Market Regime Council for Dynamic Credit Assignment in Multi-Agent LLM Decision Systems

Yunhua Pei
University of Bristol, UK
ge22472@bristol.ac.uk

Zerui Ge
Independent Researcher
gezerui1997@gmail.com

Jin Zheng
University of Bristol, UK
jin.zheng@bristol.ac.uk

John Cartlidge
University of Bristol, UK
john.cartlidge@bristol.ac.uk

Abstract

Multi-agent LLM decision systems for portfolio management still lack a principled way to assign credit across specialist agents, remain vulnerable to cold-start dominance under regime shifts, and offer limited transparency into how final allocations are formed. We propose Market Regime Council (MRC), an online credit-assignment mechanism over coalitions of specialist agents, with portfolio management as the downstream task. MRC computes exact Shapley credits across all single, pairwise, and Grand-coalition outputs for online agent weighting. Instantiated with $N{=}3$ specialist agents, at each trading period, MRC recomputes coalition-based Shapley weights from exponentially weighted performance histories, uses a Bayesian adaptive mixture to stabilize early periods, applies regime-dependent multipliers to adjust agent authority, and records each rebalance through a five-layer causal trace. Over 1,037 trading days across 13 crypto assets and five seeds, MRC achieves a Sharpe ratio of 1.51 and a cumulative return of 440.1%, ranking first on CR, SR, and IR among active baselines and attaining the lowest MDD among active methods. Ablation results show that the gains come from Shapley-weighted integration across coalition outputs rather than from any single stage in isolation. Code and demo data are available open source.[1]

1 Introduction

Cryptocurrency portfolio management is a sequential decision problem that depends on heterogeneous signals with different modalities, update frequencies, and predictive horizons. Price trends, on-chain activity, macro conditions, and market sentiment can all matter, but their usefulness changes markedly across bull, volatile, and bear markets. In practice, strong portfolio decisions are usually formed by combining partial views from different sources rather than by relying on any single signal [31].

LLM-based multi-agent systems are a natural fit for this setting [14, 12, 19], and recent financial LLM systems have reported strong risk-adjusted performance [53, 56, 25, 45]. However, three problems remain. First, most existing systems use fixed rules or heuristic confidence scores to weight agents, which provides little meaningful credit assignment, that is, little ability to identify which agent or coalition actually improved realized performance [53, 46, 54, 25]. Second, naive online adaptation is vulnerable to cold-start dominance and can retain stale weights after a regime shift [16, 13, 11]. Third, the final allocation is usually hard to trace back to its intermediate decisions [2, 7, 26]. Section 2 compares how existing systems assign credit.

[1] https://github.com/PEIYUNHUA/Market_Regime_Council

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

Table 1: How the compared trading systems assign credit across agents.

System	Credit across agents	Coalitions
FinCon [53]	Verbal reinforcement from one manager agent	no
TradingAgents [46]	Fixed role pipeline	no
FinMem [54]	Memory recency, no credit between agents	no
CryptoTrade [18]	Reflection within one agent	no
FSReason [42]	Equal-weight fusion of two agents	no
MAPPO [52], MADDPG [23]	Centralised critic	no
MRC (Ours)	Exact Shapley values over all 2^N-1 coalitions	yes

Market Regime Council (MRC) treats multi-agent portfolio management as an online cooperative game, using Shapley values as the credit-assignment signal [34]. At each period, MRC evaluates all 2^N-1 coalition outputs, computes exact Shapley credits, and converts them into agent weights. With $N=3$ specialist agents covering three orthogonal financial information channels, regime-dependent multipliers adapt these weights to non-stationary conditions, and a five-layer causal trace renders each rebalancing decision fully auditable. Our contributions can be summarized as:

- **Online Shapley credit assignment.** We formulate multi-agent cooperation as a transferable-utility game over all 2^N-1 coalitions, compute exact Shapley values as online credit signals, and combine them with a Bayesian mixture that transitions from a uniform prior to Shapley-derived weights as evidence accumulates.
- **Regime-aware adaptation.** We combine Shapley weighting with a continuous regime score that adjusts agent authority and drives the downstream risk-control overlays across bull, volatile, and bear conditions.
- **Full-chain explainability.** We trace each rebalance through five layers, from raw inputs to agent portfolios, Shapley credit, blend ratios, overlay effects, and the final allocation.
- **Empirical validation.** Over 1,037 trading days across 13 assets and five seeds, MRC ranks first on CR, SR, and IR among active LLM and DRL baselines, and achieves the lowest MDD among active methods, instantiated at $N=3$.

2 Related Work

None of the other systems in Table 1 measures how much each agent or coalition adds to realized performance. MRC instead assigns coalition-based credit grounded in the Shapley axioms [34]. The agent weights it applies are not a raw Shapley allocation but this credit mixed with a uniform prior and then adjusted by the Regime-Aware Multiplier and the Selective Winner-Takes-All override (Section 4). Appendix A gives the full discussion.

3 Preliminaries

Problem Formulation. Let $\mathcal{K}$ be the set of K crypto assets, $\mathcal{N} = \{1, \ldots, N\}$ a set of N specialist agents, and $\mathcal{T} = \{1, \ldots, T\}$ the sequence of decision periods. At each period t, every non-empty coalition $S \subseteq \mathcal{N}$ can produce a joint decision, and realized feedback is recorded in history $\mathcal{H}_t$. In this work we instantiate $N=3$, where $\mathcal{N} = \{1, 2, 3\}$ indexes three orthogonal financial information channels: price/technical (A_1), on-chain network activity (A_2), and macro/sentiment (A_3). The system receives heterogeneous data streams organized into three agent-specific feature bundles $\mathbf{x}_1^{(t)}, \mathbf{x}_2^{(t)}, \mathbf{x}_3^{(t)}$ spanning six modality groups (full inventory in Appendix D, Table 5). The system must output a portfolio weight vector $\mathbf{w}^{(t)} \in \mathcal{W}$, where $\mathcal{W} = \{\mathbf{w} : \mathbf{1}^\top \mathbf{w} + c = 1,\ w_k \in [0, w_{\max}],\ c \in [0, c_{\max}]\}$ and c denotes the cash allocation, enforcing per-asset concentration limits and a bounded cash position. Let $\mathbf{r}^{(t)}$ denote the asset return vector at period t; the objective is to maximise the long-run risk-adjusted return under the non-stationarity imposed by shifting market regimes [11]. Let $\mathcal{P} = \{(1,2), (1,3), (2,3)\}$ denote the set of pairwise coalitions.

Cooperative Game System. We model multi-agent credit assignment as a transferable-utility cooperative game $(\mathcal{N}, v)$ [34, 40, 38, 39], where the characteristic function $v : 2^{\mathcal{N}} \to \mathbb{R}$ maps

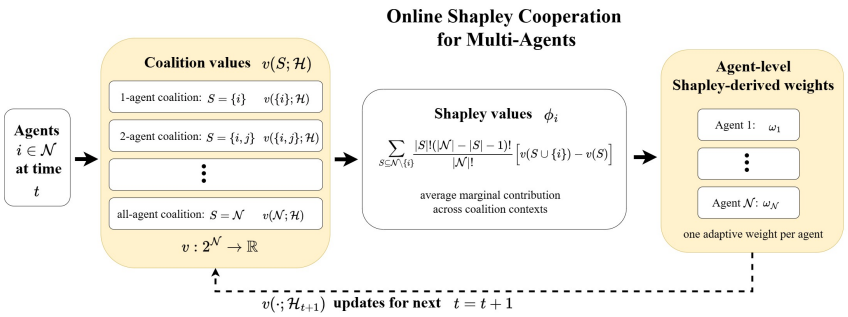

Figure 1: The general N-agent Online Shapley Cooperation mechanism.

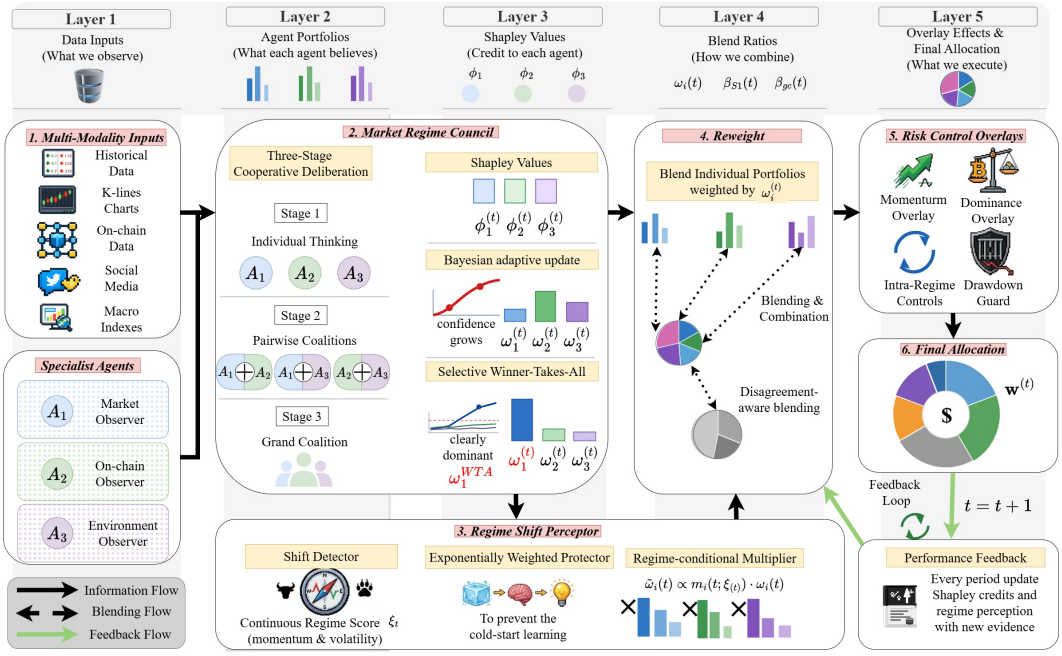

Figure 2: The MRC mechanism applied with N=3 agents for crypto portfolio management.

each non-empty coalition $S \subseteq \mathcal{N}$ to a scalar utility measuring its realized performance, with the convention $v(\emptyset) = 0$. The Shapley value ϕ_i of agent i is the unique credit allocation satisfying four-axiom verification (Efficiency, Symmetry, Dummy Player, Additivity) [34] (see Appendix G):

$$\phi_i = \sum_{S\subseteq\mathcal{N}\setminus\{i\}} \frac{|S|!\,(|\mathcal{N}|-|S|-1)!}{|\mathcal{N}|!}\big[v(S\cup\{i\})-v(S)\big]. \tag{1}$$

Equation (1) averages agent i's marginal contribution across all coalitions that exclude i, weighted by the probability of the corresponding arrival ordering. For N agents, evaluating all 2^N-1 coalitions requires $O(N\cdot 2^N)$ utility calls per period; at N=3 this reduces to exactly 7 evaluations, with closed-form expressions in Appendix G. What grows fast with N is the number of coalition portfolios, not the Shapley arithmetic, whose cost and error are reported in Table 19 (Appendix I). Figure 1 illustrates how these credits are computed online and converted to adaptive weights each period. Section 4 specifies how MRC instantiates this mechanism for financial portfolio management.

4 Method

Figure 2 summarizes MRC. The key design questions from Section 3 are: how to define the characteristic function v for financial coalition outputs, how to stabilize Shapley estimates during early periods, and how to adapt agent authority when market regimes shift. MRC addresses each of these

in turn. At each trading period, three specialist agents produce individual portfolios, pairwise debates yield coalition portfolios, and a grand-coalition readout aggregates the shared context into a synthesis portfolio. MRC then computes coalition utilities from realized historical performance and updates agent weights online via Shapley-based credit assignment. Complete pseudocode for the online weight update and per-period inference, see Appendix C.

