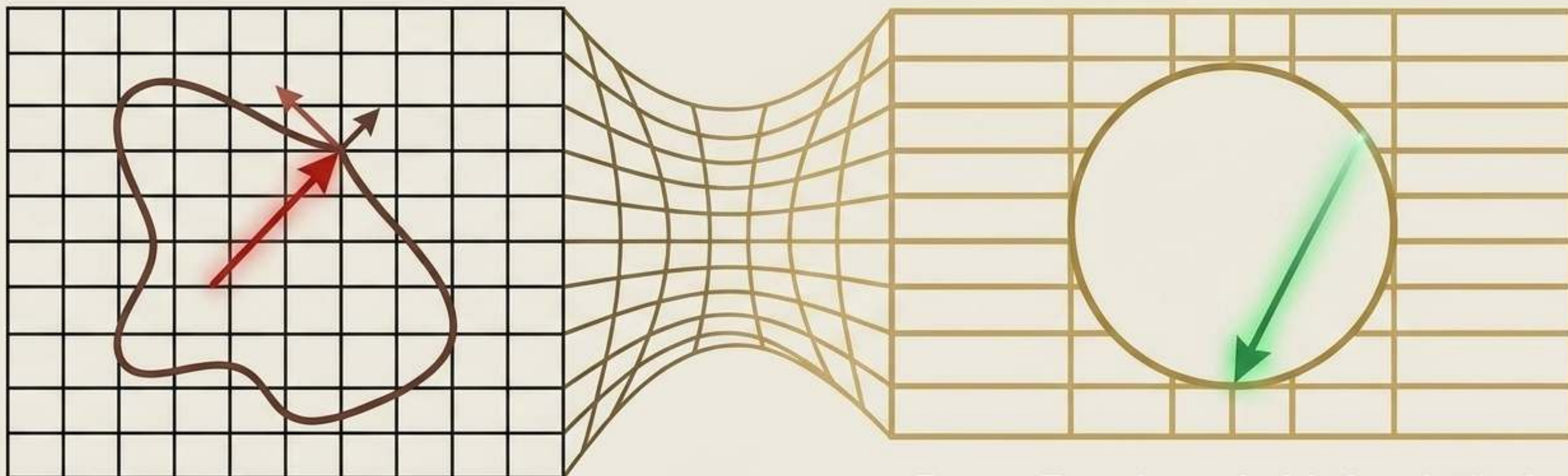


Conclusão: O uso de projeção Euclidiana em métodos AdaMM anula a convergência, prendendo o vetor em uma falsa trajetória.

A Lente de Mahalanobis e o Mapeamento de Gradiente

$$x^+ = \arg \min_{x \in X} \left\{ \langle g, x \rangle + \frac{1}{\omega} D_H(x, x^-) \right\}$$

D_H = Distância de Mahalanobis parametrizada pela Variância.



Espaço Original (x): Geometria restrita colide com taxas adaptativas.

Espaço Transformado (y): A projeção de Mahalanobis deforma a topologia, permitindo um deslizamento exato para o ótimo.

Garantias Teóricas de Convergência

Cenário Não-Convexo (Teoremas 1 e 2)

Sem restrição:

$$\mathcal{O}\left(\frac{\sqrt{d}}{\sqrt{T}} + \frac{d^{1.5}}{T}\right)$$

Com restrição:

$$\mathcal{O}\left(\frac{d}{T} + \frac{d}{b} + \frac{d^2}{q}\right)$$

A projeção induz viés estocástico. Exige estimador de variância reduzida com mini-batches.

Cenário Convexo (Proposição 4)

Avaliando o regret médio R_T :

$$\text{Convergência a } \mathcal{O}\left(\frac{d^{1.5}}{\sqrt{T}}\right)$$

O Insight Central: O ZO-AdaMM sofre uma penalidade dimensional estrutural de $\mathcal{O}(\sqrt{d})$ em relação à primeira ordem. Trata-se de um limite matemático fundamental imposto pela concentração geométrica em esferas de alta dimensão.

