BREAKING THE PSEUDO-MULTI-AGENT ILLUSION: A PARASITE–IMMUNE MULTI-AGENT TRADING FRAMEWORK

Anonymous authors
Paper under double-blind review

ABSTRACT

Large language model (LLM) based multi-agent systems offer a new paradigm for collaborative decision-making in complex financial environments. However, existing systems often exhibit a "pseudo-multi-agent illusion": despite differing role labels, prompts, or names, agents rely on nearly identical information perspectives and produce highly correlated strategies. To address this issue, we propose a parasite-immune multi-agent trading (PI-MAT) framework. PI-MAT organizes trend-following, mean-reversion, and event-driven alpha agents into a heterogeneous cognitive ecosystem. Agents generate complementary signals from dedicated information perspectives, while parasitic replication, controlled mutation, and symbiotic complementarity expand strategic diversity. An online-performance-feedback-driven meta-learning arbiter forms the framework's cognitive immune layer, dynamically suppressing low-quality signals and reallocating capital to enable adaptive strategy integration in non-stationary markets. In a preregistered out-of-sample evaluation covering 2020–2023, PI-MAT achieved an annualized net return of 18.72%, a Sharpe ratio of 0.729, and a maximum drawdown of 28.61%. Relative to a price-only baseline, PI-MAT improved annualized net returns by 3.52 percentage points, with a paired 95% confidence interval of $[0.51, 6.89]$ percentage points. The code for PI-MAT is publicly available at `https://github.com/cxw0928/parasite-immune-framework-corecode`.

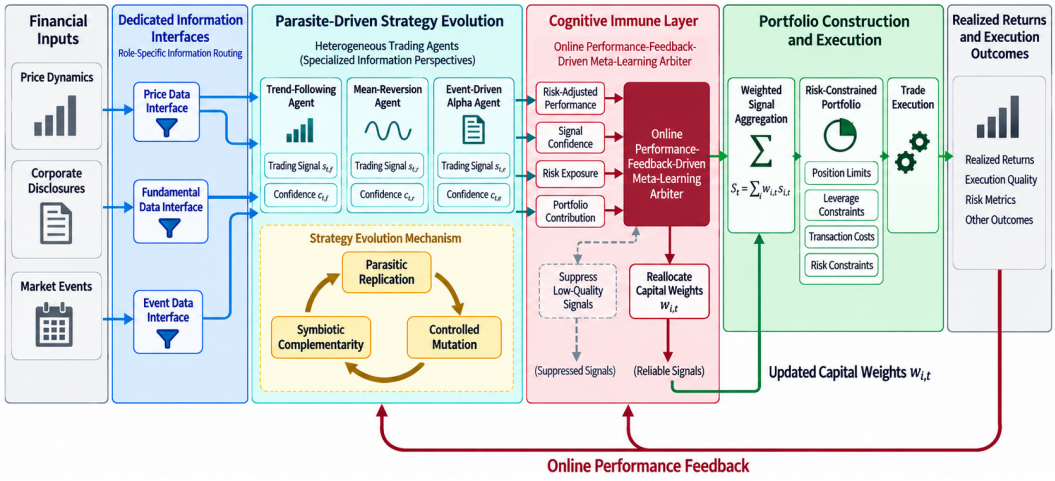


Figure 1: Overview of the PI-MAT framework. Agent-specific information interfaces support heterogeneous signal generation across trend-following, mean-reversion, and event-driven alpha agents. Parasite-driven strategy evolution expands strategic diversity, while the cognitive immune layer suppresses low-quality signals and dynamically reallocates capital based on online performance feedback.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

1 INTRODUCTION

Financial trading is fundamentally a sequential decision-making problem under incomplete information and non-stationary market conditions Qi et al. (2026); Merton (1980). Market participants must jointly process structured price data and unstructured financial text while constructing portfolios under limited information, predictive uncertainty, and conflicting strategies. Traditional statistical and factor models provide transparent and stable baseline signals, but have limited capacity to process complex semantic information and adapt to structural changes Bollerslev (1986); Sirignano & Cont (2019). Although deep learning models enhance representational capacity, they may compromise decision transparency and economic interpretability.

The development of large language models has introduced new approaches to financial information processing Vaswani et al. (2017); Devlin et al. (2019); Brown et al. (2020); Ouyang et al. (2022). Financial-domain models can extract semantic information from financial reports, corporate disclosures, and news, while supporting financial question answering, sentiment analysis, information extraction, and market prediction Chowdhery et al. (2022); Wu et al. (2023b); Yang et al. (2023); Wang et al. (2023a); Lopez-Lira & Tang (2023); Han et al. (2026). Recent systems have further incorporated domain adaptation, instruction tuning, multimodal inputs, and tool-augmented reasoning to combine textual evidence with structured market data. Other approaches augment language models with factor discovery, retrieval and memory modules, reflection, external tools, and reinforcement learning, enabling agents to retrieve historical experience, execute analytical operations, critique prior decisions, and refine actions using market feedback Schick et al. (2023); Yao et al. (2023); Shinn et al. (2023); Zhu & Zhu (2025); Yu et al. (2024a).

In recent years, financial large language models and multi-agent systems have been increasingly applied to financial information understanding, market prediction, portfolio construction, and trading decision-making Zhang et al. (2024); Yu et al. (2024b); Xiao et al. (2025). Representative systems explore multimodal financial agents, synthesized LLM collaboration, layered memory, dynamic portfolio risk management, offline reinforcement learning, and risk-sensitive optimization Zhang et al. (2024); Yu et al. (2024b;a); Li et al. (2024); Lee & Moon (2023); Lin et al. (2024). Financial benchmarks and deployment audits have also emphasized reproducible evaluation, model comparison, and the distinction between reported alpha and deployable performance Li et al. (2025); Ye et al. (2026). Existing systems commonly use role assignment, hierarchical planning, memory mechanisms, debate, and communication protocols to distribute information collection, analytical reasoning, and investment decisions across agents Wooldridge & Jennings (1995); Stone & Veloso (2000); Li et al. (2023); Wu et al. (2023a); Hong et al. (2024); Qian et al. (2024); Park et al. (2023).

Despite improving the modularity of financial decision-making, current multi-agent approaches face two major challenges:

1. **Insufficient informational and strategic separation.** Nominal role separation does not necessarily imply genuine diversity. Agents may consume the same price series, news streams, prompts, and intermediate conclusions, thereby producing highly correlated signals. This creates a "pseudo-multi-agent" illusion in which increasing the number of agents does not yield economically meaningful strategy diversity Du et al. (2023); Wang et al. (2023b; 2024).

2. **Rigid agent coordination.** Existing coordination mechanisms typically rely on static voting, discussion, fixed aggregation operators, or manually specified portfolio rules. These mechanisms cannot allocate capital dynamically in response to changes in agent performance, nor can they reallocate exposure when market regimes and signal quality evolve. This rigidity limits the ability of multi-agent trading systems to adapt to non-stationary market conditions.

To address these challenges, we propose the Parasite–Immune Multi-Agent Trading (PI-MAT) framework. PI-MAT organizes trend-following, mean-reversion, and event-driven alpha agents into a heterogeneous cognitive ecosystem. By assigning dedicated information perspectives, the framework enables agents to generate complementary signals. It further expands strategic diversity through parasitic replication, controlled mutation, and symbiotic complementarity. In addition, PI-MAT introduces an online-performance-feedback-driven meta-learning arbiter that forms a cog-

nitive immune layer. This arbiter dynamically suppresses low-quality signals and reallocates capital, enabling adaptive strategy integration in non-stationary markets.