4.1 Market Regime Council

4.1.1 Three-Stage Cooperative Deliberation (N=3 Instantiation)

MRC organizes agent interaction as a sequence of broader coalitions. In Stage 1, three specialized LLM agents $\mathcal{A} = \{A_1, A_2, A_3\}$ independently process their dedicated data modalities and each produce a private portfolio $\mathbf{w}_i^{(t)}$ and regime label $\hat{r}_i^{(t)} \in \{\text{bull, volatile, bear}\}$. In Stage 2, each pair (A_i, A_j) engages in a structured cross-examination producing a joint portfolio $\mathbf{w}_{ij}^{(t)}$, and three debates run in parallel: (A_1, A_2), (A_1, A_3), (A_2, A_3). In Stage 3, the grand coalition $\mathcal{N} = \{1, 2, 3\}$ portfolio $\mathbf{w}_{123}^{(t)}$ is the outcome of all three specialists jointly deliberating with A_4 as the readout operator, guided by a Shapley report summarizing agent weights, coalition-level Sharpe ratios, and recent performance (format in Appendix D). A_4 is a readout and is not a player in $\mathcal{N}$. For each non-empty coalition $S \subseteq \mathcal{N}$, the coalition portfolio is $\mathbf{w}_i^{(t)}$ ($|S| = 1$), $\mathbf{w}_{ij}^{(t)}$ ($|S| = 2$), or $\mathbf{w}_{123}^{(t)}$ ($|S| = 3$). The characteristic function is the composite performance measure (weights γ_ρ and γ_μ, see Appendix D)

$$v(S;\, \mathcal{H}_t) = \gamma_\rho \cdot \sqrt{365}\, \frac{\hat{\mu}_{\mathrm{EW}}(S;\, \mathcal{H}_t)}{\hat{\sigma}_{\mathrm{EW}}(S;\, \mathcal{H}_t)} + \gamma_\mu \cdot \mu_{\mathrm{ann}}(S;\, \mathcal{H}_t), \quad \gamma_\rho + \gamma_\mu = 1, \tag{2}$$

where $\mathcal{H}_t$ denotes the realized return history of coalition S up to time t, $\hat{\mu}_{\mathrm{EW}}(S; \mathcal{H}_t)$ and $\hat{\sigma}_{\mathrm{EW}}(S; \mathcal{H}_t)$ are the EWP-estimated daily mean and daily standard deviation of excess returns, and $\mu_{\mathrm{ann}}(S; \mathcal{H}_t) = 365\, \hat{\mu}_{\mathrm{EW}}(S; \mathcal{H}_t)$ is the annualised mean. This mean is scaled to the same order of magnitude as the Sharpe term to prevent one component from dominating the blend. EWP estimation is defined in §4.2.2. Equal-weight history averaging creates path dependence because early cold-start data permanently biases estimates, whereas the EWP (Appendix F) exponentially attenuates this influence.

4.1.2 Bayesian Adaptive Update.

Applying Eq. (1) to the characteristic function values yields Shapley values ϕ_i; agents with $\phi_i < 0$ receive zero weight via $\phi_i^+ = \max(\phi_i, 0)$ [39, 3], a projection required by the long-only portfolio constraint. Closed-form three-player expressions and four-axiom verification are in Appendix G. Shapley estimates from a few observations are unreliable, so we mix the truncated Shapley-derived allocation with the uninformative uniform prior [9]:

$$\alpha(t) = 1 - \exp(-t/\lambda)\,, \tag{3}$$

$$\omega_i^{(t)} = \alpha(t) \cdot \frac{\phi_i^+}{\sum_j \phi_j^+} + \big(1 - \alpha(t)\big) \cdot \tfrac{1}{|\mathcal{N}|} \tag{4}$$

where t is the number of completed periods and λ governs cold-start duration. As $t \to \infty$, $\alpha \to 1$ and the weights converge to the normalized non-negative Shapley-derived allocation. The update follows a three-phase schedule: $\alpha \approx 0$ corresponds to the uniform-prior regime, $0 < \alpha < 1$ to Bayesian mixing, and $\alpha \approx 1$ to a fully data-driven allocation. If every ϕ_i^+ is zero, the Shapley share in Eq. (4) is replaced by the prior, so $\omega_i^{(t)}$=$1/|\mathcal{N}|$ before the regime multipliers.

4.1.3 Selective Winner-Takes-All.

When one agent's rolling-window Sharpe exceeds the mean of the others by factor θ_{WTA}, the Bayesian mixture is temporarily replaced by:

$$\omega_i^{\mathrm{WTA}} = \begin{cases} \omega_{\mathrm{wta}} & i = i^* \\ (1 - \omega_{\mathrm{wta}}) \cdot \dfrac{\omega_i}{\sum_{j \neq i^*} \omega_j} & i \neq i^* \end{cases} \tag{5}$$

where ω_i^{WTA} is the WTA-overridden weight that temporarily replaces the Bayesian mixture weight ω_i, while ω_{wta} is the fixed share allocated to the dominant agent. $i^* = \arg\max_i \rho_i^{(n_{\mathrm{win}})}$ is the dominant agent, and $\rho_i^{(n_{\mathrm{win}})}$ is the rolling Sharpe of agent i over the most recent periods n_{win}. θ_{WTA} is the dominance threshold. The override reverts automatically once dominance ends. It never fires when the mean rolling Sharpe of the other agents is not positive, so an undefined or negative reference keeps the Bayesian mixture.

4.1.4 Multi-Stage Portfolio Integration

Both blend ratios are driven continuously by performance gaps through $\tanh$:

$$\beta_{S1}^{(t)} = \bar{\beta}_{S1} + \delta_{S1} \cdot \tanh\Big(-\Delta_{S2}^{(t)}/\tau_{S1}\Big), \tag{6}$$

$$\beta_{\mathrm{gc}}^{(t)} = \max\Big(0,\ \bar{\beta}_{\mathrm{gc}} + \delta_{\mathrm{gc}} \cdot \tanh\Big(\Delta_{\mathrm{gc}}^{(t)}/\tau_{\mathrm{gc}}\Big)\Big), \tag{7}$$

where $\Delta_{S2}^{(t)} = v_{\mathrm{ens}}^{S2} - v_{\mathrm{ens}}^{S1}$ and $\Delta_{\mathrm{gc}}^{(t)} = v(\mathcal{N}; \mathcal{H}_t) - v_{\mathrm{ens}}^{S1}$ measure the performance advantages of debate and grand coalition over Stage 1. $\bar{\beta}_{S1}$, δ_{S1}, τ_{S1} and their gc counterparts are the center, scale, and bandwidth of the respective blend functions, set to $\bar{\beta}_{S1}$=0.90, δ_{S1}=0.09, τ_{S1}=0.075 and $\bar{\beta}_{\mathrm{gc}}$=0.15, δ_{gc}=0.20, τ_{gc}=0.10. And

$$v_{\mathrm{ens}}^{S1} = \sum_{i \in \mathcal{N}} \tilde{\omega}_i^{(t)} \cdot v(\{i\}; \mathcal{H}_t), \qquad v_{\mathrm{ens}}^{S2} = \sum_{(i,j) \in \mathcal{P}} p_{ij}^{(t)} \cdot v(\{i,j\}; \mathcal{H}_t) \tag{8}$$

are the Shapley-weight-averaged characteristic function values (Eq. (2)) of the Stage 1 individual and Stage 2 pairwise coalitions respectively. When debate underperforms ($\Delta_{S2}^{(t)} < 0$), $\beta_{S1}^{(t)}$ automatically rises. The Shapley-weighted plurality score is introduced as

$$\kappa^{(t)} = \max_{r \in \{\text{bull, volatile, bear}\}} \sum_{i : \hat{r}_i^{(t)} = r} \omega_i^{(t)}, \tag{9}$$

to further discounts the grand-coalition blend when agents disagree on regime, consistent with the expert-disagreement principle in information economics [29]:

$$\beta_{\mathrm{gc}}^{\mathrm{final}} = \beta_{\mathrm{gc}}^{(t)} \cdot \left(\tfrac{1}{2} + \tfrac{1}{2}\kappa^{(t)}\right). \tag{10}$$

Given these blend ratios, the final Council output is:

$$\mathbf{w}_{\mathrm{council}}^{(t)} = \beta_{\mathrm{gc}}^{\mathrm{final}}\, \mathbf{w}_{123}^{(t)} + \big(1 - \beta_{\mathrm{gc}}^{\mathrm{final}}\big)\Big[\beta_{S1}^{(t)} \sum_i \tilde{\omega}_i^{(t)}\, \mathbf{w}_i^{(t)} + (1 - \beta_{S1}^{(t)}) \sum_{(i,j)} p_{ij}^{(t)}\, \mathbf{w}_{ij}^{(t)}\Big], \tag{11}$$

where $\tilde{\omega}_i^{(t)}$ are the regime-adjusted agent weights (Eq. (15), §4.2). The pairwise weights $p_{ij}^{(t)}$ are the Bayesian-normalized weights over the three elements of $\mathcal{P} = \{(1,2),(1,3),(2,3)\}$, obtained by applying the same adaptive update as Eq. (4) to the pairwise coalition values.

4.2 Regime Shift Perceptor

4.2.1 Shift Detector

To detect the current market regime, we define the regime score $\xi^{(t)} \in (-1, +1)$ as a momentum-to-volatility ratio:

$$\xi^{(t)} = \tanh\left(\frac{r_{30d}^{(t)}}{\sigma_{30d}^{(t)}}\right) \tag{12}$$

where $r_{30d}^{(t)}$ is the equal-weight 30-day log return and $\sigma_{30d}^{(t)}$ is its realized daily standard deviation. A short-term directional conflict check attenuates $\xi^{(t)}$ when recent momentum contradicts the longer-term trend, reflecting directional ambiguity. Positive values indicate bull conditions and negative values indicate bear conditions. Three nominal regime labels are derived from the thresholds $\xi_+ > 0$ and $\xi_- < 0$, which define the bull and bear boundaries respectively (values in Appendix D). The resulting regime segmentation is also visible in the bottom panel of Figure 3.

4.2.2 Exponentially Weighted Protector

During the initial periods of deployment, Shapley-based agent weights have high estimation variance due to the limited return history, which is known as the cold-start. To address path dependence, we introduce the Exponentially Weighted Protector (EWP), which assigns decaying weights geometrically to historical returns when estimating the coalition characteristic functions in Eq. (2):

$$w_\tau = e^{-(t-\tau)/h}, \quad \tau = 1, \ldots, t \tag{13}$$

$$\hat{\mu}_{\mathrm{EW}}(S; \mathcal{H}_t) = \frac{\sum_{\tau=1}^{t} w_\tau \, R_S^{(\tau)}}{\sum_{\tau=1}^{t} w_\tau}, \qquad \hat{\sigma}^2_{\mathrm{EW}}(S; \mathcal{H}_t) = \frac{\sum_{\tau=1}^{t} w_\tau \big(R_S^{(\tau)} - \hat{\mu}_{\mathrm{EW}}(S; \mathcal{H}_t)\big)^2}{\sum_{\tau=1}^{t} w_\tau} \tag{14}$$

where $\mathcal{H}_t$ is the realized return history of coalition S up to period t, h=252 is the e-fold decay period, and $R_S^{(\tau)}$ is the realized return of coalition S at period τ. EWP is grounded in discounted online learning [57, 13] and EWMA [22]. [41] further connects its geometric structure to Shapley's Additivity axiom (formal proof in Appendix F, Proposition 1, and verification in Appendix G).

4.2.3 Regime-Aware Multiplier

The Bayesian Shapley weights reflect the average historical contribution across all market regimes. In practice, agent contributions are regime-dependent: technical signals dominate during trending markets, on-chain signals provide the clearest per-token divergence in bear markets, while macroeconomic signals carry the highest cross-sectional authority in sustained bear regimes due to their role at regime turning points. To prevent any single agent from accumulating disproportionate weight during a favorable regime, we apply a Regime-Aware Multiplier: each Shapley weight is scaled by a continuous, agent-specific scalar obtained by monotone linear interpolation of agent i's bull, volatile, and bear multiplier values at $\xi^{(t)}$ (values in Appendix D), and the result is renormalized over $i \in \mathcal{N}$:

$$\tilde{\omega}_i^{(t)} \;\propto\; \omega_i^{(t)} \cdot \psi_i\Big(\xi^{(t)}\Big), \qquad \sum_{i \in \mathcal{N}} \tilde{\omega}_i^{(t)} = 1, \tag{15}$$

where ψ_i is the regime-conditional scalar for agent i. An additional dominant-dimension override amplifies the leading agent's multiplier when the Shapley-weighted plurality vote exhibits strong consensus, with intensity proportional to the normalized consensus strength $\dfrac{\kappa^{(t)} - 1/|\mathcal{N}|}{1 - 1/|\mathcal{N}|} \in [0, 1]$ (for N=3: $(\kappa^{(t)} - \frac{1}{3})/\frac{2}{3}$), ensuring that a uniform vote ($\kappa^{(t)} = 1/|\mathcal{N}|$) produces no amplification.