Matriz Diagnóstica: O Estado da Arte em ZO

Algoritmo	Oráculo	Suavização (μ)	Taxa	Consultas
ZO-SGD	Gaussiano	Restrita	Lenta	$\mathcal{O}(T)$
ZO-SCD	Coordenadas	Restrita	Lenta	$\mathcal{O}(dT)$
ZO-signSGD	Uniforme	Ampla	Viés Assintótico	$\mathcal{O}(bqT)$
ZO-AdaMM	Uniforme	$\mathcal{O}(1/\sqrt{dT})$	Sem Viés Residual	$\mathcal{O}(T)$

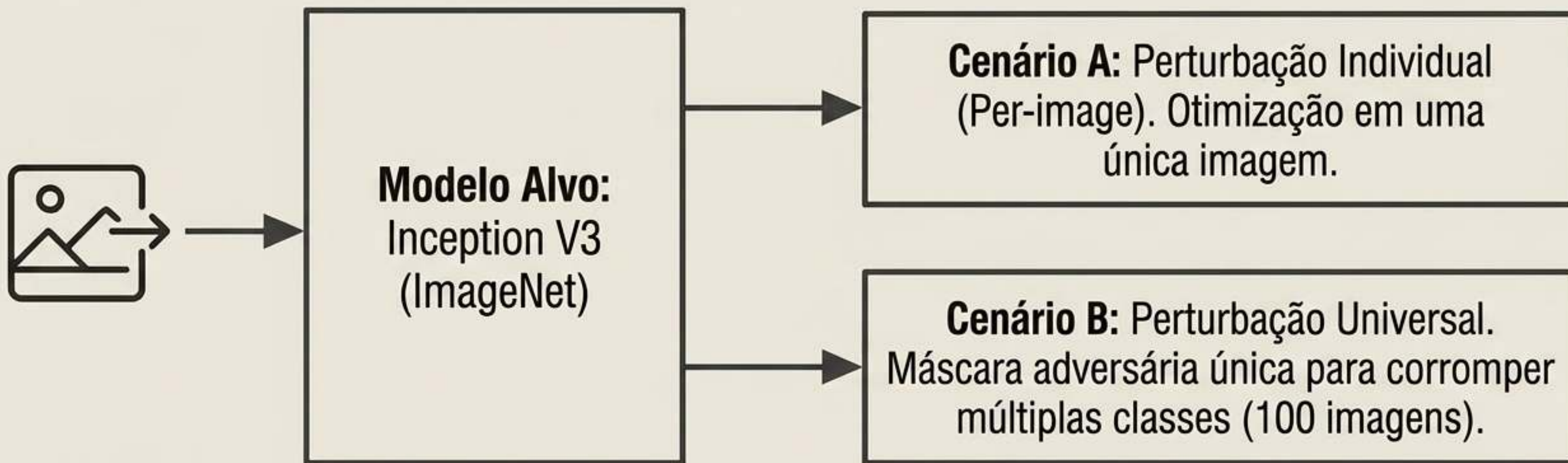
O **ZO-AdaMM** alcança o equilíbrio perfeito: combina a eficiência direcional do momentum com a viabilidade teórica sem viés residual.