The main contributions of this work are as follows:

1. **Heterogeneous strategy generation.** To address the "pseudo-multi-agent" illusion in existing multi-agent systems, we propose the Parasite–Immune Multi-Agent Trading (PI-MAT) framework. PI-MAT organizes trend-following, mean-reversion, and event-driven alpha agents into a heterogeneous cognitive ecosystem, enabling complementary signal generation through dedicated information perspectives and expanding strategic diversity through parasitic replication, controlled mutation, and symbiotic complementarity;
2. **Adaptive and reversible capital coordination.** To overcome the static and rigid nature of existing agent-coordination strategies, we introduce an online-performance-feedback-driven meta-learning arbiter that forms PI-MAT's cognitive immune layer. The arbiter dynamically suppresses low-quality signals and reallocates capital, thereby enabling adaptive strategy integration in non-stationary market environments;
3. **Experimental validation.** We conduct a rigorous evaluation using a preregistered out-of-sample test covering 2020–2023. PI-MAT achieves an annualized net return of 18.72%, a Sharpe ratio of 0.729, and a maximum drawdown of 28.61%. Relative to a price-only baseline, PI-MAT improves annualized net returns by 3.52 percentage points, with a paired 95% confidence interval of $[0.51, 6.89]$ percentage points.

2 Overview of the PI-MAT Framework

To address the "pseudo-multi-agent" phenomenon in existing multi-agent systems—where agents differ only in role labels, prompts, or names, while exhibiting highly similar information perspectives and strategy outputs—we propose the PI-MAT (Parasite–Immune Multi-Agent Trading) framework. The framework is illustrated in figure 1.

At decision time t, the system receives price dynamics, corporate disclosures, and market events:

$$\mathcal{O}_t = \left(\mathbf{x}_t^p, \mathbf{x}_t^f, \mathbf{x}_t^e\right),$$

where the superscripts p, f, and e denote price, fundamental, and event information. Dedicated interfaces transform these inputs into agent-specific representations:

$$\mathbf{z}_{i,t} = \mathcal{I}_i\left(\mathcal{O}_t, \mathcal{H}_t\right), \qquad i = 1, \ldots, N,$$

where $\mathcal{H}_t$ denotes the historical market and trading state. The agents shown in figure 1 correspond to $i \in \{f, r, e\}$, representing trend-following, mean-reversion, and event-driven strategies, respectively.

Each agent is assigned an information perspective $\mathcal{P}_i$ and trading objective $\mathcal{G}_i$, based on which it generates a signal and confidence estimate:

$$s_{i,t} = \pi_{\theta_i}\left(\mathbf{z}_{i,t}, \mathbf{h}_t \mid \mathcal{P}_i, \mathcal{G}_i\right), \qquad c_{i,t} = \Gamma_{\phi_i}\left(\mathbf{z}_{i,t}, \mathbf{h}_t\right).$$

The trend-following agent captures momentum and breakouts, the mean-reversion agent identifies equilibrium deviations, and the event-driven agent evaluates the pricing effects of disclosures, policies, and market events. Their outputs are coordinated by a stateful trading graph:

$$\mathbf{h}_{t+1} = \mathcal{F}_\psi\left(\mathbf{h}_t, \mathbf{X}_t, \mathbf{s}_t, \mathbf{c}_t\right),$$

where

$$\mathbf{s}_t = \left(s_{t,f}, s_{t,r}, s_{t,e}\right)^\top, \qquad \mathbf{c}_t = \left(c_{t,f}, c_{t,r}, c_{t,e}\right)^\top.$$

To maintain strategic diversity, PI-MAT employs parasitic replication, controlled mutation, and symbiotic complementarity. A high-performing agent i produces a candidate strategy through

$$\theta_k^{\mathrm{rep}} = \theta_i + \boldsymbol{\epsilon}_k, \qquad \boldsymbol{\epsilon}_k \sim \mathcal{D}_{\mathrm{mut}}(\sigma_i).$$

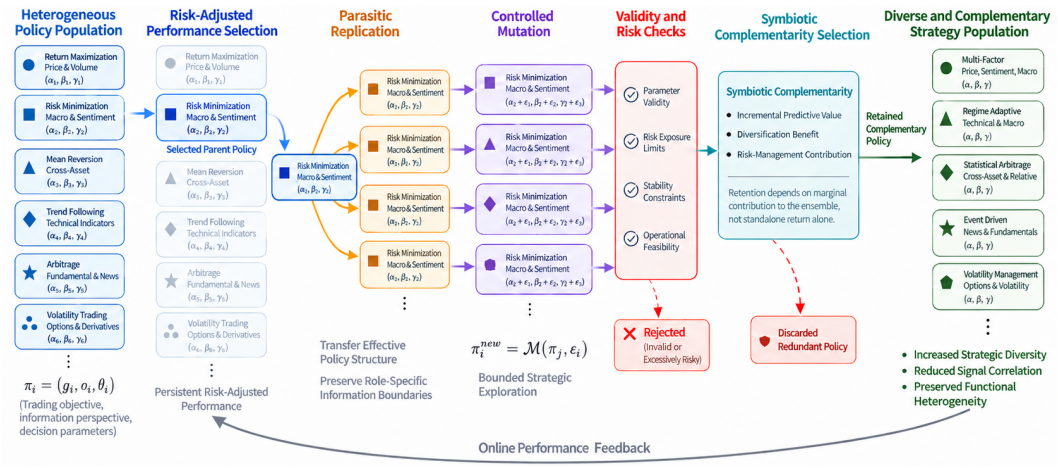


Figure 2: Parasite-driven strategy evolution in PI-MAT.

The cognitive immune layer reallocates capital according to risk-adjusted performance, signal confidence, and portfolio contribution:

$$w_{i,t} = \frac{\exp[\beta\,(\alpha q_{i,t} + \gamma c_{i,t} + \delta \Delta \mathcal{D}_{i,t})]}{\sum_{j=1}^{N} \exp[\beta\,(\alpha q_{j,t} + \gamma c_{j,t} + \delta \Delta \mathcal{D}_{j,t})]}.$$

Consequently, unreliable or excessively risky strategies are suppressed, whereas robust and complementary strategies receive greater capital weights.

The aggregate portfolio signal is

$$S_t = \sum_{i=1}^{N} w_{i,t} s_{i,t} = w_{f,t} s_{t,f} + w_{r,t} s_{t,r} + w_{e,t} s_{t,e}.$$

The target portfolio is obtained by solving

$$\mathbf{p}_t = \arg\max_{\mathbf{p}} \left\{ \mathbb{E}_t \left[\mathbf{p}^\top \mathbf{r}_{t+1} \right] - \lambda_{\text{risk}} \mathbf{p}^\top \mathbf{\Sigma}_t \mathbf{p} - \lambda_{\text{tc}} \left\| \mathbf{p} - \mathbf{p}_{t-1} \right\|_1 \right\},$$

subject to the relevant position, leverage, transaction-cost, and risk constraints.

Finally, realized returns, execution quality, and risk metrics are fed back to update strategy parameters and capital weights:

$$(\mathbf{w}_{t+1}, \Theta_{t+1}) = \mathcal{U}\left(\mathbf{w}_t, \Theta_t, \mathcal{Y}_{t+1}\right).$$

Consequently, the PI-MAT architecture enhances the adaptability, robustness, and interpretability of multi-agent trading systems under non-stationary financial market conditions.

2.1 Parasite-Driven Strategy Evolution

PI-MAT evolves a heterogeneous population of trading policies through four coupled stages: risk-adjusted selection, parasitic replication, controlled mutation, and symbiotic complementarity. Figure 2 illustrates the mechanism of parasite-driven strategy evolution. The policy of agent i is represented as

$$\pi_i = (g_i, o_i, \theta_i), \qquad \theta_i = (\alpha_i, \beta_i, \gamma_i),$$

where g_i denotes the trading objective, o_i the information perspective, and θ_i the decision-parameter vector. The components $(\alpha_i, \beta_i, \gamma_i)$ represent strategy-specific coefficients, such as signal weights, risk preferences, or reasoning parameters.