4.3 Risk Control

Momentum Overlay. Each council weight is multiplicatively scaled by a regime-gated cross-sectional momentum score, with stronger tilts in trending regimes and near-zero tilts in volatile regimes. Full specification is in Appendix E. **Dominance Overlay.** A BTC dominance signal $d^{(t)}$ captures the relative outperformance of BTC versus the equal-weight basket over a rolling window. The signal continuously tilts the portfolio toward BTC (or altcoins) via a $\tanh$-transformed score:

$$d^{(t)} = \tanh\left(\frac{\Delta^{(t)}_{\mathrm{BTC-EW}}}{\tau_d}\right) \in (-1, 1) \tag{16}$$

In volatile regimes, a hard BTC floor constraint additionally enforces a minimum BTC allocation, motivated by the empirical finding that BTC significantly outperforms the altcoin equal-weight basket during high-volatility periods. **Intra-Regime Controls.** Three regime-specific controls further refine the portfolio. A bear on-chain tilt shifts weight toward BTC based on network-activity signals. A volatile cash target rises continuously as the regime score approaches zero. A bull-to-volatile transition buffer scales down positions when the regime score drops sharply, see Appendix E. **Drawdown Protection.** Realized drawdown $\mathrm{DD}^{(t)}$ from the rolling portfolio peak triggers a regime-gated position reduction, analogous to the risk-sensitive value factorization approach of [36]:

$$s_{\mathrm{dd}}^{(t)} = 1 - g\Big(\xi^{(t)}\Big) \cdot \tanh\left(\frac{\mathrm{DD}^{(t)}}{\tau_{\mathrm{dd}}}\right) \cdot \nu_{\mathrm{dd}} \tag{17}$$

Table 2: Main results over 1,037 trading days (2023-03-01 to 2025-12-31). Non-benchmark methods: mean $\pm$ std (5 seeds); benchmarks: deterministic. Red: best mean per column among non-benchmark methods; Blue: second best. Superscripts on MRC denote two-sample t-test significance versus the second-best method per metric: $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$ (Welch's t-test, $n=5$ seeds each). The transaction cost is 0bps. For more cost sensitivity analysis, please refer to Table 9 (Appendix H).

Category	Method	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)
Benchmarks	EW	279.93	1.14	42.78	0.00
	BTC	270.94	1.23	32.02	−0.07
	ETH	78.45	0.63	63.75	−0.59
DRL Baselines	A2C	$196.0_{\pm 64.4}$	$1.03_{\pm 0.18}$	$38.5_{\pm 8.2}$	$-0.53_{\pm 0.31}$
	PPO	$215.2_{\pm 58.5}$	$1.04_{\pm 0.13}$	$41.1_{\pm 3.6}$	$-0.51_{\pm 0.41}$
	DQN	$249.7_{\pm 76.0}$	$1.16_{\pm 0.18}$	$34.6_{\pm 4.4}$	$-0.45_{\pm 0.55}$
	MAPPO	$254.5_{\pm 4.5}$	$1.14_{\pm 0.01}$	$39.9_{\pm 0.4}$	$-1.14_{\pm 0.11}$
	MADDPG	$251.5_{\pm 13.3}$	$1.14_{\pm 0.03}$	$40.2_{\pm 0.6}$	$-1.16_{\pm 0.34}$
LLM Baselines	FinCon	$262.1_{\pm 6.4}$	$1.14_{\pm 0.02}$	$41.2_{\pm 1.2}$	$-0.98_{\pm 0.19}$
	TradingAgents	$277.7_{\pm 15.7}$	$1.19_{\pm 0.03}$	$38.6_{\pm 2.9}$	$-0.23_{\pm 0.15}$
	FinMem	$320.3_{\pm 107.6}$	$1.25_{\pm 0.20}$	$39.9_{\pm 1.9}$	$0.04_{\pm 0.50}$
	CryptoTrade	$356.3_{\pm 39.9}$	$1.32_{\pm 0.06}$	$41.3_{\pm 3.6}$	$0.33_{\pm 0.23}$
	FSReason	$282.6_{\pm 29.8}$	$1.20_{\pm 0.06}$	$38.4_{\pm 1.7}$	$-0.14_{\pm 0.18}$
Full System	**MRC (Ours)**	$440.1^{**}_{\pm 51.4}$	$1.51^{***}_{\pm 0.07}$	$34.1_{\pm 2.5}$	$0.47_{\pm 0.20}$

$$\mathbf{w}^{(t)} = s_{\mathrm{dd}}^{(t)} \cdot \mathbf{w}^{(\ell)}, \qquad c^{(t)} = c^{(\ell)} + \left(1 - s_{\mathrm{dd}}^{(t)}\right) \mathbf{1}^{\top} \mathbf{w}^{(\ell)} \tag{18}$$

where ℓ indexes the earlier overlay step, $g(\xi^{(t)}) = \max(0, -\xi^{(t)})$ is a regime gate suppressing protection in bull regimes, $\tau_{\mathrm{dd}}=0.15$ is the drawdown bandwidth, and $\nu_{\mathrm{dd}}=0.40$ is the maximum position-reduction coefficient. After all overlays, $\mathbf{w}^{(t)}$ is projected onto $\mathcal{W}$, and the realized return $R^{(t)} = (\mathbf{w}^{(t)})^{\top} \mathbf{r}^{(t+1)}$ updates $v(\cdot; \mathcal{H}_{t+1})$, which closes the online learning loop.

5 Experiments

5.1 Settings

Dataset. We evaluate MRC on a multi-modal web3 dataset covering 13 assets from 2023-03-01 to 2025-12-31, yielding 1,037 daily decision periods. On day t every agent sees only rows dated t or earlier (per-source timing in Appendix D, Table 5). We use crypto as a stress-test domain for non-stationary multi-agent credit assignment because it runs 24/7, transparent on-chain activity, and rapid regime shifts. All competing LLM and DRL baselines are evaluated on comparably domain-specific asset universes without cross-asset transfer claims [53, 46, 18]. **Baselines.** We compare against 13 baselines in three categories. *Market benchmarks*: Equal-Weight [6], BTC Buy-and-Hold, and ETH Buy-and-Hold. *LLM multi-agent baselines*: FinCon [53] (hierarchical CVR), TradingAgents [46] (multi-role pipeline), FinMem [54] (layered memory), CryptoTrade [18] (single LLM + reflection), and FSReason [42] (dual Fact/Subjectivity agents, equal-weight fusion); all share MRC's backbone (Qwen3-VL) and temperature ($T_{\mathrm{LLM}}=0.7$). *DRL baselines*: A2C [28], PPO [33] ($\varepsilon=0.2$), DQN [27] (10 discrete templates), MAPPO [52] (three temporal actors, centralised critic), and MADDPG [23] (three actors + critics, 302-dim input); all trained online with 260-dim log-return features following the FinRL protocol [21, 20]. Full dataset and implementation details, please refer to Appendix D.

5.2 Experiment Results

We evaluate all methods using four metrics (definitions in Appendix B): cumulative return **CR** % (↑), Sharpe ratio **SR** (↑), maximum drawdown **MDD** % (↓), and information ratio **IR** vs. EW (↑).

RQ1: Does MRC outperform other baselines? Table 2 reports the overall comparison. MRC attains the best CR, SR, IR, and lowest MDD among all active methods, remaining within 2.1 percentage points of passive BTC on drawdown. This pattern matters because the strongest competing methods each fail on a different axis: FinMem has high return but high seed variance, CryptoTrade improves Sharpe, but not drawdown, and DRL baselines remain below equal weight on benchmark-relative skill. MRC also shows the tightest seed-level SR variance among LLM methods (± 0.07), suggesting that the Bayesian uniform prior and EWP cold-start mechanism improve stability rather than amplify early-period sensitivity. Full per-group visual comparisons are provided in Appendix H, and transaction cost sensitivity is reported in Table 9: at both 5 and 10 bps one-way per unit of

Table 3: Shapley Council ablation: each coalition as a standalone decision maker (5 seeds, mean $\pm$ std). Metrics as in Table 2. Superscript ** on MRC IR: two-sample t-test $p<0.05$ versus the second-best method (Grand Coalition) on IR.

Stage	Method	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)
Benchmarks	EW	279.93	1.14	42.78	0.00
	BTC	270.94	1.23	32.02	−0.07
	ETH	78.45	0.63	63.75	−0.59
Stage 1	Market Agent (A_1)	$393.5_{\pm 135.9}$	$1.54_{\pm 0.22}$	$47.5_{\pm 6.0}$	$0.08_{\pm 0.30}$
	On-chain Agent (A_2)	$238.3_{\pm 20.6}$	$1.18_{\pm 0.05}$	$30.8_{\pm 3.0}$	$-0.46_{\pm 0.11}$
	Macro Agent (A_3)	$129.5_{\pm 35.0}$	$0.83_{\pm 0.11}$	$40.9_{\pm 1.3}$	$-1.39_{\pm 0.29}$
Stage 2	$A_1 \oplus A_2$	$246.9_{\pm 106.2}$	$1.34_{\pm 0.25}$	$35.2_{\pm 4.9}$	$-0.43_{\pm 0.34}$
	$A_1 \oplus A_3$	$170.8_{\pm 95.2}$	$1.10_{\pm 0.28}$	$42.9_{\pm 5.6}$	$-0.68_{\pm 0.34}$
	$A_2 \oplus A_3$	$151.1_{\pm 55.6}$	$0.98_{\pm 0.17}$	$31.8_{\pm 2.0}$	$-0.95_{\pm 0.30}$
Stage 3	Grand Coalition	$286.0_{\pm 77.6}$	$1.33_{\pm 0.17}$	$38.2_{\pm 3.2}$	$-0.26_{\pm 0.31}$
Full System	**MRC (Ours)**	$440.1_{\pm 51.4}$	$1.51_{\pm 0.07}$	$34.1_{\pm 2.5}$	$0.47^{**}_{\pm 0.20}$

turnover, MRC still exceeds BTC Buy-and-Hold in cumulative return, while several high-turnover LLM baselines lose their zero-cost advantage. We use 10 bps one way as the reference rate for costed results. MRC turns over 128.7 ± 5.5 times a year and at that rate keeps SR 1.23 ± 0.06, the highest of all active methods and above Equal-Weight charged the same rate (1.13, Table 18). We therefore interpret the main result as a relative robustness claim under matched execution constraints, rather than as an unconditional claim about deployable alpha at arbitrary cost levels.

RQ2: Does MRC really work? Table 3 reports the coalition ablation. Stage 1 individual agents span a wide performance range (SR 0.83 to 1.54) with substantial inter-seed variance (SR ± 0.05 to 0.22). A_1 achieves the highest point-estimate SR (1.54) but also the deepest MDD (47.5%), while A_3 records the weakest SR (0.83) together with the widest MDD among Stage 1 agents (40.9%), showing that no single agent balances return and risk at the same time. Stage 2 pairwise coalitions consistently underperform the best Stage 1 individual portfolio on SR (1.34, 1.10, and 0.98 versus 1.54) while preserving or amplifying inter-seed variance, with $A_1 \oplus A_3$ CR variance reaching $\pm 95.2\%$, and offer no compensating improvement in MDD (see Appendix J for analysis of why pairwise debate attenuates rather than improves signal in portfolio construction). The Stage 3 grand coalition alone recovers SR to 1.33 but records MDD of 38.2%, leaving its risk-return profile worse than A_1 on both dimensions taken separately. The full MRC system records SR = 1.51 and MDD = 34.1%, reducing MDD by 13.4 percentage points relative to A_1 alone at a cost of only 0.03 in SR, tightening seed-level SR variance to ± 0.07, and producing the only statistically significant IR advantage in the table ($p<0.05$). For sensitivity analyses for temperature, burn-in, and transaction cost, see Appendix H. With the coalition outputs and risk-control overlays held fixed, the credit mechanism beats uniform weighting in all five seeds (Sharpe $+0.137$), and Appendix I compares it with three alternative aggregation rules and reports leave-one-out and robustness checks.

RQ3: Does MRC remain robust across different market regimes? Table 4 is a single-seed case study verifying that the regime-conditional mechanisms behave as designed. MAPPO and MADDPG achieve SR ≈ 2.26 in the bull regime but collapse to SR = 0.03 and 0.02 in volatile, and DQN turns negative (SR = -0.04) in bear, a pattern consistent with price-only state representations that carry no regime signal and sustain bull-market postures into fundamentally different conditions. At the agent level, A_1 alone holds 55.7% cash in bear, an overcautious posture driven by prior bull-period weight accumulation rather than current conditions, while A_3 collapses to SR ≈ 0.01 across both bear and volatile. These are precisely the dominance and collapse pathologies that the Bayesian Shapley update and regime-aware multiplier are designed to prevent. MRC achieves SR = 2.60 / 0.41 / 0.83 and mean cash of 8.0% / 28.4% / 13.6% across bull, bear, and volatile, a monotonically risk-adjusted cash posture absent in every DRL baseline where allocations remain near-static between 7% and 11% regardless of market condition, confirming that the adaptive mechanisms engage as intended. Over all five seeds (Table 17, Appendix I), MRC has the best mean SR rank of the 14 methods across the three regimes. It ranks first in bull (2.55 ± 0.05) and bear (0.63 ± 0.27), second in volatile (0.45 ± 0.22).