O Campo de Batalha: Ataques Adversários



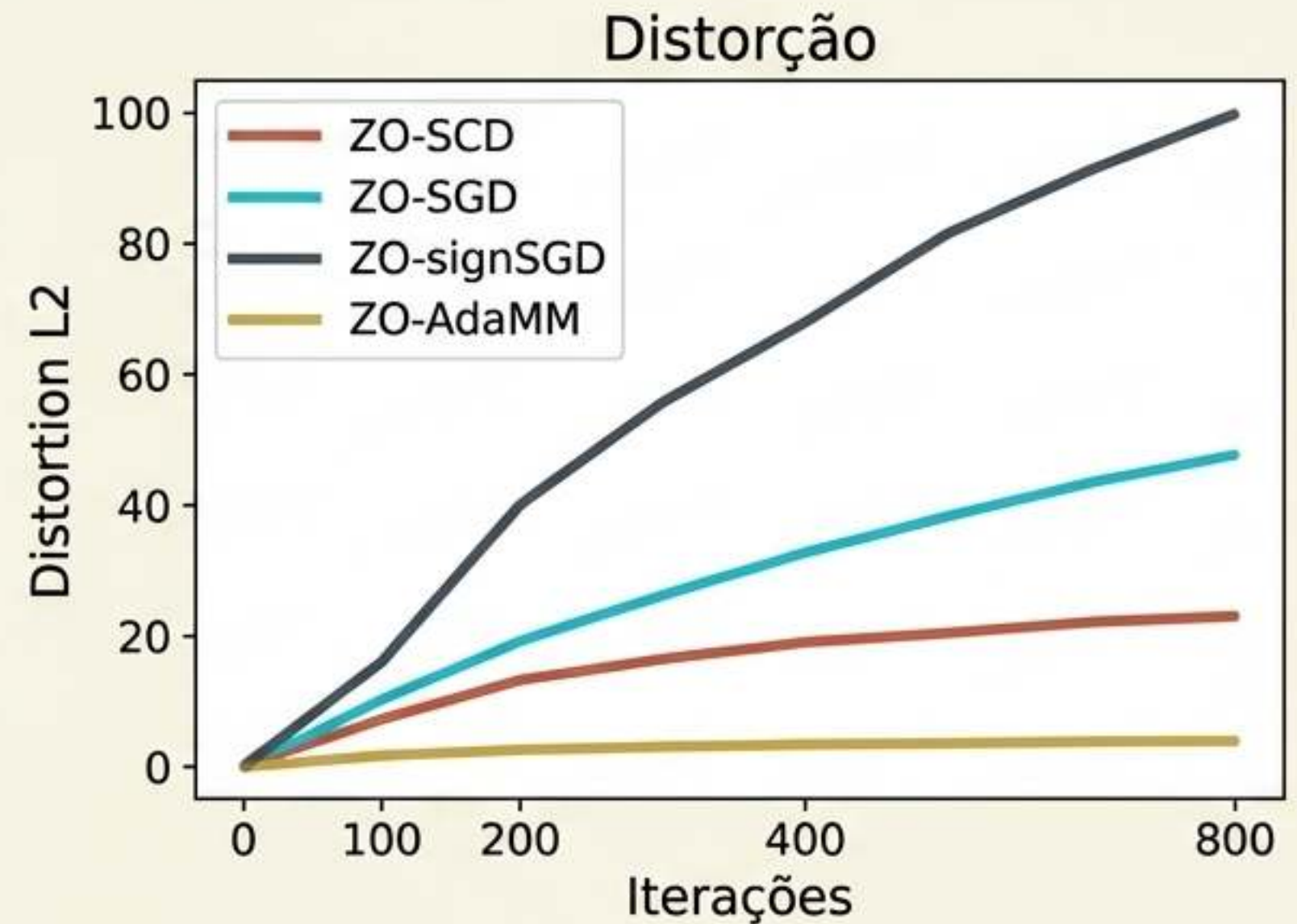