Each policy is evaluated over a rolling window using

$$\mathcal{S}_i = \widehat{\mu}_i - \lambda_\sigma \widehat{\sigma}_i - \lambda_{\text{CVaR}} \widehat{\text{CVaR}}_{\alpha,i} + \lambda_{\text{IC}} \widehat{\text{IC}}_i - \lambda_{\text{TO}} \widehat{\text{TO}}_i,$$

where CVaR denotes conditional value-at-risk and captures tail-loss exposure, IC denotes the information coefficient measuring the correlation between predicted and realized returns, and TO denotes portfolio turnover and its associated implementation cost. To emphasize persistent performance, PI-MAT updates

$$\overline{\mathcal{S}}_i^{(t)} = (1-\rho)\overline{\mathcal{S}}_i^{(t-1)} + \rho\mathcal{S}_i^{(t)},$$

and selects parent policies according to

$$p_i^{\text{parent}} = \frac{\exp(\overline{\mathcal{S}}_i/\tau)}{\sum_{k\in\mathcal{P}}\exp(\overline{\mathcal{S}}_k/\tau)}.$$

Here, ρ controls temporal smoothing and τ governs the exploration–exploitation trade-off. Thus, policies with persistent risk-adjusted performance are more likely to be selected while population diversity is retained.

For a selected parent j, parasitic replication transfers its effective policy structure while preserving its role-specific information perspective. Controlled mutation is then applied directly to the parent policy:

$$\pi_j = (g_j, o_j, \theta_j) \longrightarrow \pi_i^{\text{new}} = \mathcal{M}(\pi_j, \epsilon_i), \qquad \theta_i^{\text{new}} = \theta_j + \epsilon_i.$$

Here,

$$\epsilon_i = (\epsilon_\alpha, \epsilon_\beta, \epsilon_\gamma), \qquad \|\epsilon_i\|_W \le \delta_i,$$

and W is a positive-definite scaling matrix. Equivalently,

$$(\alpha_i^{\text{new}}, \beta_i^{\text{new}}, \gamma_i^{\text{new}}) = (\alpha_j + \epsilon_\alpha, \beta_j + \epsilon_\beta, \gamma_j + \epsilon_\gamma).$$

As shown below, the mutation radius adapts to population diversity, with declining diversity triggering broader strategic exploration.

$$\delta_i = \delta_{\min} + \kappa_\delta\big(1 - \mathcal{D}(\mathcal{P})\big),$$

Before admission, each candidate undergoes parameter-validity, exposure, tail-risk, stability, and operational-feasibility checks:

$$\mathbb{I}_{\text{valid}}(\pi_i) = \mathbb{I}[\theta_i \in \Theta,\ \text{Exposure}(\pi_i) \le E_{\max},\ \text{CVaR}_\alpha(\pi_i) \le C_{\max},\ \text{Stability}(\pi_i) \ge s_{\min}].$$

Candidates violating these constraints are rejected, preventing mutation from producing infeasible or excessively risky strategies.

Finally, PI-MAT evaluates policies by their marginal contribution to the active ensemble $\mathcal{A}$, rather than by standalone return alone. Let $U(\mathcal{A})$ denote ensemble utility, incorporating return, risk, cost, and diversity. The contribution of candidate i is

$$\Delta_i = U(\mathcal{A} \cup \{i\}) - U(\mathcal{A}).$$

A policy is retained when $\Delta_i \ge \gamma$ and its redundancy with existing agents remains below a prescribed threshold.

2.2 Cognitive Immune Layer

The cognitive immune layer serves as the reliability-aware coordination module of PI-MAT. As illustrated in figure 3, heterogeneous agents, including trend-following, mean-reversion, and event-driven experts, generate signals $s_{i,t}$ and confidence estimates $c_{i,t}$. These signals are combined into a portfolio-level decision,

$$S_t = \sum_{i=1}^{N} w_{i,t} s_{i,t},$$

which produces the realized net return

$$r_{t+1} = S_t^\top R_{t+1} - \mathcal{C}_{t+1},$$

where $\mathcal{C}_{t+1}$ includes transaction costs, slippage, and execution losses. The realized contribution of agent i is approximated by

$$\Delta r_{i,t+1} = w_{i,t} s_{i,t}^\top R_{t+1} - \lambda_{\text{int}}|w_{i,t}|\|s_{i,t} - S_t\|_2,$$

At each time step, the online meta-learning arbiter estimates agent reliability from recent contribution, risk exposure, confidence, and diversification benefit. Exponentially weighted statistics are updated as

$$\widehat{\mu}_{i,t} = (1-\rho)\widehat{\mu}_{i,t-1} + \rho\Delta r_{i,t},$$

$$\widehat{\sigma}^2_{i,t} = (1-\rho)\widehat{\sigma}^2_{i,t-1} + \rho(\Delta r_{i,t} - \widehat{\mu}_{i,t})^2,$$

where ρ controls the adaptation rate. The resulting information state is

$$x_{i,t} = [\widehat{\mu}_{i,t}, -\widehat{\sigma}_{i,t}, c_{i,t}, \widehat{\Delta r}_{i,t}, \widehat{\mathrm{Div}}_{i,t}]^\top.$$

The meta-learner computes

$$q^{\mathrm{raw}}_{i,t} = f_{\theta_t}(x_{i,t}),$$

yields smoothed, risk-adjusted reliability state

$$q_{i,t} = \gamma q_{i,t-1} + (1-\gamma)q^{\mathrm{raw}}_{i,t} - \lambda_\sigma \widehat{\sigma}_{i,t} - \lambda_d \widehat{\mathrm{MDD}}_{i,t}.$$

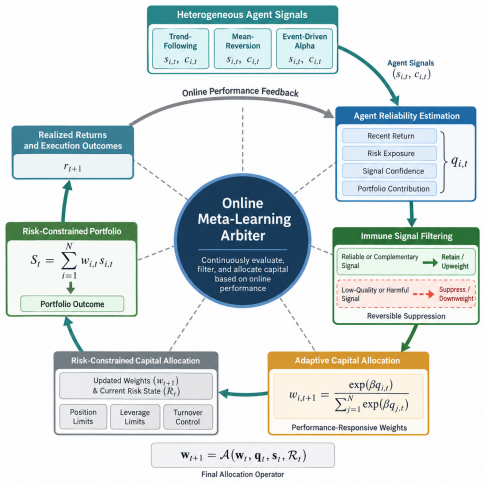


Figure 3: Core mechanism of the PI-MAT cognitive immune layer.

Here, $\widehat{\mathrm{MDD}}_{i,t}$ denotes the estimated maximum drawdown over the recent evaluation window and measures cumulative downside risk. Immune filtering is implemented through the reversible gate

$$g_{i,t} = \sigma(\kappa(q_{i,t} - \tau_t)), \qquad \widetilde{s}_{i,t} = g_{i,t}s_{i,t},$$

where τ_t is the reliability threshold. Reliable signals are retained or upweighted, whereas unreliable signals are progressively suppressed without permanently removing their agents. Performance-responsive weights are then computed as

$$w_{i,t+1} = \frac{\exp(\beta q_{i,t})}{\sum_{j=1}^{N}\exp(\beta q_{j,t})},$$

where β controls the sensitivity to reliability differences. The final allocation is

$$\mathbf{w}_{t+1} = \mathcal{A}_{\theta_t}(\mathbf{w}_t, \mathbf{q}_t, \mathbf{s}_t, \mathcal{R}_t),$$

where $\mathcal{R}_t$ summarizes current volatility, leverage, concentration, and turnover conditions. The resulting portfolio outcome is fed back to the arbiter, enabling continuous adaptation under non-stationary market conditions.