RQ4: How do the learned Shapley weights evolve? Figure 3 visualises how the Bayesian Shapley weights $\omega_i^{(t)}$, their underlying EWP-estimated Sharpe drivers $\rho_i^{(t)}$ and the conditioning regime score $\xi^{(t)}$ evolve over the evaluation window. Three observations stand out. First, no agent dominates the first 100 days on average. Over five seeds, the mean regime-adjusted weights of A_1, A_2 and A_3 are

Table 4: Regime-conditional performance decomposition (1,037 days). Bull / Bear / Volatile identified by $\xi^{(t)}$. Metrics as in Table 2. Red: best per column among MRC and LLM/DRL baselines (Shapley ablations and benchmarks excluded); Blue: second best. Single seed (T_{LLM}=0.7).

		Bull (430 d)				**Bear** (285 d)				**Volatile** (322 d)			
Category	**Method**	CR % (↑)	SR (↑)	MDD % (↓)	Cash %	CR % (↑)	SR (↑)	MDD % (↓)	Cash %	CR % (↑)	SR (↑)	MDD % (↓)	Cash %
Stage 1	Market (A_1)	315.8	2.61	20.8	22.7%	17.3	0.83	15.8	55.7%	44.7	1.38	31.6	41.0%
	On-chain (A_2)	217.5	2.37	23.6	15.1%	19.3	0.73	20.6	22.1%	−4.9	0.06	32.1	19.2%
	Macro (A_3)	226.9	2.22	17.9	7.2%	−9.1	0.02	31.1	10.5%	−8.4	0.01	34.5	9.3%
Stage 2	$A_1 \oplus A_2$	227.9	2.50	24.7	24.4%	15.3	0.79	16.9	42.7%	41.4	1.40	21.9	35.2%
	$A_1 \oplus A_3$	266.5	2.66	18.5	23.5%	1.3	0.21	17.2	48.2%	16.8	0.70	32.4	37.7%
	$A_2 \oplus A_3$	206.3	2.44	19.5	20.5%	23.6	0.86	20.6	26.0%	−7.7	−0.06	28.6	23.6%
Stage 3	Grand Coalition	274.7	2.55	18.3	17.7%	5.4	0.37	19.8	34.7%	31.2	1.05	28.3	29.6%
Benchmarks	EW	275.83	2.26	19.84	0.00	10.43	0.50	28.35	0.00	−8.46	0.04	40.34	0.00
	BTC	132.71	1.77	17.46	0.00	13.13	0.57	23.02	0.00	40.90	1.11	33.03	0.00
	ETH	297.71	2.08	24.19	0.00	−42.26	−0.79	50.14	0.00	−22.29	−0.19	51.63	0.00
LLM Baselines	FinCon	271.1	2.26	19.7	0.8%	7.8	0.45	28.8	0.7%	−8.4	0.04	40.3	0.8%
	TradingAgents	228.4	2.24	17.7	8.0%	9.4	0.48	28.5	8.0%	7.6	0.41	36.5	8.0%
	FinMem	305.6	2.54	17.9	8.0%	21.2	0.76	30.0	8.0%	17.5	0.65	32.4	8.0%
	CryptoTrade	299.7	2.51	19.7	7.7%	3.2	0.33	25.8	7.8%	16.0	0.60	37.3	7.8%
	FSReason	281.2	2.49	17.1	7.0%	4.9	0.37	28.3	7.5%	−1.0	0.20	33.5	7.1%
DRL Baselines	A2C	236.0	2.44	19.1	19.3%	19.2	0.74	20.5	19.7%	−0.2	0.20	31.3	19.9%
	PPO	186.5	2.00	18.4	11.2%	−0.1	0.25	28.6	10.6%	8.2	0.42	31.9	10.7%
	DQN	231.7	2.48	18.1	9.7%	−10.3	−0.04	28.5	9.1%	15.5	0.59	27.5	9.0%
	MAPPO	245.3	2.26	18.5	7.0%	10.9	0.52	26.5	7.0%	−7.7	0.03	38.1	7.0%
	MADDPG	246.3	2.26	18.6	7.1%	14.1	0.59	25.8	7.0%	−8.0	0.02	38.3	7.0%
Full System	**MRC (Ours)**	**353.7**	**2.60**	**20.9**	**8.0%**	**6.8**	**0.41**	**19.4**	**28.4%**	**25.9**	**0.83**	**29.8**	**13.6%**

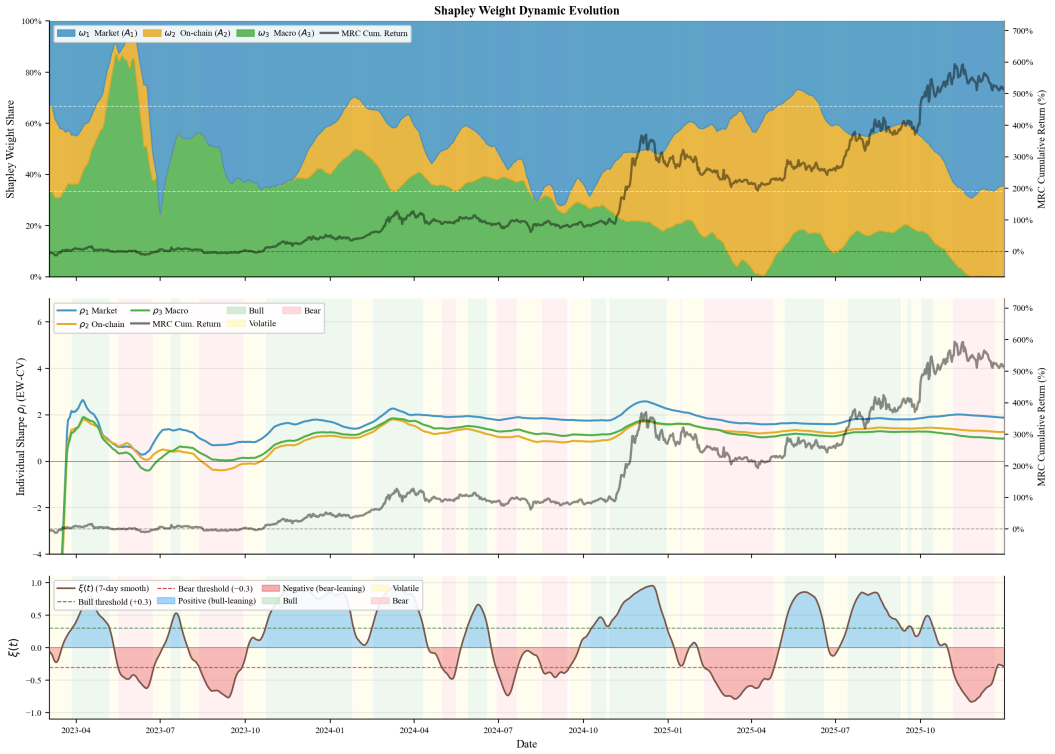

Figure 3: Shapley weight evolution in one seed. Top: Bayesian weights $\omega_i^{(t)}$ for three agents. Middle: $\rho_i^{(t)}$ for each Stage-1 individual coalition. Bottom: conditioning regime score $\xi^{(t)}$.

0.34, 0.24 and 0.42, while a single run can already follow early evidence, as the early peak of A_3 in Figure 3 shows. Second, after the cold-start period, A_1 (Market) becomes dominant, especially during bull regimes (green shading), consistent with its higher individual Sharpe. Third, over five seeds A_3 (Macro) falls to a mean regime-adjusted weight of 0.15 ± 0.05 after day 100 (against 0.50 ± 0.07 for A_1), reflecting its persistently lower ρ_3 throughout the risk-on environment of 2023 to 2025. Its weight would recover under a sustained bear regime. The interaction between Shapley weights and regime-consensus signals, and how the β_{gc} blend ratio responds to agent disagreement, is further illustrated in Figure 9. Additional backbone ablations are in Appendix H (Tables 10–13), token-level return breakdowns are in Appendix H, and agent-level profiles are in Appendix K.

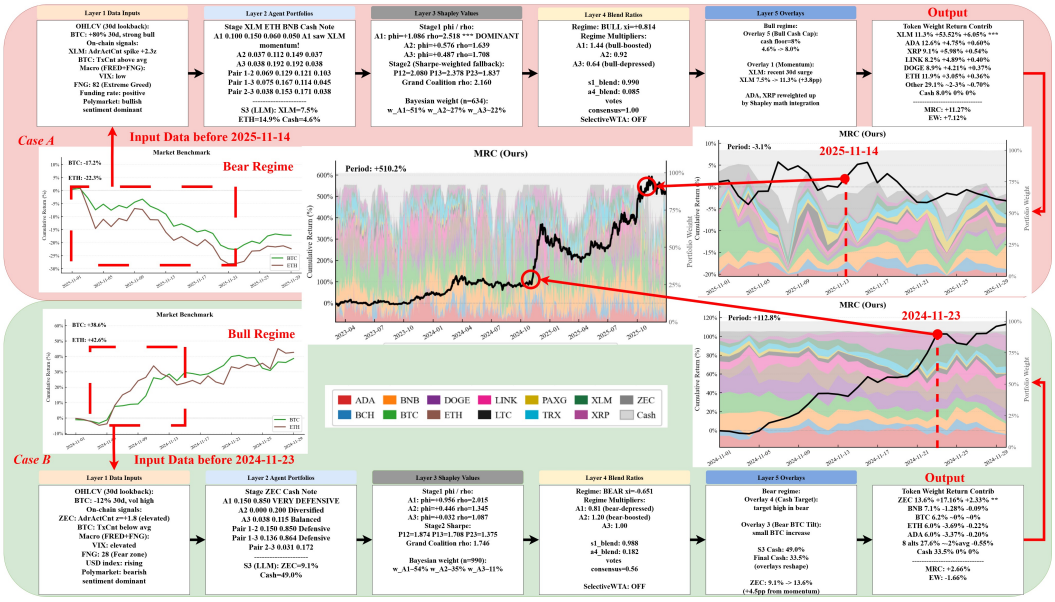

Figure 4: Full-chain explainability under regime-dependent rebalancing.

RQ5: How can MRC explain rebalancing decisions? Figure 4 shows how MRC performs rebalancing under opposite market regimes. Case A illustrates the bear-market path. When market momentum deteriorates, volatility rises, and sentiment turns bearish, MRC shifts toward a more defensive allocation with higher cash while preserving selective exposure supported by on-chain evidence. Case B illustrates the bull-market path. When momentum, sentiment, and regime indicators turn strongly positive, MRC assigns greater authority to the market specialist and increases exposure to trend-following assets. Together, the two cases show that MRC's explainability is not only post-hoc. The same five-layer Shapley report exposes how regime-dependent inputs drive the reallocation decision from raw signals to the final executed portfolio.

6 Conclusion

We introduced MRC, a Shapley-based cooperation mechanism for online credit assignment in heterogeneous multi-agent decision systems, validated at $N=3$ to cover the three orthogonal financial information channels. Across 1,037 trading days, MRC achieves the best CR, SR, and IR among the active baselines under matched execution constraints, while remaining close to passive BTC on drawdown. Coalition ablation indicates that the main gain comes from Shapley-weighted integration across coalition outputs rather than from any isolated stage alone. This provides evidence that principled credit reassignment can improve robustness in heterogeneous multi-agent decision systems.

The limitation is that the evaluation covers a single cryptocurrency market over one time window and relies on backtests. Live trading would add slippage, limited liquidity, and drift in specialist quality. Trading costs lower the returns of every active method, although MRC keeps its edge over Equal-Weight up to a higher cost than any other active method. Our comparison with non-LLM specialists uses a rule-based council rather than a fully matched-information baseline, which we leave to future work. With both kinds of specialists, the credit mechanism improves on uniform weighting, so the contribution does not depend on the LLM layer. Our future work will extend MRC beyond the current 13 cryptocurrency assets and low-cost execution setting to broader asset classes and cross-domain validation, and will test whether the Shapley-based credit assignment system remains effective as the number of agents scales.

Acknowledgments and Disclosure of Funding

This work was funded by UKRI EPSRC (grant No. EP/Y028392/1): AI for Collective Intelligence (AI4CI).

References

[1] A. Ang, J. Chen, and Y. Xing. Downside risk. *The Review of Financial Studies*, 19(4): 1191–1239, 2006.

[2] A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. García, S. Gil-López, D. Molina, R. Benjamins, et al. Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58:82–115, 2020.