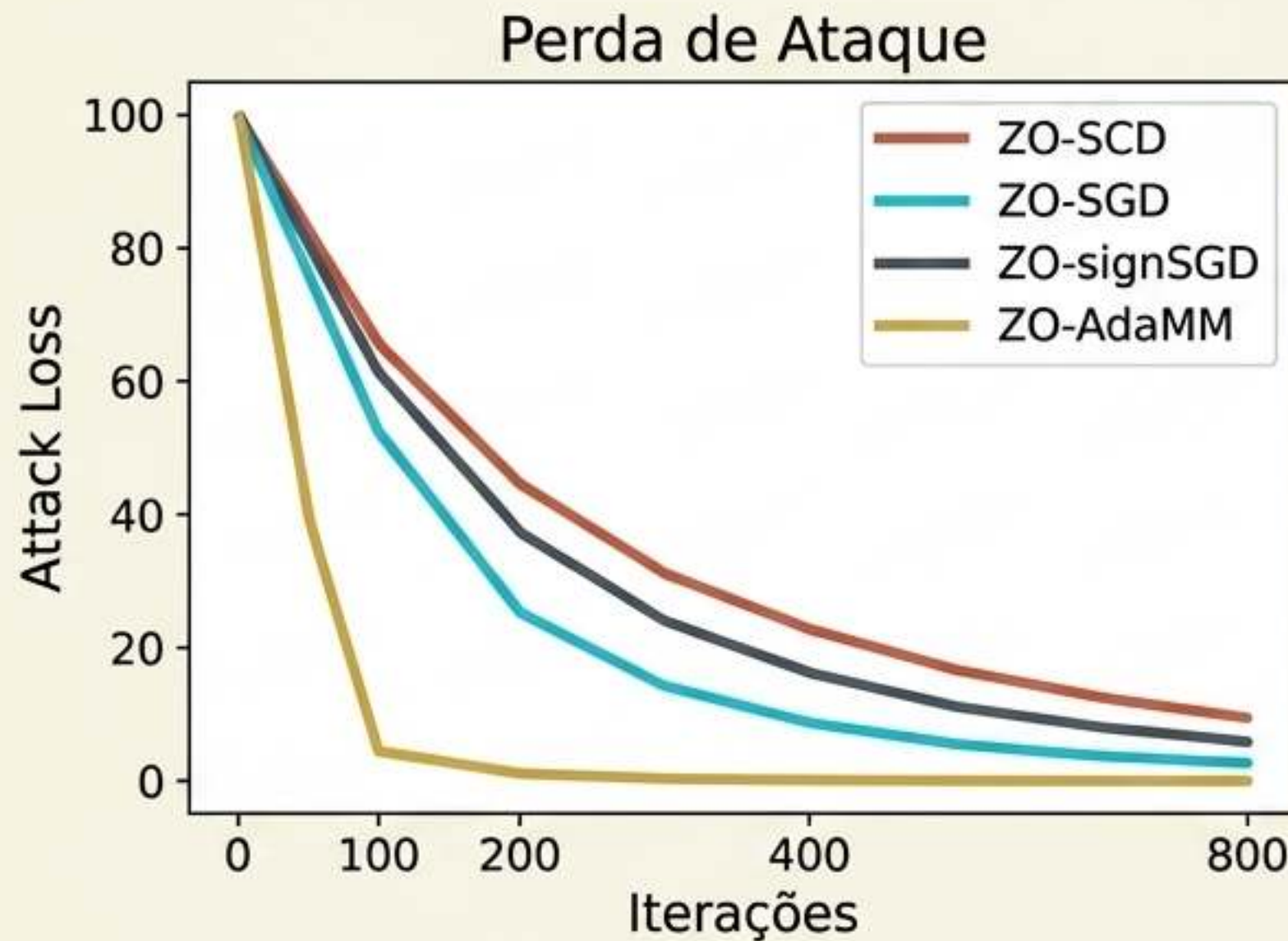
Desafio de Caixa-Preta