3 Experiments

We evaluate PI-MAT out of sample on daily AAPL data from 2020 to 2023. The experiments address three research questions: whether parasite-driven evolution produces substantively distinct trading strategies; whether the cognitive immune layer enables dynamic capital allocation based on online performance feedback; and whether these mechanisms improve overall trading performance. Performance is evaluated using annualized return, Sharpe ratio, maximum drawdown, cumulative wealth, and paired confidence intervals; details are provided in the appendix.

3.1 Does Parasite-Driven Evolution Improve Strategy Diversity?

We first examine whether the heterogeneous agents generate genuinely distinct trading views. As shown in table 1 panel (a), the pairwise correlations among the trend-following, mean-reversion, and event-driven proxy signals are low, with a mean absolute off-diagonal correlation of only 0.188. By contrast, the shared-input placebo produces perfectly correlated signals ($\rho = 1.000$). This comparison indicates that dedicated information interfaces substantially reduce signal redundancy and help alleviate the pseudo-multi-agent illusion. Figure 4 provides complementary evidence: the agents exhibit clearly separated cumulative-wealth paths, suggesting that their signals differ not only statistically but also in the timing and persistence of realized returns.

Table 1: Signal correlations and standalone out-of-sample performance of the three agents.

(a) Signal correlations

Agent	Trend	Mean-rev.	Event
Trend-following	1.000	−0.216	−0.114
Mean-reversion	−0.216	1.000	−0.234
Event-driven proxy	−0.114	−0.234	1.000

$\overline{|\rho|}_{\text{off}} = 0.188$

(b) Standalone performance

Agent	Ann. return (%)	Sharpe	Max. DD (%)
Trend-following	9.16	1.053	9.02
Mean-reversion	2.13	0.969	0.78
Event-driven proxy	224.84	8.050	0.03

Table 2: Adaptive allocation and reversible suppression statistics.

Agent	$\bar{w}$	SD(w)	w_T	$w < .05$ (%)	Episodes	Recovered	Recovery (%)	Median
Trend-following	0.302	0.370	0.798	19.96	12	12	100.0	15.0
Mean-reversion	0.383	0.396	0.000	21.52	15	14	93.3	9.0
Event-driven proxy	0.315	0.336	0.202	0.00	38	38	100.0	9.5

We next evaluate whether this signal heterogeneity translates into standalone economic value. Table 1 panel (b) reports the out-of-sample performance of the three agents under a fixed evaluation protocol. All agents achieve positive annualized returns. The trend-following and mean-reversion agents achieve annualized returns of 9.16% and 2.13%, with corresponding Sharpe ratios of 1.053 and 0.969, respectively. The event-driven proxy exhibits substantially higher performance, attaining an annualized return of 224.84%, a Sharpe ratio of 8.050, and a maximum drawdown of 0.03%. These results suggest that the differentiated strategies provide economically meaningful, although potentially heterogeneous, predictive value.

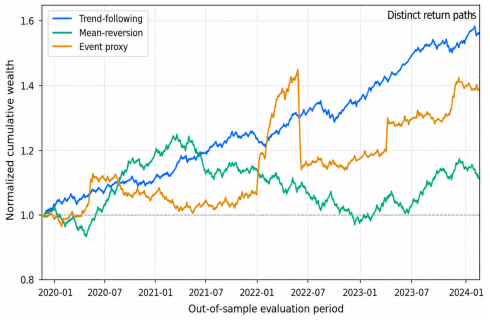
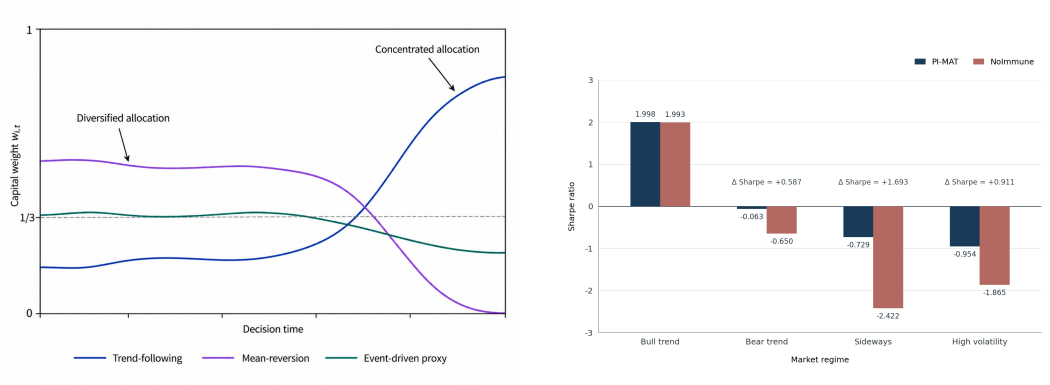


Figure 4: Normalized wealth paths of the three agents.

3.2 Does the Cognitive Immune Layer Enable Adaptation?

Figure 5 (a) and table 2 show that the arbiter does not maintain a static equal-weight allocation of $1/3$ across agents. Instead, it continuously adjusts capital weights according to online estimates of agent reliability. Over the evaluation period, the trend-following, mean-reversion, and event-driven proxy agents receive mean weights of 0.302, 0.383, and 0.315, respectively. At the final decision point, these weights become 0.798, 0.000, and 0.202, indicating a substantial reallocation toward the currently preferred expert. The suppression statistics further indicate that low-quality signals are temporarily downweighted. Near-zero allocation is defined as $w_{i,t} < 0.05$. Under this criterion, the trend-following and mean-reversion agents receive near-zero weights at 19.96% and 21.52% of decision points, respectively. Their subsequent recovery rates are 100.0% and 93.3%, while the event-driven proxy achieves a 100.0% recovery rate under the episode-level definition. Overall, these results demonstrate that the cognitive immune layer performs dynamic and reversible capital coordination: it concentrates capital in currently more reliable agents, temporarily suppresses deteriorating signals, and preserves the possibility of subsequent weight recovery.

Figure 5 (b) compares the Sharpe ratios of PI-MAT and the NoImmune ablation across four market regimes. PI-MAT achieves a higher Sharpe ratio in every regime, although the magnitude of the improvement varies substantially. The largest gain occurs in sideways markets, with a Sharpe difference of $+1.693$, followed by high-volatility periods ($+0.911$) and bear-trend periods ($+0.587$). The improvement is marginal in bull-trend periods ($+0.006$). These results indicate that the relative effectiveness of the cognitive immune layer is regime dependent, with larger risk-adjusted performance gains observed in sideways and high-volatility markets.



(a) Dynamic capital weights.

(b) Sharpe ratios by market regime.

Figure 5: Adaptive allocation and regime-dependent performance.

Table 3: Main Results, Ablations and Temporal Robustness of PI-MAT.