[3] D. Beechey, T. M. Smith, and Ö. Şimşek. Explaining reinforcement learning with Shapley values. In *International Conference on Machine Learning*, pages 2003–2014. PMLR, 2023.

[4] J. Castro, D. Gómez, and J. Tejada. Polynomial calculation of the Shapley value based on sampling. *Computers & Operations Research*, 36(5):1726–1730, 2009.

[5] T. M. Cover. Universal portfolios. *Mathematical Finance*, 1(1):1–29, 1991.

[6] V. DeMiguel, L. Garlappi, and R. Uppal. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *The Review of Financial Studies*, 22(5):1915–1953, 2009.

[7] F. Doshi-Velez and B. Kim. Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*, 2017.

[8] Y. Du, S. Li, A. Torralba, J. B. Tenenbaum, and I. Mordatch. Improving factuality and reasoning in language models through multiagent debate. In *Forty-first International Conference on Machine Learning*, 2024.

[9] Y. Freund and R. E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119–139, 1997.

[10] R. C. Grinold and R. N. Kahn. *Active Portfolio Management*. McGraw Hill New York, 2000.

[11] S. Gu, B. Kelly, and D. Xiu. Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33(5):2223–2273, 2020.

[12] T. Guo, X. Chen, Y. Wang, R. Chang, S. Pei, N. V. Chawla, O. Wiest, and X. Zhang. Large language model-based multi-agents: a survey of progress and challenges. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, pages 8048–8057, 2024.

[13] E. Hazan. Introduction to online convex optimization. *Foundations and Trends in Optimization*, 2(3-4):157–325, 2016.

[14] S. Hong, M. Zhuge, J. Chen, X. Zheng, Y. Cheng, J. Wang, C. Zhang, Z. Wang, S. K. S. Yau, Z. Lin, et al. MetaGPT: Meta programming for a multi-agent collaborative framework. In *The Twelfth International Conference on Learning Representations*, 2023.

[15] J. Hull. *Risk Management and Financial Institutions*, volume 733. John Wiley & Sons, 2012.

[16] A. Jadbabaie, A. Rakhlin, S. Shahrampour, and K. Sridharan. Online optimization: Competing with dynamic comparators. In *Artificial Intelligence and Statistics*, pages 398–406. PMLR, 2015.

[17] J. Li, Y. Liu, W. Liu, S. Fang, L. Wang, C. Xu, and J. Bian. MarS: a financial market simulation engine powered by generative foundation model. In *The Thirteenth International Conference on Learning Representations*, 2025.

[18] Y. Li, B. Luo, Q. Wang, N. Chen, X. Liu, and B. He. CryptoTrade: A reflective LLM-based agent to guide zero-shot cryptocurrency trading. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 1094–1106, 2024.

[19] Y. Lin, S. Ye, Y. Yao, Z. Fang, B. Han, and T. Liu. SquidAgent: Parallelize wisely, coordinate efficiently. *arXiv preprint arXiv:2610.08647*, 2026.

[20] X.-Y. Liu, H. Yang, Q. Chen, R. Zhang, L. Yang, B. Xiao, and C. D. Wang. FinRL: A deep reinforcement learning library for automated stock trading in quantitative finance. *Deep RL Workshop, NeurIPS 2020*, 2020.

[21] X.-Y. Liu, Z. Xia, J. Rui, J. Gao, H. Yang, M. Zhu, C. Wang, Z. Wang, and J. Guo. FinRL-Meta: Market environments and benchmarks for data-driven financial reinforcement learning. *Advances in Neural Information Processing Systems*, 35:1835–1849, 2022.

[22] J. Longerstaey and M. Spencer. *RiskMetrics—Technical Document*. J.P.Morgan/Reuters, 1996.

[23] R. Lowe, Y. I. Wu, A. Tamar, J. Harb, O. Pieter Abbeel, and I. Mordatch. Multi-agent actor-critic for mixed cooperative-competitive environments. *Advances in Neural Information Processing Systems*, 30, 2017.

[24] S. M. Lundberg and S.-I. Lee. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 2017.

[25] Y. Luo, Y. Feng, J. Xu, P. Tasca, and Y. Liu. LLM-powered multi-agent system for automated crypto portfolio management. *arXiv preprint arXiv:2501.00826*, 2025.

[26] T. Miller. Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267:1–38, 2019.

[27] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, et al. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533, 2015.

[28] V. Mnih, A. P. Badia, M. Mirza, A. Graves, T. Lillicrap, T. Harley, D. Silver, and K. Kavukcuoglu. Asynchronous methods for deep reinforcement learning. In *International Conference on Machine Learning*, pages 1928–1937. PMLR, 2016.

[29] M. Ottaviani and P. Sørensen. Information aggregation in debate: who should speak first? *Journal of Public Economics*, 81(3):393–421, 2001.

[30] J. S. Park, J. O'Brien, C. J. Cai, M. R. Morris, P. Liang, and M. S. Bernstein. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, pages 1–22, 2023.

[31] Y. Pei, J. Cartlidge, A. Mandal, D. Gold, E. Marcilio, and R. Mazzon. Cross-modal temporal fusion for financial market forecasting. In *ECAI-2025, 28th European Conference on Artificial Intelligence including 14th Conference on Prestigious Applications of Intelligent Systems (PAIS)*, volume 413, pages 5360–5367, 2025.

[32] M. T. Ribeiro, S. Singh, and C. Guestrin. " Why should I trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1135–1144, 2016.

[33] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

[34] L. S. Shapley. A value for n-person games. In H. W. Kuhn and A. W. Tucker, editors, *Contributions to the Theory of Games II*, pages 307–317. Princeton University Press, Princeton, NJ, 1953. doi: 10.1515/9781400881970-018.

[35] W. F. Sharpe. The Sharpe ratio. *The Journal of Portfolio Management*, 21(1):49–58, 1994. doi: 10.3905/jpm.1994.409501.

[36] S. Shen, C. Ma, C. Li, W. Liu, Y. Fu, S. Mei, X. Liu, and C. Wang. RiskQ: risk-sensitive multi-agent reinforcement learning value factorization. *Advances in Neural Information Processing Systems*, 36:34791–34825, 2023.

[37] N. Shinn, F. Cassano, A. Gopinath, K. Narasimhan, and S. Yao. Reflexion: Language agents with verbal reinforcement learning. *Advances in Neural Information Processing Systems*, 36: 8634–8652, 2023.

[38] J. Wang, Y. Zhang, T.-K. Kim, and Y. Gu. Shapley Q-value: A local reward approach to solve global reward games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 7285–7292, 2020.

[39] J. Wang, Y. Zhang, Y. Gu, and T.-K. Kim. SHAQ: Incorporating Shapley value theory into multi-agent Q-learning. *Advances in Neural Information Processing Systems*, 35:5941–5954, 2022.

[40] J. Wang, Y. Li, Y. Zhang, W. Pan, and S. Kaski. Open ad hoc teamwork with cooperative game theory. In *International Conference on Machine Learning*, volume 235 of *PMLR*, pages 50902–50930, 2024. URL `https://proceedings.mlr.press/v235/wang24an.html`.

[41] J. Wang, Y. Li, S. Kaski, and J. Lawry. Shapley machine: A game-theoretic framework for n-agent ad hoc teamwork. *arXiv preprint arXiv:2506.11285*, 2025.

[42] Q. Wang, Y. Gao, Z. Tang, B. Luo, N. Chen, and B. He. Exploring LLM cryptocurrency trading through fact-subjectivity aware reasoning. In *Proceedings of the ICLR 2025 Workshop on Advances in Financial AI*, 2025.

[43] J. Wei, X. Wang, D. Schuurmans, M. Bosma, F. Xia, E. Chi, Q. V. Le, D. Zhou, et al. Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, 35:24824–24837, 2022.

[44] Q. Wu, G. Bansal, J. Zhang, Y. Wu, B. Li, E. Zhu, L. Jiang, X. Zhang, S. Zhang, J. Liu, et al. AutoGen: Enabling next-gen LLM applications via multi-agent conversations. In *First Conference on Language Modeling*, 2024.

[45] S. Wu, O. Irsoy, S. Lu, V. Dabravolski, M. Dredze, S. Gehrmann, P. Kambadur, D. Rosenberg, and G. Mann. BloombergGPT: A large language model for finance. *arXiv preprint arXiv:2303.17564*, 2023.

[46] Y. Xiao, E. Sun, D. Luo, and W. Wang. TradingAgents: Multi-agents LLM financial trading framework. *arXiv preprint arXiv:2412.20138*, 2024.

[47] Q. Xie, W. Han, Z. Chen, R. Xiang, X. Zhang, Y. He, M. Xiao, D. Li, Y. Dai, D. Feng, et al. FinBen: A holistic financial benchmark for large language models. *Advances in Neural Information Processing Systems*, 37:95716–95743, 2024.

[48] H. Yang, X.-Y. Liu, and C. D. Wang. FinGPT: Open-source financial large language models. *arXiv preprint arXiv:2306.06031*, 2023. URL `https://arxiv.org/abs/2306.06031`. FinLLM Workshop at IJCAI 2023.

[49] Y. Yang, Y. Tang, and K. Y. Tam. InvestLM: A large language model for investment using financial domain instruction tuning. *arXiv preprint arXiv:2309.13064*, 2023.

[50] S. Yao, D. Yu, J. Zhao, I. Shafran, T. Griffiths, Y. Cao, and K. Narasimhan. Tree of thoughts: Deliberate problem solving with large language models. *Advances in Neural Information Processing Systems*, 36:11809–11822, 2023.

[51] S. Yao, J. Zhao, D. Yu, N. Du, I. Shafran, K. Narasimhan, and Y. Cao. ReAct: Synergizing reasoning and acting in language models. In *11th International Conference on Learning Representations, ICLR 2023*, 2023.

[52] C. Yu, A. Velu, E. Vinitsky, J. Gao, Y. Wang, A. Bayen, and Y. Wu. The surprising effectiveness of PPO in cooperative multi-agent games. *Advances in Neural Information Processing Systems*, 35:24611–24624, 2022.

[53] Y. Yu, Z. Yao, H. Li, Z. Deng, Y. Jiang, Y. Cao, Z. Chen, J. W. Suchow, Z. Cui, R. Liu, et al. FinCon: A synthesized LLM multi-agent system with conceptual verbal reinforcement for enhanced financial decision making. *Advances in Neural Information Processing Systems*, 37: 137010–137045, 2024.

[54] Y. Yu, H. Li, Z. Chen, Y. Jiang, Y. Li, J. W. Suchow, D. Zhang, and K. Khashanah. FinMem: A performance-enhanced LLM trading agent with layered memory and character design. *IEEE Transactions on Big Data*, 2025.

[55] Y. Yuzhe, Y. Zhang, M. Wu, K. Zhang, Y. Zhang, H. Yu, Y. Hu, and B. Wang. TwinMarket: A scalable behavioral and social simulation for financial markets. In *Thirty-Ninth Annual Conference on Neural Information Processing Systems*, 2026.

[56] W. Zhang, L. Zhao, H. Xia, S. Sun, J. Sun, M. Qin, X. Li, Y. Zhao, Y. Zhao, X. Cai, 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*, pages 4314–4325, 2024.

[57] M. Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th International Conference on Machine Learning (ICML-03)*, pages 928–936, 2003.

Appendix

A Extended Related Work

A.1 LLM-based Multi-Agent Trading System

LLM multi-agent systems for financial decision-making [53, 46, 54, 56, 25, 48, 45, 49] have demonstrated strong empirical performance, but none resolves the credit assignment problem. General-purpose LLM agent frameworks [14, 44, 30] confirm that role specialisation and structured communication improve collective reasoning, yet lack principled mechanisms for dynamic credit assignment under non-stationary conditions. **FinCon** [53] is the closest predecessor: it relies on Conceptual Verbal Reinforcement (CVR)—a language-based feedback loop in which a Manager Agent distills prior analysis into "conceptual signals" for the downstream Analyst Agents—without any cooperative game-theoretic credit assignment, online agent reweighting, or regime conditioning. **TradingAgents** [46] and **FinMem** [54] rely on heuristic weighting and layered memory respectively, and neither can identify which modality drove a given decision. DRL methods (A2C, PPO, DQN) via **FinRL** [21] handle credit implicitly through policy gradient objectives, but treat the portfolio as a monolithic black box over price-only features, with no concept of heterogeneous agent coalitions.