O modelo devolve apenas a probabilidade final.
Sem arquitetura, sem backpropagation.



Competidores Analisados: ZO-SGD | ZO-SCD | ZO-signSGD | ZO-NES | ZO-PSGD | ZO-SMD

Resultados Empíricos: Ataques Individuais (Inception V3)



O ZO-AdaMM exibe o melhor trade-off termodinâmico entre velocidade de evasão (Loss 0) e energia de distorção mínima.

Escalabilidade Tática: Ataques Universais

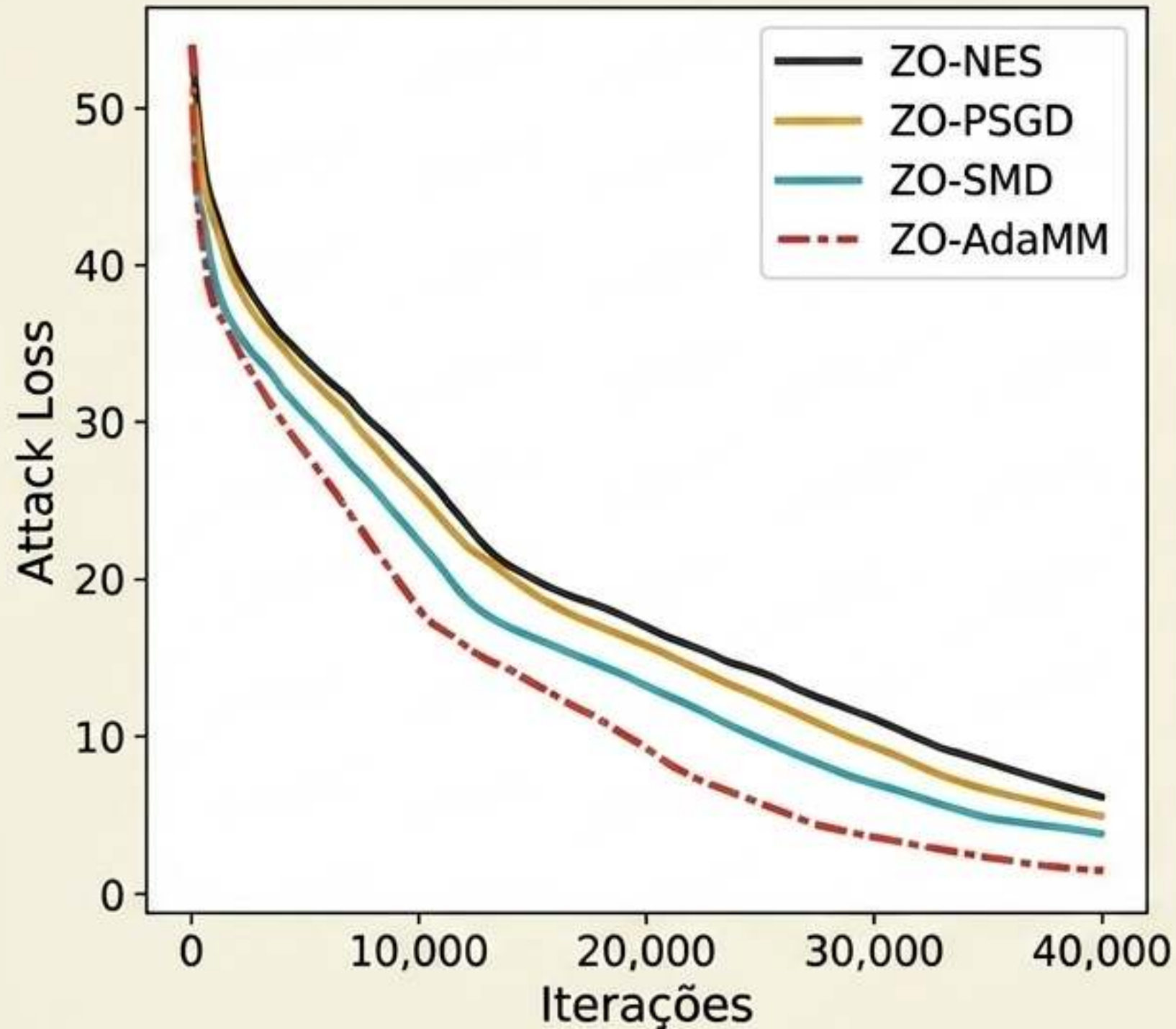


Tabela de Domínio

ZO-NES: 74% de sucesso | Distorção: 67.74

ZO-PSGD: 78% de sucesso | Distorção: 49.92

ZO-SMD: 79% de sucesso | Distorção: 47.36

ZO-AdaMM: 84% de sucesso | Distorção: 38.40

O ZO-AdaMM é o único algoritmo a combinar o maior sucesso tático com o menor footprint de distorção no espaço universal.

A Barreira do Agente Único



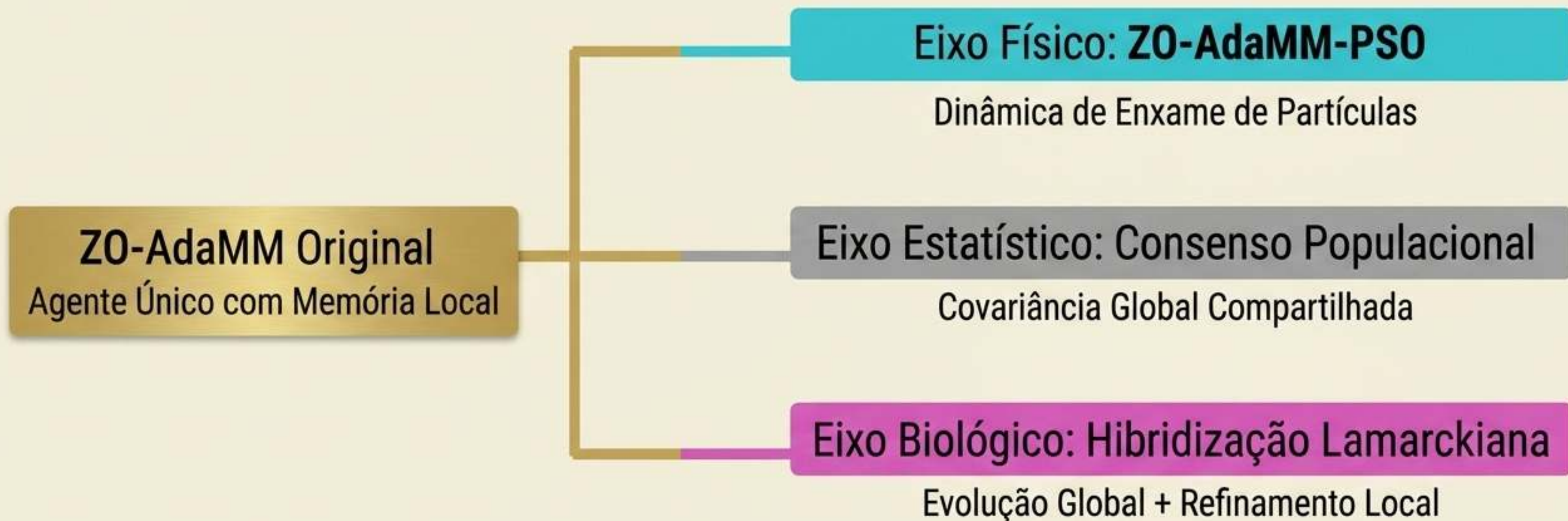
O Paradoxo da Exploração:

Por mais ciente do seu próprio histórico (matriz de variância) que o agente seja, a topologia de altíssima dimensão cria um gargalo físico. O limite de eficiência atinge um teto rígido penalizado pelo fator $\mathcal{O}(\sqrt{d})$.