Panel A(Main Results): Pre-registered evaluation, 2020–2023

Method	Ann. net return (%)	Sharpe ratio	Max. DD (%)	Abs. gain (pp)	Rel. gain (%)	Paired 95% CI (pp)
PI-MAT	**18.72**	**0.729**	28.61	**+3.52**	**+23.16**	**[0.51, 6.89]**
Price-only (implied)	15.20	–	–	0.00	0.00	–

Panel B(Ablations and Temporal Robustness): Split-sample performance and rolling-window robustness

	Ann. return (%)		Sharpe ratio		PI-MAT win rate	
Method	Early	Late	Early	Late	Return	Sharpe
PI-MAT	**6.65**	**11.33**	**0.785**	1.027	–	–
NoEvo	−10.10	5.16	−1.018	**1.513**	9/16	8/16
NoImmune	−8.26	−3.27	0.098	−0.505	9/16	9/16
SharedInput	2.12	2.25	0.253	0.316	9/16	8/16
EqualWeight	−1.33	−5.12	−0.294	−1.358	10/16	7/16

3.3 Does PI-MAT Improve Trading Performance?

As shown in panel A of table 3, PI-MAT achieves an annualized net return of 18.72% and a Sharpe ratio of 0.729. Relative to the Price-only baseline, its annualized return increases by 3.52 percentage points, corresponding to a 23.16% relative improvement. The paired 95% confidence interval is [0.51, 6.89] percentage points and excludes zero, providing evidence of a positive return gain. PI-MAT achieves positive returns in both consecutive subperiods, with its annualized return increasing from 6.65% in the early half to 11.33% in the late half, while its Sharpe ratio rises from 0.785 to 1.027. By contrast, NoEvo (without strategy evolution), NoImmune (without the cognitive immune layer), and EqualWeight (fixed equal-weight allocation) each exhibit a negative return or Sharpe ratio in at least one subperiod. In the rolling-window analysis, PI-MAT achieves annualized-return win rates of 56.3%–62.5% against the respective baselines; its Sharpe-ratio win rates range from 43.8% to 56.3%. Overall, PI-MAT does not dominate across all periods and evaluation metrics, but adapts to changing market conditions while maintaining positive performance across subperiods and improving aggregate returns.

Table 4 shows that PI-MAT achieves a Sharpe ratio of 0.729, exceeding those of InvestorBench 0.479, MASA-MLP 0.27, and Sparse Mean–CVaR 0.253. Its cumulative return is 18.72%, higher than the 14.336% reported for InvestorBench and the 8.87% reported for MASA-MLP. Although FinAgent, FinMem, and TradingAgents report higher returns or Sharpe ratios, their results are obtained under different asset universes, evaluation periods, return definitions, and transaction-cost assumptions; consequently, they do not support a strict like-for-like ranking. Overall, PI-MAT achieves a relatively balanced profile of return and risk-adjusted performance, supporting the effectiveness of its heterogeneous-expert coordination and dynamic capital-allocation mechanisms.

Table 4: Reported quantitative results from recent financial-agent systems and PI-MAT. NR denotes not reported; N/A denotes not applicable.

Method **Citation**	**Task**	**Asset /** **data scope**	**Return (%)**	**Sharpe**	**MDD (%)**
FinAgent Zhang et al. (2024) KDD 2024	Multimodal LLM agent	AAPL, AMZN GOOGL, MSFT TSLA, ETHUSD	31.90	1.43	10.40
InvestorBench Li et al. (2025) ACL 2025	Financial-agent benchmark	Stocks, crypto, ETFs AAPL subset	14.336	0.479	12.029
TradingAgents Xiao et al. (2025) Preprint 2025	Role-based LLM trading	AAPL, GOOGL AMZN	26.62	8.21	0.91
MASA-MLP Li et al. (2024) AAMAS 2024	Portfolio risk management	CSI 300, DJIA S&P 500	8.87	0.27	31.78
TACR Lee & Moon (2023) AAMAS 2023	Offline RL trading	NDX, MDAX CSI	NR	Relative improvement 13.1–177.7%	NR
TRIBE Vidler & Walsh (2025) AAMAS 2025	Bond-market ABM	Australian government bonds	N/A	Asset trading 90/0/7%	N/A
Beliefs–Shocks–Roles Albers et al. (2024) AAMAS 2024	Asset-market ABM	30 agents/3 assets 500 agents/10 assets	N/A	N/A	Positive shock $+11.11\%$
FinMem Yu et al. (2024a) AAAI 2024	Memory-augmented LLM	Multi-source stocks 2021–2023	61.776	2.679	10.800
AlphaQCM Zhu & Zhu (2025) ICML 2025	Alpha discovery	CSI 300, CSI 500 A-share, S&P 500	8.49–9.55	NR	NR
Sparse Mean–CVaR Lin et al. (2024) ICML 2024	Sparse portfolio optimization	Six monthly factor datasets	NR	0.253	NR
Alpha Illusion Ye et al. (2026) Preprint 2026	Deployment audit	TSLA, NVDA, KO XOM, MSTR	2.30	0.43 gross 0.22 net	NR
PI-MAT This work	Adaptive multi-agent trading	AAPL 2020–2023	18.72	0.729	28.61

4 Conclusion

We presented PI-MAT, a parasite–immune multi-agent trading framework designed to address the pseudo-multi-agent illusion in LLM-based financial systems. By assigning dedicated information perspectives and differentiated objectives to trend-following, mean-reversion, and event-driven alpha agents, PI-MAT promotes genuinely complementary signal generation. Its parasite-driven strategy evolution mechanism further expands strategic diversity through replication, controlled mutation, and symbiotic complementarity. Meanwhile, an online-performance-feedback-driven meta-learning arbiter forms a cognitive immune layer that suppresses unreliable signals and reallocates capital according to time-varying expert performance. Out-of-sample results demonstrate that PI-MAT improves risk-adjusted trading performance over a price-only baseline in non-stationary markets. Future work will examine broader asset classes, longer evaluation horizons, and the framework's robustness and theoretical stability.

REFERENCES

Evan Albers, Mohammad T. Irfan, and Matthew J. Botsch. Beliefs, shocks, and the emergence of roles in asset markets: An agent-based modeling approach. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024)*, pp. 40–48, 2024.

Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986.

Tom B. Brown, Benjamin Mann, Nick Ryder, et al. Language models are few-shot learners. In *Advances in Neural Information Processing Systems*, 2020.

Aakanksha Chowdhery, Sharan Narang, Jacob Devlin, et al. Palm: Scaling language modeling with pathways. *arXiv preprint arXiv:2204.02311*, 2022.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL-HLT*, pp. 4171–4186, 2019.

Yilun Du, Shuang Li, Antonio Torralba, Joshua B. Tenenbaum, and Igor Mordatch. Improving factuality and reasoning in language models through multiagent debate. In *Proceedings of EMNLP*, 2023.

Jeong-Ji Han, Soyoung Jun, and Jong Woo Kim. Examining the impact of esg news sentiment on corporate performance: A comprehensive analysis by news topic and industry. *Business Ethics, the Environment & Responsibility*, 35(3):1624–1647, 2026. doi: 10.1111/beer.12855.

Sirui Hong, Mingchen Zhuge, Jonathan Chen, et al. Metagpt: Meta programming for multi-agent collaborative framework. In *International Conference on Learning Representations*, 2024.

Namyeong Lee and Jun Moon. Transformer actor-critic with regularization: Automated stock trading using reinforcement learning. In *Proceedings of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023)*, pp. 2815–2817, 2023. Extended abstract.

Guohao Li, Hasan Abed Al Kader Hammoud, Hani Itani, Dmitrii Khizbullin, and Bernard Ghanem. Camel: Communicative agents for mind exploration of large scale language model society. In *Advances in Neural Information Processing Systems*, 2023.

Haohang Li, Yupeng Cao, Yangyang Yu, Shashidhar Reddy Javaji, Zhiyang Deng, Yueru He, Yuechen Jiang, Zining Zhu, Koduvayur Subbalakshmi, Jimin Huang, Lingfei Qian, Xueqing Peng, Qianqian Xie, and Jordan W. Suchow. Investorbench: A benchmark for financial decision-making tasks with llm-based agents. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics*, volume 1, pp. 2509–2525, 2025.

Zhenglong Li, Vincent Tam, and Kwan L. Yeung. Developing a multi-agent and self-adaptive framework with deep reinforcement learning for dynamic portfolio risk management. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024)*, pp. 1174–1182, 2024.