The online credit assignment mechanism of MRC connects directly to the classical online portfolio selection literature. Universal Portfolio [5] and its variants establish that exponentially-weighted expert aggregation achieves logarithmic regret against the best fixed portfolio in hindsight, and the Hedge algorithm [9] achieves minimax-optimal regret under the same expert-aggregation model. MRC instantiates this principle at the agent-coalition level: the EWP characteristic functions implement a discounted online learning objective [57, 13], while the Bayesian adaptive mixture corresponds structurally to Hedge-style exponential weighting. The key departure is that Shapley values replace the undifferentiated loss surrogate of standard expert aggregation with an attribution derived from axiomatic Shapley values, enabling credit to track each modality's marginal coalition contribution rather than treating all agents as interchangeable.

A.2 Regime Modeling and Detection

The core difficulty in online financial learning is non-stationarity: a model optimal in one regime is harmful in another, yet standard algorithms are slow to unlearn early advantage [57, 22]. In LLM multi-agent systems, where agent parameters are frozen, all adaptation must occur at the weighting layer, making cold-start dominance an especially acute failure mode. DRL policies are equally vulnerable, which means price-only state features discard the on-chain and macro signals most predictive of regime transitions, and the same network is applied uniformly across bull, volatile, and bear conditions.

MRC decouples regime detection from agent weighting: the Regime Shift Perceptor computes a continuous market-state score independently of the LLM agents, then modulates Shapley weights through smooth regime-conditional multipliers. **MarS** [17] and **TwinMarket** [55] confirm that bull, volatile, and bear regimes reflect qualitatively distinct market dynamics warranting different agent authority structures, providing empirical motivation for this design. This is consistent with the broader financial literature showing that factor exposures and model performance are strongly regime-dependent [11]. **A2C**, **PPO**, and **DQN** are included as direct baselines following the evaluation protocol of **FinCon** [53].

A.3 Explainability in Trading Systems

Existing financial AI systems provide no quantitative attribution linking a specific allocation to a specific data source or deliberation stage, a gap highlighted by standardized benchmarks [47, 21]. **ReAct** [51] improves LLM auditability through interleaved reasoning; **Chain-of-Thought** [43], **Tree-of-Thoughts** [50], and **Reflexion** [37] further demonstrate that structured deliberation elicits more reliable outputs from LLMs. MRC extends this line of work by connecting each agent's chain-of-thought directly to final portfolio weights through the Shapley computation, making the reasoning trace causally (not merely rhetorically) linked to the decision. The information economics result [29] that transparent aggregation dominates opaque pooling when experts disagree motivates MRC's Shapley-weighted divergence discount, which surfaces inter-agent disagreement explicitly rather than

absorbing it silently. MRC's five-layer attribution system directly addresses the opacity gap shared by all competing systems [53, 46, 54]. This system complements model-level explainability tools such as **LIME** [32] and **SHAP** [24] by operating at the system level, and attributing final allocations to agent coalitions rather than input features.

B Evaluation Metrics

We evaluate all methods using four complementary metrics that jointly capture profitability, risk-adjusted efficiency, downside risk, and benchmark-relative skill.

Cumulative Return (CR%) [15, 53] measures the total compounded growth of a portfolio over the evaluation window, providing a direct summary of long-term profitability. Let $R^{(t)}$ denote the portfolio's net return on day t; the cumulative return over T days is

$$\mathrm{CR} = \prod_{t=1}^{T}(1 + R^{(t)}) - 1. \tag{19}$$

A higher CR indicates a more profitable strategy over the evaluation period.

Sharpe Ratio (SR) [35]: Sharpe Ratio quantifies risk-adjusted return by normalising the excess return of a portfolio above the risk-free rate R_f by its total return volatility σ_p, as shown in Eq. (20).

$$\mathrm{SR} = \sqrt{365} \cdot \frac{\bar{R}_p - R_f}{\hat{\sigma}_p}, \tag{20}$$

where $\bar{R}_p$ is the mean daily portfolio return, $\hat{\sigma}_p$ is the sample standard deviation of daily returns, $R_f = 0$ is the risk-free rate used in the crypto setting, and the factor $\sqrt{365}$ annualizes the ratio for daily crypto markets.

A higher SR indicates more return per unit of total risk and is therefore used to compare risk-adjusted performance across methods. Interpretation remains context-dependent [53].

Maximum Drawdown (MDD%) [1]: MDD captures the worst peak-to-trough loss a portfolio sustains before recovering to a new high, measuring the largest relative decline from a peak value P_{peak} to a subsequent trough P_{trough}, as detailed in Eq. (21).

$$\mathrm{MDD} = \max\left(\frac{P_{\text{peak}} - P_{\text{trough}}}{P_{\text{peak}}}\right). \tag{21}$$

A lower MDD reflects stronger capital preservation and greater resilience under adverse market conditions.

Information Ratio (IR) [10] measures the consistency with which a strategy outperforms a passive benchmark, penalising high tracking error. We use the equal-weight $(1/K)$ portfolio, denoted EW, as the benchmark:

$$\mathrm{IR} = \frac{\bar{R}_p - \bar{R}_b}{\hat{\sigma}(R_p - R_b)}, \tag{22}$$

where $\bar{R}_p$, $\bar{R}_b$ are the mean daily returns of the portfolio and benchmark respectively, and $\hat{\sigma}(R_p - R_b)$ is the sample standard deviation of the daily active return. A higher IR indicates more consistent outperformance relative to the benchmark.

C Algorithm Pseudocode

We provide complete pseudocode for the two core routines of MRC.

Algorithm 1 details the offline weight-update step executed once per period after realized returns are observed: it estimates coalition characteristic functions via EWP (Phase 1), computes exact three-player Shapley values (Phase 2), applies Bayesian adaptive mixing (Phase 3), and conditionally activates the Selective WTA override (Phase 4), corresponding to §4.1 of the main paper.

Algorithm 2 details the *online* per-period inference step: it runs the three-stage cooperative deliberation (Stages 1–3), blends coalition portfolios via the data-driven ratios $\beta_{S1}^{(t)}$ and $\beta_{\text{gc}}^{\text{final}}$, applies the

Algorithm 1 MRC Shapley Value Ledger Update

Require: Return history $\{R_S^{(\tau)}\}$ for all $S \subseteq \mathcal{N}$, $\tau = 1, \ldots, t$; period count t; decay period h; concentration parameter λ; WTA threshold θ_{WTA}; WTA weight ω_{wta}; uniform prior $\pi_i = 1/|\mathcal{N}| = 1/3$ for all i

Ensure: Updated agent weights $\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3)$, pairwise weights $\{p_{ij}\}$

Phase 1: EWP Estimation of Coalition Characteristic Functions

for each non-empty $S \subseteq \mathcal{N}$ **do**
  Compute $w_\tau \leftarrow \exp\big(-(t-\tau)/h\big)$ for $\tau = 1, \ldots, t$ ▷ Eq. (13)
  Compute $\hat{\mu}_{\text{EW}}(S; \mathcal{H}_t)$, $\hat{\sigma}_{\text{EW}}(S; \mathcal{H}_t)$ ▷ Eq. (14)
  $v(S) \leftarrow \gamma_\rho \cdot \sqrt{365}\, \hat{\mu}_{\text{EW}}(S; \mathcal{H}_t)/\hat{\sigma}_{\text{EW}}(S; \mathcal{H}_t) + \gamma_\mu \cdot \mu_{\text{ann}}(S; \mathcal{H}_t)$ ▷ Eq. (2)
end for

Phase 2: Exact Three-Player Shapley Computation

$(\phi_1, \phi_2, \phi_3) \leftarrow \text{ShapleyValues}(v_1, v_2, v_3, v_{12}, v_{13}, v_{23}, v_{123})$ ▷ Eqs. (29)–(31)
$\phi_i^+ \leftarrow \max(0, \phi_i)$ for each $i \in \mathcal{N}$ ▷ Dummy-player truncation

Phase 3: Bayesian Adaptive Weight Update

$\alpha \leftarrow 1 - \exp(-t/\lambda)$ ▷ Eq. (3)
for each $i \in \mathcal{N}$ **do**
  $\omega_i \leftarrow \alpha \cdot \phi_i^+ / \sum_j \phi_j^+ + (1-\alpha) \cdot \pi_i$ ▷ Eq. (4)
end for
Compute $\{p_{ij}\}$ analogously using $v(\{i,j\})$ values and pairwise priors

Phase 4: Selective Winner-Takes-All Override

$i^* \leftarrow \arg\max_i \rho_i^{(n_{\text{win}})}$ ▷ rolling Sharpe over n_{win}-period window
if $\rho_{i^*}^{(n_{\text{win}})} \,/\, \text{mean}_{j \neq i^*}(\rho_j^{(n_{\text{win}})}) \;\geq\; \theta_{\text{WTA}}$ **then**
  $\omega_{i^*} \leftarrow \omega_{\text{wta}}$
  $\omega_j \leftarrow (1 - \omega_{\text{wta}}) \cdot \omega_j / \sum_{k \neq i^*} \omega_k$ for all $j \neq i^*$
end if
return $\boldsymbol{\omega}$, $\{p_{ij}\}$

risk-control cascade (§4.3), and then calls Algorithm 1 to update the Shapley ledger $\mathcal{L}$ with the new realized return.

The two algorithms together implement the closed online learning loop described in §4.

D Implementation Details

Dataset Table 5 summarizes the full dataset inventory used by MRC, organized by observer agent, stream type, source, frequency, coverage window, and feature detail. The paragraphs below expand each input group and describe the corresponding preprocessing pipeline. On decision day t every agent sees only rows dated t or earlier. Daily values for day t are final when day t closes, which is also when the portfolio is executed, and the portfolio earns the return from day t to day $t+1$.

Table 5: Dataset inventory organized by MRC observer agent. *ts* = time-series; *img* = image; *text* = natural language. Row date: the date each row carries. [a]The four daily FRED series shown to A_3 are US market quotes fixed before midnight UTC of their observation date, and FRED posts them one or more business days later. [b]Sentiment labels and engagement counts come from a collection made after the evaluation window.

Agent	Category	Stream	Type	Source	Freq.	Row date	Raw Coverage	Detail
A_1	Historical Data	OHLCV prices	*ts*	Binance	Daily	UTC day	2017-08–2025-12	5 feat., 13 assets
	K-line Charts	Candlestick w/ MA overlays	*img*	`mplfinance`	Daily	UTC day	2017-08–2025-12	LB = 90 d, MA(5,20,60)
A_2	Historical Data	OHLCV prices	*ts*	Binance	Daily	UTC day	2017-08–2025-12	5 feat., 13 tokens
	On-chain Data	Per-token on-chain activity	*ts*	CoinMetrics / DeFiLlama / Dune	Daily	UTC day	2020-12–2025-12	3 feat., 13 tokens
A_3	Social Media	Twitter/X crypto sentiment	*text*	X (Twitter)	Event-driven	Day of post[b]	2023-01–2025-12	7 features
	Social Media	Prediction-market probabilities	*ts*	Polymarket	Daily	Day of price	2023-02–2025-12	2 features
	Macro Indices	FRED macroeconomic indicators	*ts*	St. Louis Fed (FRED)	Daily/Monthly	US trading day[a]	2019-12–2025-12	10 series, 4 shown
	Macro Indices	Web3 macro fundamentals	*ts*	Blockchain.com / Binance / Alternative.me	Daily	UTC day	2018-01–2025-12	12 series

Market Data (A_1). Daily OHLCV bars are stored as HDF5 files (one per asset) and accessed via a 90-day rolling lookback window at each decision period. Multi-scale statistics are computed at 1-day, 7-day, and 30-day horizons, covering returns, realized volatility, RSI, and ATR. Candlestick

Algorithm 2 MRC Per-Period Inference

Require: Data streams $\mathbf{x}_1^{(t)}, \mathbf{x}_2^{(t)}, \mathbf{x}_3^{(t)}$; current agent weights $\boldsymbol{\omega}$, pairwise weights $\{p_{ij}^{(t)}\}$; Shapley ledger $\mathcal{L}$; price history $\mathcal{H}_{\text{price}}$

Ensure: Portfolio $\mathbf{w}^{(t)}$, updated ledger $\mathcal{L}$

$\xi^{(t)} \leftarrow \text{RegimeScore}(\mathcal{H}_{\text{price}})$ ▷ Eq. (12)

Stage 1: Private Observation (executed in parallel)

for $i \in \{1, 2, 3\}$ **in parallel do**

  $(\mathbf{w}_i^{(t)},\, \hat{r}_i^{(t)}) \leftarrow \text{LLM}_{A_i}\Big(\mathbf{x}_i^{(t)}\Big)$

end for

Stage 2: Pairwise Socratic Debate (executed in parallel)

for $(i, j) \in \{(1, 2), (1, 3), (2, 3)\}$ **in parallel do**

  $\mathbf{w}_{ij}^{(t)} \leftarrow \text{LLM}_{\text{Debate}}\Big(A_i, A_j, \mathbf{w}_i^{(t)}, \mathbf{w}_j^{(t)}\Big)$

end for

Regime-Aware Multiplier and Consensus

$\kappa^{(t)} \leftarrow \text{ShapleyWeightedPlurality}\Big(\boldsymbol{\omega}, \{\hat{r}_i^{(t)}\}\Big)$ ▷ Eq. (9)