A Pergunta de Transição:

E se não otimizarmos o agente, mas sim a inteligência colaborativa do próprio espaço geométrico?

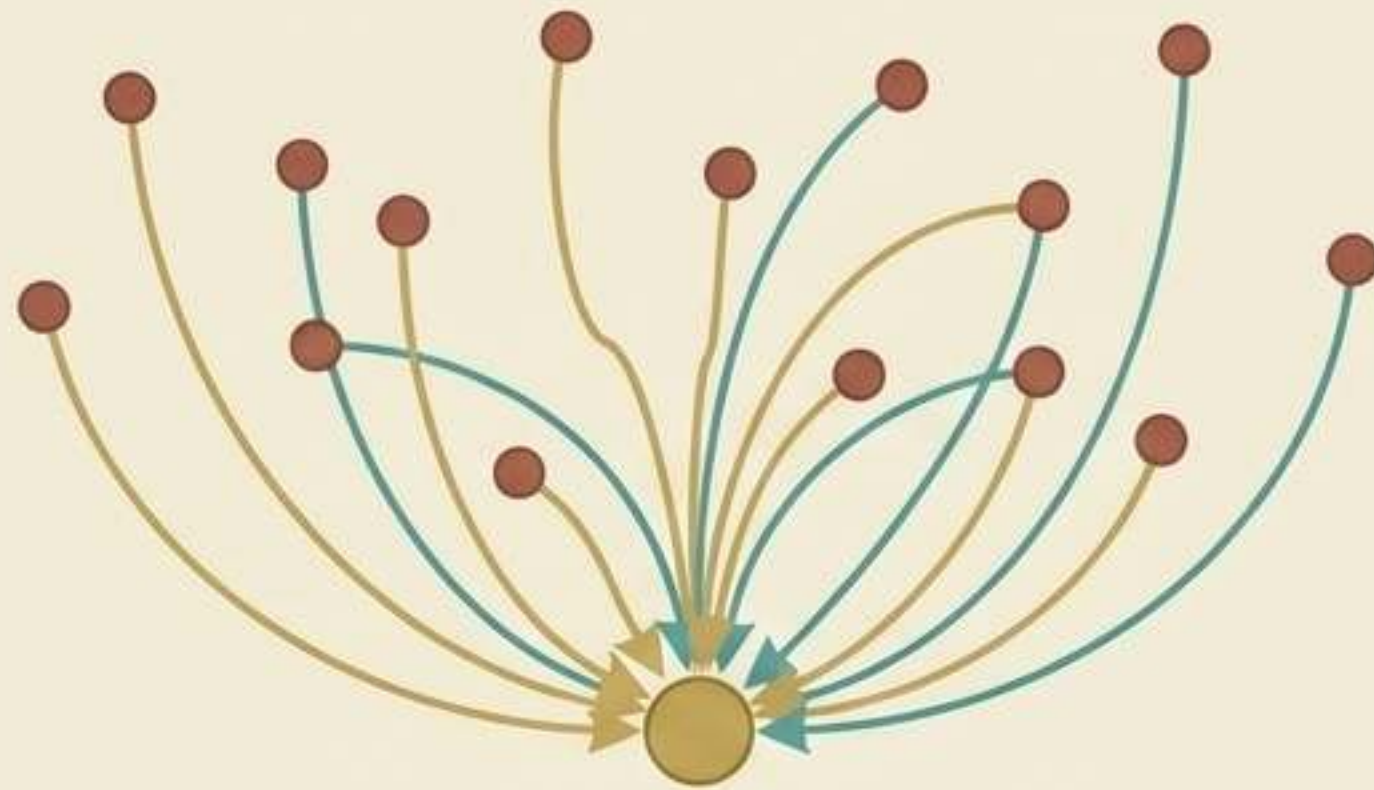
A Fronteira do Enxame: Pop-ZO-AdaMM



A transição definitiva da navegação estocástica isolada para a inteligência paramétrica massivamente distribuída.