Yizun Lin, Yangyu Zhang, Zhao-Rong Lai, and Cheng Li. Autonomous sparse mean-cvar portfolio optimization. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*. PMLR, 2024.

Alejandro Lopez-Lira and Yuehua Tang. Can chatgpt forecast stock price movements? return predictability and large language models. *arXiv preprint arXiv:2304.07619*, 2023.

Robert C. Merton. On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, 8(4):323–361, 1980.

Long Ouyang, Jeffrey Wu, Xu Jiang, et al. Training language models to follow instructions with human feedback. In *Advances in Neural Information Processing Systems*, 2022.

Joon Sung Park, Joseph C. O'Brien, Carrie J. Cai, et al. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (UIST)*, pp. 1–22, 2023. doi: 10.1145/3586183.3606763.

Yue Qi, Jianing Huang, and Yixuan Zhu. On intensively criticizing and envisioning the research on multiple-objective portfolio selection from the perspective of capital asset pricing models. *Mathematics*, 14(2):216, 2026. doi: 10.3390/math14020216.

Chen Qian, Wei Liu, Hongzhang Liu, et al. Chatdev: Communicative agents for software development. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 15174–15186, 2024.

Timo Schick, Jane Dwivedi-Yu, Roberto Dessì, et al. Toolformer: Language models can teach themselves to use tools. In *Advances in Neural Information Processing Systems*, 2023.

Noah Shinn, Federico Cassano, Ashwin Gopinath, Karthik Narasimhan, and Shunyu Yao. Reflexion: Language agents with verbal reinforcement learning. In *Proceedings of the 37th Conference on Neural Information Processing Systems (NeurIPS)*, volume 36 of *Advances in Neural Information Processing Systems*, pp. 8634–8652, 2023.

Justin Sirignano and Rama Cont. Universal features of price formation in financial markets: Perspectives from deep learning. *Quantitative Finance*, 19(9):1449–1459, 2019.

Peter Stone and Manuela Veloso. Multiagent systems: A survey from a machine learning perspective. *Autonomous Robots*, 8:345–383, 2000.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, 2017.

Alicia Vidler and Toby Walsh. Shifting power: Leveraging llms to simulate human aversion in abms of bilateral financial exchanges: A bond market study. In *Proceedings of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025)*, pp. 2777–2779, 2025. Extended abstract.

Lei Wang, Chen Ma, Xueyang Feng, et al. A survey on large language model based autonomous agents. *Frontiers of Computer Science*, 18(6):1–26, 2024. doi: 10.1007/s11704-024-40231-1.

Neng Wang et al. Fingpt: Instruction tuning benchmark for financial large language models. *arXiv preprint arXiv:2310.04793*, 2023a.

Xuezhi Wang, Jason Wei, Dale Schuurmans, et al. Self-consistency improves chain of thought reasoning in language models. In *International Conference on Learning Representations*, 2023b.

Michael Wooldridge and Nicholas R. Jennings. Intelligent agents: Theory and practice. *The Knowledge Engineering Review*, 10(2):115–152, 1995.

Qingyun Wu, Gagan Bansal, Jieyu Zhang, et al. Autogen: Enabling next-gen llm applications via multi-agent conversation. *arXiv preprint arXiv:2308.08155*, 2023a.

Shijie Wu, Ozan Irsoy, Steven Lu, et al. Bloomberggpt: A large language model for finance. *arXiv preprint arXiv:2303.17564*, 2023b.

Yijia Xiao, Edward Sun, Di Luo, et al. Tradingagents: Multi-agents llm financial trading framework. *arXiv preprint arXiv:2412.20138*, 2025. doi: 10.48550/arXiv.2412.20138.

Hongyang Yang, Xiao-Yang Liu, and Christina D. Wang. Fingpt: Open-source financial large language models. *arXiv preprint arXiv:2306.06031*, 2023.

Shunyu Yao, Jeffrey Zhao, Dian Yu, et al. React: Synergizing reasoning and acting in language models. In *International Conference on Learning Representations*, 2023.

Yuxuan Ye, Jun Han, Ao Hu, et al. The alpha illusion: Reported alpha from llm trading agents should not be treated as deployment evidence. *arXiv preprint arXiv:2605.16895*, 2026. doi: 10.48550/arXiv.2605.16895.

Yangyang Yu, Haohang Li, Zhi Chen, Yuechen Jiang, Yang Li, Denghui Zhang, Rong Liu, Jordan W. Suchow, and Khaldoun Khashanah. Finmem: A performance-enhanced llm trading agent with layered memory and character design. *arXiv preprint arXiv:2311.13743*, 2024a.

Yangyang Yu, Zhiyuan Yao, Haohang Li, Zhiyang Deng, Yuechen Jiang, Yupeng Cao, Zhi Chen, Jordan W. Suchow, Zhenyu Cui, Rong Liu, Zhaozhuo Xu, Denghui Zhang, Koduvayur Subbalakshmi, Guojun Xiong, Yueru He, Jimin Huang, Dong Li, and Qianqian Xie. Fincon: A synthesized llm multi-agent system with conceptual verbal reinforcement for enhanced financial decision making. *arXiv preprint arXiv:2407.06567*, 2024b.

Wentao Zhang, Lingxuan Zhao, Haochong Xia, et al. A multimodal foundation agent for financial trading: Tool-augmented, diversified, and generalist. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 4314–4325, 2024. doi: 10.1145/3637528.3671801.

Zhoufan Zhu and Ke Zhu. Alphaqcm: Alpha discovery in finance with distributional reinforcement learning. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*. PMLR, 2025.

AI USAGE DISCLOSURE

We acknowledge the use of OpenAI's Codex (model: GPT-5) for proofreading and grammar improvement.

APPENDIX

A EXPERIMENTAL SETUP AND EVALUATION METRICS

Table 5: Experimental setup for the PI-MAT evaluation.

Component	Setting
Asset and period	Daily AAPL observations; out-of-sample evaluation over 2020–2023.
Expert agents	Trend-following, mean-reversion, and event-driven proxy.
Comparators	Price-only (uses only price information), NoEvo (removes strategy evolution), NoImmune (removes the cognitive immune layer), SharedInput (provides the same information to all agents), and EqualWeight (assigns fixed equal capital weights).
Temporal checks	Consecutive early and late subperiods; 16 rolling evaluation windows.
Market regimes	Bull trend, bear trend, sideways, and high volatility.
Trading costs	Included in reported net performance.

Table 6: Definitions of the principal evaluation metrics.

Metric	Definition
Net return	$r_t = V_t/V_{t-1} - 1.$
Cumulative return	$R_{\mathrm{cum}} = V_T/V_0 - 1.$
Annualized return	$R_{\mathrm{ann}} = (V_T/V_0)^{K/T} - 1.$
Sharpe ratio	$\mathrm{SR} = \sqrt{K}\,(r_t - r_{f,t})/\,\mathrm{SD}(r_t - r_{f,t}).$
Maximum drawdown	$\mathrm{MDD} = \max_t \left[1 - V_t/\max_{u\le t} V_u\right].$
Signal diversity	$\overline{\lvert\rho\rvert}_{\mathrm{off}} = \frac{2}{N(N-1)} \sum_{i<j} \lvert\rho_{ij}\rvert.$
Near-zero allocation	$T^{-1} \sum_{t=1}^{T} \mathbf{1}\{w_{i,t} < 0.05\}.$
Recovery rate	Recovered suppression episodes / identified suppression episodes.
Window win rate	Windows won by PI-MAT / evaluated windows.
Paired return gain	$R_{\mathrm{ann}}^{\mathrm{PI\text{-}MAT}} - R_{\mathrm{ann}}^{\mathrm{baseline}}.$