$\tilde{\boldsymbol{\omega}}^{(t)} \leftarrow \text{RegimeAwareMultiplier}\big(\boldsymbol{\omega}, \xi^{(t)}, \kappa^{(t)}\big)$ ▷ Eq. (15)

Stage 1/2 Ensemble Blend

$\beta_{S1}^{(t)} \leftarrow \bar{\beta}_{S1} + \delta_{S1} \cdot \tanh(-\Delta_{S2}^{(t)}/\tau_{S1})$ ▷ Eq. (6)

Stage 3: Grand Coalition Synthesis

$\text{report} \leftarrow \mathcal{L}.\text{FormatShapleyReport}()$

$\mathbf{w}_{123}^{(t)} \leftarrow \text{LLM}_{A4}\Big(\{\mathbf{w}_i^{(t)}\}, \{\mathbf{w}_{ij}^{(t)}\}, \text{report}\Big)$

Council Blend with Divergence Discount

$\beta_{\text{gc}}^{(t)} \leftarrow \max\Big(0,\, \bar{\beta}_{\text{gc}} + \delta_{\text{gc}} \cdot \tanh(\Delta_{\text{gc}}^{(t)}/\tau_{\text{gc}})\Big)$ ▷ Eq. (7)

$\beta_{\text{gc}}^{\text{final}} \leftarrow \beta_{\text{gc}}^{(t)} \cdot (1/2 + \kappa^{(t)}/2)$ ▷ Eq. (10)

$\mathbf{w}_{\text{council}}^{(t)} \leftarrow \beta_{\text{gc}}^{\text{final}}\, \mathbf{w}_{123}^{(t)} + (1 - \beta_{\text{gc}}^{\text{final}})\Big[\beta_{S1}^{(t)} \sum_i \tilde{\omega}_i^{(t)} \mathbf{w}_i^{(t)} + (1 - \beta_{S1}^{(t)}) \sum_{(i,j)} p_{ij}^{(t)} \mathbf{w}_{ij}^{(t)}\Big]$ ▷ Eq. (11)

Risk Control Cascade (§4.3)

$\mathbf{w}^{(t)} \leftarrow \text{RiskControl}\Big(\mathbf{w}_{\text{council}}^{(t)},\, \xi^{(t)},\, \mathbf{x}_2^{(t)}\Big)$ ▷ see Appendix E

Update Shapley Value Ledger (Algorithm 1)

$\mathcal{L} \leftarrow \text{ShapleyLedgerUpdate}\big(\mathcal{L},\, \mathbf{w}^{(t)},\, \mathbf{r}^{(t+1)}\big)$

return $\mathbf{w}^{(t)},\ \mathcal{L}$

K-line chart images are rendered with `mplfinance` at daily stride over the same 90-day window, with moving-average overlays at periods $(5, 20, 60)$, and passed as base64-encoded PNG attachments to the vision encoder.

On-chain Feature Assignment (A_2). Per-token on-chain activity metrics come from a proprietary dataset with 1,827 daily rows and 39 feature columns, corresponding to three features per token across 13 assets. Each token's three features are drawn from five metric families: active address count (AdrActCnt), transaction count (TxCnt), native fee total (FeeTotNtv), chain TVL in USD (ChainTVL_USD), and token transfer count (TokenTxCnt). The specific assignment depends on protocol type and data availability: standard PoW/PoS L1s use (AdrActCnt, TxCnt, FeeTotNtv), DeFi-heavy chains (BNB, ETH, TRX) substitute ChainTVL_USD, and ERC-20 tokens (LINK, PAXG) substitute TokenTxCnt for the unavailable native fee series.

Macro Data (A_3). Ten FRED macroeconomic series are provided to A_3: five at daily frequency (10-year real yield, USD broad index, overnight reverse repo, VIX, WTI crude) and five at monthly frequency (CPI, core PCE, unemployment rate, nonfarm payrolls, industrial production). These are complemented by the Crypto Fear & Greed Index presented as a 30-day trailing window, Polymarket prediction-market probabilities (209,128 rows across 575 topics), and Twitter/X crypto sentiment aggregated from ≈27,000 engagement-weighted tweets with human-assigned sentiment labels spanning

all 13 assets from 2023 to 2025. The Twitter/X signal constitutes one of several feature inputs to A_3 and is interchangeable with any engagement-weighted sentiment source without architectural change.

Social Sentiment Signal Construction. Polymarket prediction-market sentiment is computed per token by filtering slugs with token-specific keywords, retaining only `outcome_norm = yes` rows, and mapping the mean implied probability $\bar{p}_k \in [0, 1]$ to a sentiment score $(\bar{p}_k - 0.5) \times 2 \in [-1, +1]$. For tokens with fewer than five matching slugs in a given window (primarily H1 2023), a general crypto-market fallback is applied using market-wide slug aggregates.

Twitter/X sentiment scores are derived per token from the curated tweet dataset as follows. Each tweet may mention multiple assets via a pipe-delimited `affected_cryptos` field; rows are first expanded so that each (tweet, token) pair becomes a separate record. A sentiment label $s_m \in \{-1, 0, +1\}$ (bearish, neutral, bullish) and human-assigned confidence $c_m \in (0, 1]$ are combined with an engagement weight

$$w_m \;=\; \log\bigl(1 + \ell_m + r_m + v_m/100\bigr) + 1,$$

where ℓ_m, r_m, v_m are likes, reposts, and views respectively (the $+1$ baseline prevents zero-weight tweets from being discarded). The token-level sentiment signal is then the engagement-weighted average

$$s_k \;=\; \frac{\sum_m s_m \, c_m \, w_m}{\sum_m w_m}.$$

Data Cleaning. All cleaning is applied after raw loading and before agent dispatch. FRED daily series (VIX, DXY, real yield, overnight reverse repo, WTI) are forward-filled with a 5-day limit to fill weekends and holidays; FRED monthly series (CPI, core PCE, unemployment, nonfarm payrolls, industrial production) are held constant until the next scheduled release. BTC blockchain and Binance-sourced macro series are forward-filled with a 5-day limit. The Fear & Greed Index and per-token on-chain series are forward-filled with a 5-day limit. No look-ahead filling is performed: forward-fill propagates only the last observed value. No survivorship bias is introduced: all 13 assets were actively traded throughout the evaluation window.

LLM Output Enforcement. All agents use `force_json=True`, implemented via a two-layer fallback: (1) the user prompt is appended with a structured output trigger instructing the model to respond with a valid JSON object matching the declared schema; (2) if the primary response fails schema validation, the system retries once with a stricter prefill (`{"`) to guide the tokenizer toward JSON continuation. LLM temperature is fixed at $T_{\text{LLM}} = 0.7$ across all agents (primary configuration; see Table 7 for temperature sensitivity); no top-p or top-k truncation is applied beyond the model defaults. The 90-day rolling lookback window is passed as a structured table in the system prompt; K-line chart images are passed as base64-encoded PNG attachments to the vision encoder.

Backbone Configuration and Evaluation Protocol. All LLM agents access the Qwen3-VL family via OpenRouter, and we evaluate seven backbone variants spanning instruct and extended-thinking modes (`8b`, `32b`, `30b-a3b`, `235b-a22b`). All Shapley, EWP, and WTA hyperparameters are fixed before the evaluation window and are not tuned on test data: Bayesian concentration λ=30, EWP e-fold decay h=252, characteristic-function weights γ_ρ=0.4 and γ_μ=0.6 (see sensitivity analysis below), WTA rolling window n_{win}=30, dominance threshold θ_{WTA}=1.8, WTA concentration weight ω_{wta}=0.80, regime thresholds $\xi_+ = +0.30$ (bull) and $\xi_- = -0.30$ (bear) with volatile otherwise, and regime-score attenuation factor 0.5 applied when the 7-day direction conflicts with the 30-day trend and the 7-day magnitude exceeds 30% of the 30-day signal. Each agent operates on a 90-day rolling lookback window, decisions use only information dated t or earlier, results in the main paper are averaged over five seeds, and all baselines share the same matched execution constraints.

Blend Ratio Hyperparameters. The Stage 1/2 blend (Eq. (6)) and grand-coalition blend (Eq. (7)) use the following fixed values, set prior to the evaluation window: $\bar{\beta}_{S1}$=0.90, δ_{S1}=0.09, τ_{S1}=0.075; $\bar{\beta}_{\text{gc}}$=0.15, δ_{gc}=0.20, τ_{gc}=0.10. The center $\bar{\beta}_{S1}$=0.90 encodes the prior that Stage-1 individual analyses dominate when Stage-2 debate provides no additional coalition value ($\Delta_{S2}^{(t)}$=0); δ_{S1} bounds the maximum deviation to $\pm$0.09, keeping $\beta_{S1}^{(t)}$ in the natural range $(0.81,\ 0.99)$ without explicit clipping. The center $\bar{\beta}_{\text{gc}}$=0.15 reflects the empirical finding that the grand-coalition LLM readout

contributes modestly on average; the wider scale δ_{gc}=0.20 allows it to rise toward 0.35 when A_4 synthesis demonstrably outperforms the Stage-1 ensemble.

	$\xi = +1$ (bull)	$\xi = 0$ (neutral)	$\xi = -1$ (bear)
ψ_1 (Technical, A_1)	1.50	1.20	0.60
ψ_2 (On-chain, A_2)	0.90	1.00	0.90
ψ_3 (Macro, A_3)	0.60	0.80	1.50

Table 6: Regime-conditional multiplier anchor values for $\psi_i(\xi^{(t)})$. Values are linearly interpolated, and the neutral column ($\xi = 0$) serves as the prior when no regime signal is present.

Regime-Aware Multiplier Values (ψ_i). The regime-conditional scalar ψ_i (Eq. (15)) is obtained by linearly interpolating between three anchor points at ξ=+1 (strong bull), ξ=0 (neutral), and ξ=−1 (strong bear). The anchor values, fixed prior to the evaluation window, are listed in Table 6.

Characteristic Function Weight Sensitivity (γ_ρ). For brevity, we omit the history argument and write $v(S) \equiv v(S; \mathcal{H}_t)$, $\text{SR}_{\text{EW}}(S) \equiv \sqrt{365}\,\hat{\mu}_{\text{EW}}(S; \mathcal{H}_t)/\hat{\sigma}_{\text{EW}}(S; \mathcal{H}_t)$. Under this shorthand, the blend is $v(S) = \gamma_\rho \cdot \text{SR}_{\text{EW}}(S) + \gamma_\mu \cdot \mu_{\text{ann}}(S)$. The weight γ_ρ=0.4 is chosen so that both terms in $v(S)$ contribute at comparable magnitudes: $\text{SR}_{\text{EW}}(S) \in [0.8, 1.6]$ and $\mu_{\text{ann}}(S) \in [0.3, 1.2]$ across the evaluation window.

This value was set by inspecting pre-evaluation return magnitudes and was not tuned on test data. A sweep over $\gamma_\rho \in \{0.2, 0.4, 0.6, 0.8\}$ shows rank correlation of ϕ_i vectors above 0.93 across all tested values.

Asymmetric EMA Smoothing. The council output $\mathbf{w}^{(t)}_{\text{council}}$ (Eq. (11)) is smoothed before execution via an asymmetric EMA:

$$\mathbf{w}^{(t)} \;=\; \eta_s^{(t)}\, \mathbf{w}^{(t)}_{\text{council}} \;+\; (1 - \eta_s^{(t)})\, \mathbf{w}^{(t-1)},$$

$$\eta_s^{(t)} \;=\; \begin{cases} 0.70 & \text{if } (\mathbf{w}^{(t)}_{\text{council}})_k \geq w_k^{(t-1)} \text{ (building position)}, \\ 0.78 & \text{if } (\mathbf{w}^{(t)}_{\text{council}})_k < w_k^{(t-1)} \text{ (de-risking)}. \end{cases}$$

The higher de-risking speed (+0.08) ensures that defensive signals are acted on promptly, while the lower build speed suppresses day-to-day noise in conviction signals. Smoothing is applied element-wise before the risk control cascade, after which the weight vector is re-normalized to the simplex.