Eixos de Evolução: Colaboração Dinâmica e Estatística

Caminho A: ZO-AdaMM-PSO

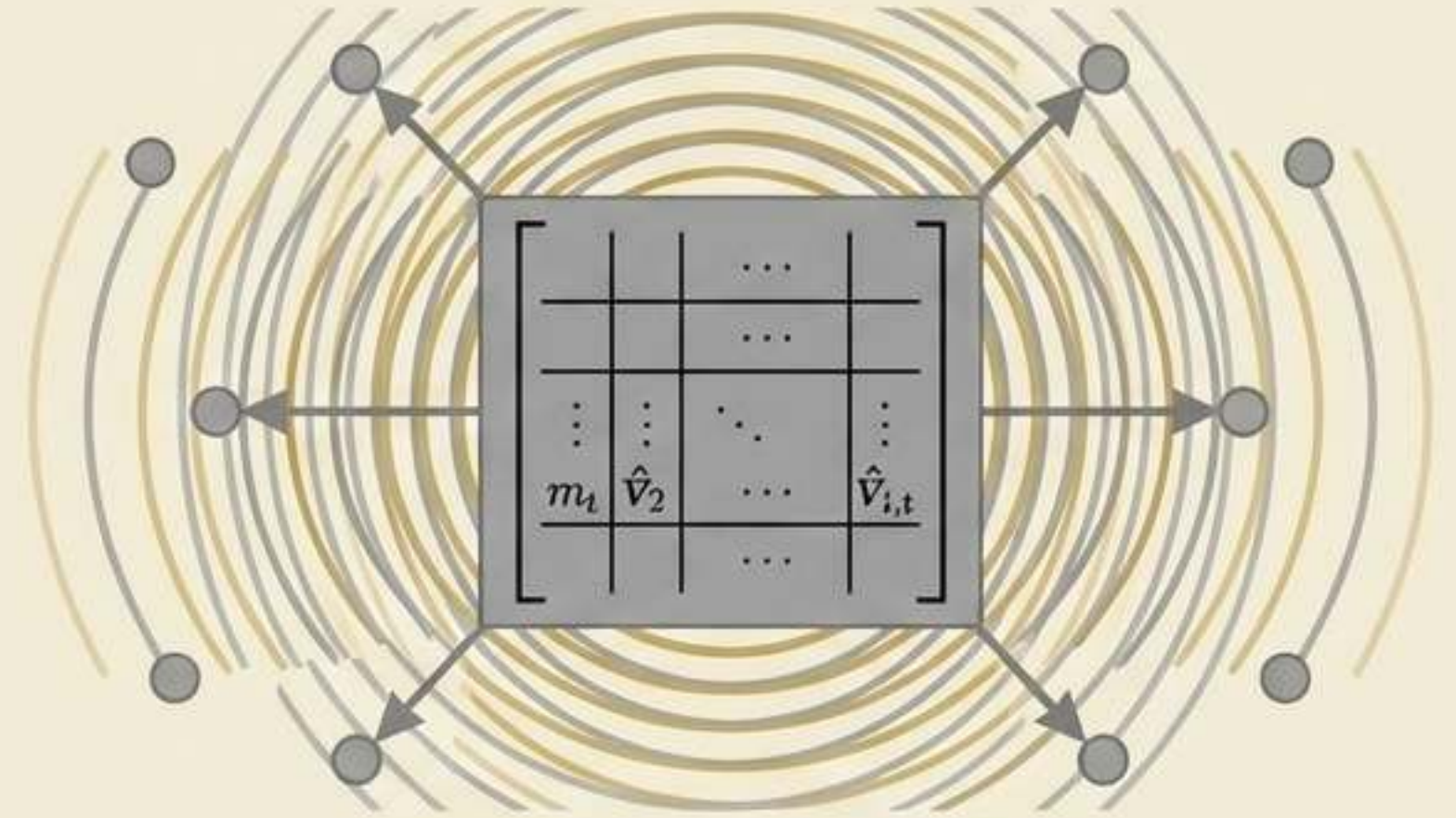


Mecânica Físico-Gravitacional:

$$m_t \leftarrow f(m_{\text{local}}, m_{\text{global_swarm}})$$

O vetor não confia apenas em sua memória. Ele é puxado por um momento acoplado, evitando mínimos locais através da gravidade do melhor vetor do enxame.