B PSEUDOCODE FOR PARASITE-DRIVEN STRATEGY EVOLUTION

Algorithm 1 Parasite-Driven Strategy Evolution in PI-MAT

Require: Policy population $\mathcal{P} = \{\pi_i = (g_i, o_i, \theta_i)\}$; active ensemble $\mathcal{A}$; evaluation window $\mathcal{W}$; number of offspring K; smoothing rate ρ; selection temperature τ; mutation distribution $\mathcal{D}_{\text{mut}}$; feasible parameter set Θ; limits $E_{\max}, C_{\max}, s_{\min}$; thresholds $\gamma, \eta_{\max}$

Ensure: Updated population $\mathcal{P}^+$ and ensemble $\mathcal{A}^+$

for all $\pi_i = (g_i, o_i, \theta_i) \in \mathcal{P}$ **do**

  Estimate $\widehat{\mu}_i, \widehat{\sigma}_i, \widehat{\mathrm{CVaR}}_{\alpha,i}, \widehat{\mathrm{IC}}_i, \widehat{\mathrm{TO}}_i$ over $\mathcal{W}$

  Compute the risk-adjusted score

$$\mathcal{S}_i \leftarrow \widehat{\mu}_i - \lambda_\sigma \widehat{\sigma}_i - \lambda_{\mathrm{CVaR}} \widehat{\mathrm{CVaR}}_{\alpha,i} + \lambda_{\mathrm{IC}} \widehat{\mathrm{IC}}_i - \lambda_{\mathrm{TO}} \widehat{\mathrm{TO}}_i$$

  Update the persistent score

$$\overline{\mathcal{S}}_i \leftarrow (1-\rho)\overline{\mathcal{S}}_i + \rho \mathcal{S}_i$$

end for

Compute parent-selection probabilities

$$p_i^{\text{parent}} \leftarrow \frac{\exp(\overline{\mathcal{S}}_i/\tau)}{\sum_{\pi_k \in \mathcal{P}} \exp(\overline{\mathcal{S}}_k/\tau)}$$

Estimate population diversity $\mathcal{D}(\mathcal{P})$ and set

$$\delta \leftarrow \delta_{\min} + \kappa_\delta \big(1 - \mathcal{D}(\mathcal{P})\big)$$

for $k = 1, \ldots, K$ **do**

  Sample parent $\pi_{j_k} = (g_{j_k}, o_{j_k}, \theta_{j_k})$ according to p^{parent}

  Sample bounded perturbation

$$\epsilon_k \sim \mathcal{D}_{\text{mut}}(\sigma_{j_k}), \qquad \|\epsilon_k\|_W \le \delta$$

  Generate offspring by parasitic replication and mutation

$$\pi_k^{\text{new}} \leftarrow \mathcal{M}(\pi_{j_k}, \epsilon_k) = (g_{j_k}, o_{j_k}, \theta_{j_k} + \epsilon_k)$$

  Evaluate $\mathrm{Exposure}(\pi_k^{\text{new}})$, $\mathrm{CVaR}_\alpha(\pi_k^{\text{new}})$, and $\mathrm{Stability}(\pi_k^{\text{new}})$

  if $\theta_k^{\text{new}} \notin \Theta$ **or** $\mathrm{Exposure}(\pi_k^{\text{new}}) > E_{\max}$ **or** $\mathrm{CVaR}_\alpha(\pi_k^{\text{new}}) > C_{\max}$ **or** $\mathrm{Stability}(\pi_k^{\text{new}}) < s_{\min}$ **then**

    Reject π_k^{new}

    continue

  end if

  Compute marginal ensemble contribution

$$\Delta_k \leftarrow U(\mathcal{A} \cup \{\pi_k^{\text{new}}\}) - U(\mathcal{A})$$

  Compute $\mathrm{Redundancy}(\pi_k^{\text{new}}, \mathcal{A})$

  if $\Delta_k \ge \gamma$ **and** $\mathrm{Redundancy}(\pi_k^{\text{new}}, \mathcal{A}) \le \eta_{\max}$ **then**

    $\mathcal{A} \leftarrow \mathcal{A} \cup \{\pi_k^{\text{new}}\}$

    $\mathcal{P} \leftarrow \mathcal{P} \cup \{\pi_k^{\text{new}}\}$

  else

    Discard π_k^{new} as redundant

  end if

end for

return $\mathcal{P}^+ \leftarrow \mathcal{P}$, $\mathcal{A}^+ \leftarrow \mathcal{A}$

C Pseudocode for the Cognitive Immune Layer

Algorithm 2 Online Cognitive Immune Arbitration in PI-MAT

Require: Agent signals $\mathbf{s}_t = \{s_{i,t}\}_{i=1}^N$; confidence estimates $\mathbf{c}_t = \{c_{i,t}\}_{i=1}^N$; previous weights $\mathbf{w}_t$; previous statistics $\widehat{\mu}_{i,t-1}, \widehat{\sigma}^2_{i,t-1}, q_{i,t-1}$; asset returns R_{t+1}; execution cost $\mathcal{C}_{t+1}$; risk state $\mathcal{R}_t$

Ensure: Updated weights $\mathbf{w}_{t+1}$ and filtered signals $\widetilde{\mathbf{s}}_t$

Aggregate agent signals

$$S_t \leftarrow \sum_{i=1}^{N} w_{i,t} s_{i,t}$$

Observe realized portfolio performance

$$r_{t+1} \leftarrow S_t^\top R_{t+1} - \mathcal{C}_{t+1}$$

for $i = 1$ to N **do**

Estimate the realized agent contribution

$$\Delta r_{i,t+1} \leftarrow w_{i,t} s_{i,t}^\top R_{t+1} - \lambda_{\text{int}} |w_{i,t}| \|s_{i,t} - S_t\|_2$$

Update exponentially weighted return and risk statistics

$$\widehat{\mu}_{i,t} \leftarrow (1-\rho)\widehat{\mu}_{i,t-1} + \rho \Delta r_{i,t}$$

$$\widehat{\sigma}^2_{i,t} \leftarrow (1-\rho)\widehat{\sigma}^2_{i,t-1} + \rho(\Delta r_{i,t} - \widehat{\mu}_{i,t})^2$$

Construct the information state

$$x_{i,t} \leftarrow [\widehat{\mu}_{i,t}, -\widehat{\sigma}_{i,t}, c_{i,t}, \widehat{\Delta r}_{i,t}, \widehat{\text{Div}}_{i,t}]^\top$$

Compute the smoothed, risk-adjusted reliability state

$$q_{i,t} \leftarrow \gamma q_{i,t-1} + (1-\gamma) f_{\theta_t}(x_{i,t}) - \lambda_\sigma \widehat{\sigma}_{i,t} - \lambda_d \widehat{\text{MDD}}_{i,t}$$

Apply reversible immune filtering

$$g_{i,t} \leftarrow \sigma(\kappa(q_{i,t} - \tau_t)), \qquad \widetilde{s}_{i,t} \leftarrow g_{i,t} s_{i,t}$$

end for

Compute performance-responsive weights

for $i = 1$ to N **do**

$$w^{\text{raw}}_{i,t+1} \leftarrow \frac{\exp(\beta q_{i,t})}{\sum_{j=1}^N \exp(\beta q_{j,t})}$$

end for

Apply risk-aware arbitration

$$\mathbf{w}_{t+1} \leftarrow \mathcal{A}_{\theta_t}(\mathbf{w}_t, \mathbf{q}_t, \mathbf{s}_t, \mathcal{R}_t)$$

Return $\mathbf{w}_{t+1}$ and $\widetilde{\mathbf{s}}_t$

D NOTATION AND VARIABLE DEFINITIONS

Table 7: Notation and variable definitions in PI-MAT.