Caminho B: Consenso Populacional



Mecânica Estatística:

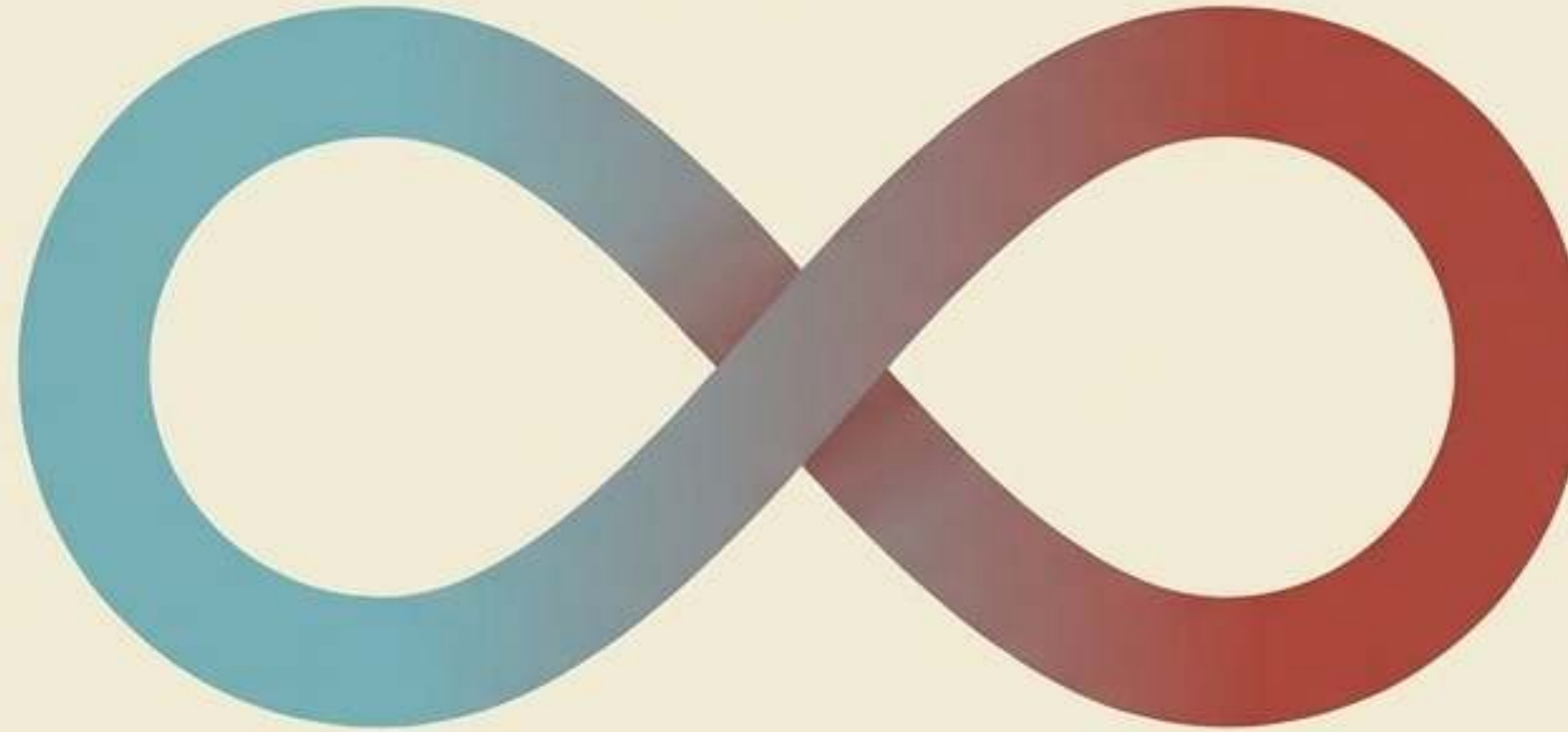
$$\hat{V}_{\text{global}} = \sum w_i \hat{V}_{i,t}$$

Compartilhamento da topologia de erro. A Matriz de Mahalanobis deixa de ser egoísta e deforma a projeção perfeitamente para todos os agentes simultaneamente.

O Ápice Híbrido: Evolução Lamarckiana em ZO

Fase Global (Cruzamento Genético)

Algoritmos genéticos espalham uma população massiva pelo espaço de busca indiscriminadamente.



Fase Local (A Magia Lamarckiana)

Cada 'filho' aplica o ZO-AdaMM como motor hiper-veloz para absorver sua bacia de atração. Os gradientes acumulados (m_t) são devolvidos à matriz genética.

Síntese Final

O Pop-ZO-AdaMM não apenas navega na névoa matemática; ele assimila e reprograma a própria topologia do problema. Uma nova era para a Otimização de Caixa-Preta irrestrita.