Symbol	Category	Definition
t	Time index	Decision time or evaluation period.
i, j, k	Agent indices	Indices of trading agents, candidate policies, or parent strategies.
N	System dimension	Number of active trading agents.
$\mathcal{O}_t$	Market observation	Current market information, including price, fundamental, and event data.
$\mathbf{x}_t^p, \mathbf{x}_t^f, \mathbf{x}_t^e$	Information modalities	Price, fundamental, and event-related observations at time t.
$\mathcal{H}_t$	Historical state	Historical market observations, trading records, and execution states.
$\mathcal{I}_i$	Information interface	Agent-specific transformation that extracts the information available to agent i.
$\mathbf{z}_{i,t}$	Agent representation	Agent-specific representation generated from current observations and historical information.
$\mathcal{P}_i$	Information perspective	Dedicated information perspective assigned to agent i.
$\mathcal{G}_i$	Trading objective	Strategic objective of agent i, such as trend following, mean reversion, or event-driven alpha generation.
π_{θ_i}	Signal policy	Parameterized policy that maps agent representations to trading signals.
$s_{i,t}$	Agent signal	Trading signal generated by agent i at time t.
$c_{i,t}$	Signal confidence	Confidence estimate associated with agent i's signal.
Γ_{ϕ_i}	Confidence model	Parameterized function that estimates signal confidence.
$\mathbf{s}_t, \mathbf{c}_t$	Agent outputs	Vectors collecting all agent signals and confidence estimates.
$\mathbf{h}_t$	Trading-graph state	State of the stateful trading graph used to coordinate agent outputs.
$\mathcal{F}_\psi$	State-transition operator	Operator that updates the trading-graph state using agent outputs and market information.
π_i	Trading policy	Policy representation of agent i, including its objective, information perspective, and parameters.
g_i	Trading objective label	Objective associated with policy π_i.
o_i	Information perspective label	Information perspective associated with policy π_i.
θ_i	Policy parameters	Decision parameters of policy π_i.
$\alpha_i, \beta_i, \gamma_i$	Policy coefficients	Strategy-specific coefficients controlling signal weighting, risk preference, or reasoning behavior.
θ_i^{new}	Mutated parameters	Parameters obtained after applying controlled mutation to a parent strategy.
ϵ_i	Mutation vector	Perturbation applied to policy parameters during strategy evolution.
$\mathcal{D}_{\text{mut}}$	Mutation distribution	Distribution governing controlled parameter perturbations.
δ_i	Mutation radius	Maximum scale of the mutation applied to policy i.

Continued on next page

Table 7 – continued from previous page

Symbol	Category	Definition
$\mathcal{D}(\mathcal{P})$	Population diversity	Measure of strategic diversity within the current policy population.
$\mathcal{S}_i$	Strategy score	Risk-adjusted performance score used to evaluate policy i.
$\overline{\mathcal{S}}_i$	Smoothed strategy score	Temporally smoothed version of the strategy score.
$\widehat{\mu}_i$	Mean return estimate	Estimated average return of policy i.
$\widehat{\sigma}_i$	Volatility estimate	Estimated return volatility of policy i.
$\widehat{\mathrm{CVaR}}_{\alpha,i}$	Tail-risk estimate	Conditional value-at-risk of policy i at confidence level α.
$\widehat{\mathrm{IC}}_i$	Information coefficient	Correlation between predicted and realized returns for policy i.
$\widehat{\mathrm{TO}}_i$	Turnover estimate	Estimated portfolio turnover associated with policy i.
p_i^{parent}	Parent-selection probability	Probability that policy i is selected as a parent during strategy evolution.
τ	Selection temperature	Parameter controlling exploration versus exploitation in parent selection.
$\mathcal{M}$	Mutation operator	Operator that generates a new policy from a selected parent and mutation vector.
$\mathbb{I}_{\mathrm{valid}}$	Validity indicator	Indicator specifying whether a candidate policy satisfies feasibility and risk constraints.
$\mathcal{A}$	Active ensemble	Set of currently admitted trading policies.
$U(\mathcal{A})$	Ensemble utility	Utility of an active ensemble, incorporating return, risk, cost, and diversity.
Δ_i	Marginal contribution	Incremental utility obtained by adding policy i to the active ensemble.
$w_{i,t}$	Agent capital weight	Capital allocation weight assigned to agent i at time t.
$\mathbf{w}_t$	Weight vector	Vector of all agent capital weights at time t.
S_t	Aggregate signal	Weighted combination of agent signals used for portfolio construction.
$\Delta r_{i,t+1}$	Agent contribution	Realized contribution of agent i to the next-period portfolio outcome.
r_{t+1}	Portfolio return	Realized portfolio return after execution at time $t+1$.
R_{t+1}	Asset-return vector	Vector of realized asset returns at time $t+1$.
$\mathcal{C}_{t+1}$	Execution cost	Transaction costs, slippage, market impact, and other execution losses.
λ_{int}	Redundancy penalty	Coefficient penalizing highly redundant or insufficiently complementary signals.
$\widehat{\Delta r}_{i,t}$	Contribution estimate	Estimated marginal contribution of agent i to portfolio performance.
$\widehat{\mathrm{Div}}_{i,t}$	Diversification estimate	Estimated diversification or complementarity benefit of agent i.
$q_{i,t}^{\mathrm{raw}}$	Raw reliability score	Instantaneous reliability estimate produced by the online meta-learner.
$q_{i,t}$	Reliability state	Smoothed and risk-adjusted reliability state used for signal filtering and capital allocation.
$\widehat{\mathrm{MDD}}_{i,t}$	Maximum drawdown estimate	Estimated maximum cumulative loss of agent i over the recent evaluation window.
γ	Smoothing coefficient	Coefficient controlling temporal smoothing of reliability estimates.

Continued on next page

Table 7 – continued from previous page

Symbol	Category	Definition
ρ	Adaptation rate	Update rate for exponentially weighted return and risk statistics.
$\lambda_\sigma, \lambda_d$	Risk-penalty coefficients	Coefficients penalizing volatility and maximum drawdown in reliability estimation.
$g_{i,t}$	Immune gate	Soft reliability gate controlling the effective influence of agent i's signal.
$\widetilde{s}_{i,t}$	Filtered signal	Signal remaining after reliability-aware immune filtering.
τ_t	Reliability threshold	Dynamic threshold separating reliable and unreliable signal states.
κ	Gate sharpness	Parameter controlling the sharpness of the immune-filtering transition.
$e_{i,t}$	Suppression state	Exponentially smoothed evidence of persistent signal unreliability.
η	Evidence decay rate	Coefficient controlling the persistence of the suppression state.
β	Allocation sensitivity	Temperature parameter controlling the sensitivity of capital weights to reliability differences.
$\mathcal{R}_t$	Current risk state	Portfolio-level state summarizing volatility, leverage, concentration, turnover, and their admissible limits.
$\mathcal{A}_{\theta_t}$	Arbitration operator	Online operator mapping current weights, reliability states, signals, and risk state to updated weights.
$\mathbf{p}_t$	Target portfolio	Asset-level portfolio position obtained from risk-aware portfolio optimization.
$\mathbf{\Sigma}_t$	Covariance matrix	Estimated covariance matrix of asset returns at time t.
λ_{risk}	Risk-aversion coefficient	Penalty imposed on portfolio variance in the portfolio objective.
λ_{tc}	Turnover penalty	Penalty imposed on changes in portfolio positions.
$\mathbf{Y}_{t+1}$	Realized feedback	Observed returns, execution outcomes, and risk metrics used for online updates.
$\mathcal{U}$	Update operator	Operator updating agent weights and strategy parameters from realized feedback.