Portfolio Constraints. Each token weight is capped at 0.40 and the cash allocation at 0.30 before the risk overlays. After the full overlay cascade, the weight vector is projected onto the feasible set $\{\mathbf{w} \geq 0,\ \mathbf{1}^\top \mathbf{w} + c = 1,\ w_k \leq 0.40,\ c \leq 0.30\}$ via proportional rescaling. Transaction costs are set to zero in the main evaluation, and the portfolio is rebalanced at the close of each trading day.

Risk Control Parameters. The following parameter values are fixed prior to the evaluation window: BTC dominance bandwidth $\tau_d = 0.15$; transition-buffer depth $c_{\text{trans}} = 0.35$, bandwidth $\tau_{\text{trans}} = 0.25$; drawdown scale $\tau_{\text{dd}} = 0.15$, protection coefficient $\nu_{\text{dd}} = 0.40$. Full overlay cascade specifications are in Appendix E.

E Risk Control Overlay Cascade: Full Specification

The Risk Control layer is organized into four groups matching §4.3, comprising seven overlays applied in a fixed cascade. All overlays are continuous functions of their driving signals with no binary switches, and all parameter values are fixed prior to the evaluation window.

Momentum Overlay. Let $z_k^{(t)}$ denote the cross-sectional 30-day return z-score of asset k. The driving signal is $\tilde{z}_k^{(t)} = \tanh(z_k^{(t)}/1.5)$, a soft-capped version that bounds the tilt magnitude. Each weight is scaled by

$$w_k \cdot \big(1 + \eta(\xi^{(t)}) \cdot \tilde{z}_k^{(t)}\big),$$

where $\eta(\xi) = 0.08 + 0.35\max(0, \xi) \in [0.08,\ 0.43]$.

Dominance Overlay. The driving BTC Dominance signal is the 30-day BTC return minus the EW altcoin basket return, $\Delta^{(t)}_{\text{BTC-EW}}$, compressed via $d^{(t)} = \tanh(\Delta^{(t)}_{\text{BTC-EW}}/\tau_d)$, where $\tau_d = 0.15$. In a BTC season ($d > 0$, volatile or bear regime), weight $d \times c_d$ is shifted from {ETH, ADA, LINK, DOGE, XLM, XRP, BCH} to BTC (60%), TRX (25%), and ZEC (15%), each capped at 0.30; the coefficient is $c_d = 0.45$ in volatile and 0.30 in bear. In an alt season ($d < 0$, bull regime), weight $|d| \times 0.20$ is cut from {BTC, TRX} and redistributed proportionally to {ADA, XLM, DOGE, XRP, BCH, ETH}. There is no activation gate; scaling is continuous in d.

In the volatile regime ($|\xi| < 0.30$), the BTC allocation is lifted to a Floor minimum of 18%:

$$\max(w_{\text{BTC}},\ 0.18).$$

Any shortfall is recovered proportionally from altcoin longs. Empirical motivation: BTC returned $+40.9\%$ versus altcoin EW -8.5% in volatile periods, and 47% of volatile days had $w_{\text{BTC}} < 18\%$ before this floor was introduced.

Intra-Regime Controls. Active only in the bear regime ($\xi < -0.30$). The signal is the mean on-chain z-score differential between BTC and the altcoin basket across two network-activity metrics (active addresses and transaction count): $\delta^{(t)}_{\text{oc}} = \bar{z}_{\text{BTC}}(\text{AdrActCnt},\ \text{TxCnt}) - \bar{z}_{\text{alts}}$. The BTC allocation is increased by

$$\Delta^{\text{oc}}_{\text{BTC}} \ = \ 0.08 \cdot \tanh\big(\delta^{(t)}_{\text{oc}}/1.5\big),$$

provided $\Delta^{\text{oc}}_{\text{BTC}} > 0.005$, with w_{BTC} capped at 0.30. The maximum tilt is 8 pp, and the deficit is taken proportionally from altcoin longs. This control is skipped if A_2 on-chain data are unavailable.

In the volatile regime, the target cash allocation is a smooth function of regime conviction:

$$c_{\text{tgt}} \ = \ 0.08 + 0.17\exp\big(-|\xi^{(t)}|/0.12\big),$$

yielding $c_{\text{tgt}} = 0.25$ at $|\xi| = 0$ (deep volatile) and $c_{\text{tgt}} \approx 0.094$ at $|\xi| = 0.30$ (near the bull/bear boundary). In the bull regime a hard cap of $c_{\text{tgt}} = 0.08$ applies. Adjustment is bidirectional: cash below c_{tgt} is raised by trimming longs, while cash above c_{tgt} is redeployed into longs.

Applied to bull-to-volatile transitions. The downward score drop is $\delta^{(t)}_{\xi} = \max(0,\ \xi^{(t-1)} - \xi^{(t)})$, and the scaling factor is

$$s^{(t)}_{\text{trans}} = 1 - c_{\text{trans}} \cdot \tanh\big(\delta^{(t)}_{\xi}/\tau_{\text{trans}}\big), \tag{23}$$

($c_{\text{trans}} = 0.35$, $\tau_{\text{trans}} = 0.25$), so that $\mathbf{w} \leftarrow s^{(t)}_{\text{trans}} \cdot \mathbf{w}$ with freed weight moved to cash. A mild drop ($\delta_\xi = 0.05$) gives $s^{(t)}_{\text{trans}} \approx 0.93$ (7% de-risk); a severe collapse ($\delta_\xi = 0.60$) gives $s^{(t)}_{\text{trans}} \approx 0.66$ (34% de-risk).

Drawdown Protection. Driven by the current drawdown from the rolling peak, $\text{DD}^{(t)}$, gated by $g(\xi^{(t)}) = \max(0, -\xi^{(t)})$ so that the overlay is suppressed in bull regimes ($g = 0$):

$$s^{(t)}_{\text{dd}} \ = \ 1 - g(\xi^{(t)}) \cdot \tanh\big(\text{DD}^{(t)}/\tau_{\text{dd}}\big) \cdot \nu_{\text{dd}}, \qquad \mathbf{w} \ = \ s^{(t)}_{\text{dd}} \cdot \mathbf{w}.$$

Freed weight goes to cash, with a dynamic cash cap $c^{(t)}_{\text{cap}} = 0.08 + 0.22 \cdot g(\xi^{(t)}) \in [0.08,\ 0.30]$. As a concrete example, with $g(\xi^{(t)}) = 1$ and $\text{DD}^{(t)} = \tau_{\text{dd}} = 15\%$: $s^{(t)}_{\text{dd}} = 1 - \tanh(1) \times \nu_{\text{dd}} \approx 1 - 0.762 \times 0.40 \approx 0.695$, implying a 30.5% de-risking (see Eq. (17)).

Cascade Order and Continuity. The four groups are applied in the listed order, with each group acting on the weight vector as modified by all preceding groups. This ordering reflects an implicit priority structure: the Momentum and Dominance Overlays establish the directional tilt, the Intra-Regime Controls refine the BTC allocation and cash level within each regime, and Drawdown Protection operates as a final risk-budget constraint. All seven overlays use $\tanh$, $\exp$, or clamp operations with no Heaviside functions, ensuring that $\mathbf{w}^{(t)}$ is a Lipschitz-continuous function of the driving signals $(\xi^{(t)}, \{z^{(t)}_k\}, \text{DD}^{(t)}, \delta^{(t)}_{\text{oc}})$. After the full cascade, the weight vector is normalized to $\mathbf{1}^\top \mathbf{w} + c = 1$.

F Proof of EWP Early-Period Influence Attenuation

We provide a formal statement and proof sketch showing that the Exponentially Weighted Protector (EWP) exponentially attenuates the influence of cold-start observations on the coalition characteristic function estimate, in contrast to the EW estimator whose path dependence decays only as $O(1/t)$.

Proposition 1 (EWP Early-Period Influence Attenuation)**.** *Let $\{\hat{\mu}_{\mathrm{EW}}(S)\}$ be the EWP estimate of the coalition mean return (Eq.* (14)*) with e-fold decay period $h > 0$, computed at time t from return history $\{R_S^{(\tau)}\}_{\tau=1}^{t}$. Define the early-period influence ratio as:*

$$\iota_{\mathrm{EW}}(n_0, t) = \frac{\sum_{\tau=1}^{n_0} w_\tau}{\sum_{\tau=1}^{t} w_\tau} \tag{24}$$

where n_0 is the length of the cold-start phase ($n_0 \ll t$). Then:

$$\iota_{\mathrm{EW}}(n_0, t) = \frac{e^{-(t-n_0)/h} - e^{-t/h}}{1 - e^{-t/h}} \;\le\; e^{-(t-n_0)/h} \tag{25}$$

which converges to zero at exponential rate $O(e^{-(t-n_0)/h})$ as $t \to \infty$. In contrast, the EW estimator has early-period influence ratio n_0/t, which converges to zero only at rate $O(1/t)$.

Proof Sketch. Recall the EWP decay weights $w_\tau = e^{-(t-\tau)/h}$ for $\tau = 1, \ldots, t$ (Eq. (13)). The sum of all weights is:

$$\sum_{\tau=1}^{t} w_\tau = \sum_{\tau=1}^{t} e^{-(t-\tau)/h} = \frac{1 - e^{-t/h}}{1 - e^{-1/h}} \;\approx\; h \cdot (1 - e^{-t/h}) \tag{26}$$

for large h (continuous approximation). The contribution of the early periods $\tau \le n_0$ to the total weight is:

$$\sum_{\tau=1}^{n_0} w_\tau = \sum_{\tau=1}^{n_0} e^{-(t-\tau)/h} = e^{-t/h} \sum_{\tau=1}^{n_0} e^{\tau/h} = e^{-t/h} \cdot \frac{e^{(n_0+1)/h} - e^{1/h}}{e^{1/h} - 1} \;\approx\; h \cdot e^{-(t-n_0)/h} \tag{27}$$

Therefore, the influence ratio satisfies:

$$\iota_{\mathrm{EW}}(n_0, t) \approx \frac{h \cdot e^{-(t-n_0)/h}}{h \cdot (1 - e^{-t/h})} = \frac{e^{-(t-n_0)/h}}{1 - e^{-t/h}} \;\le\; 2e^{-(t-n_0)/h} \tag{28}$$

for $t \ge h \ln 2$. This is $O(e^{-(t-n_0)/h})$, which decays exponentially in $(t - n_0)$. At $t = n_0 + 3h$ (three e-fold periods after cold start), the early influence is bounded by $2e^{-3} \approx 10\%$; at $t = n_0 + 5h$, it is below 1.4%.

By contrast, the EW mean estimator assigns weight $1/t$ to every historical observation, giving early-period influence ratio exactly n_0/t. For $n_0 = 100$ cold-start periods, this ratio decays only as $100/t$: it takes $t = 10{,}000$ periods to reduce early influence below 1%, versus $t = n_0 + 5h$ for EWP with any practical h (e.g., $h = 60$ gives $t = 400$ periods).

Applying this result to the EWP estimate $\hat{\mu}_{\mathrm{EW}}(S)$: the absolute perturbation of the characteristic function estimate due to cold-start returns $\{R_S^{(\tau)}\}_{\tau \le n_0}$ is bounded by $2\|R\|_\infty \cdot e^{-(t-n_0)/h}$, which converges to zero exponentially. The raw EWP-based Shapley values therefore converge to their oracle values (computed from the true post-cold-start distribution) at an exponential rate, justifying the early-period influence attenuation claim. □

Remark (Path Dependence). The proposition implies that any random seed, initialization artefact, or unusual market event confined to the first n_0 periods of deployment has negligible influence on the characteristic function after $t \approx n_0 + 3h$ periods. For the MRC configuration with $h = 252$ trading days (one calendar year), the influence of the cold-start phase falls below $2e^{-3} \approx 10\%$ after $t \approx n_0 + 756$ periods and below 1.4% after $t \approx n_0 + 5h = 1{,}260$ periods. In the experimental window of 1,037 periods, cold-start artefacts are therefore substantially suppressed by period 300 onward.

Connection to Shapley Machine and Axiomatic Justification Beyond its path-dependence benefits, the EWP's geometric structure provides an axiomatic grounding for MRC's credit assignment via the Shapley Machine system of Wang et al. [41]. Their central result (Theorem 2 therein) establishes that TD(λ_{TD})-style geometric discounting is equivalent to the Additivity axiom of Shapley values: basis-game decomposition coefficients $k_C = \sum_{T \subseteq C} (-1)^{|C|-|T|} v(T)$, ordered by coalition size, naturally form the geometric sequence $(1-\lambda_{\text{TD}})\lambda_{\text{TD}}^{n-1}$ identifying λ_{TD}-returns. (Here λ_{TD} denotes the TD discount factor from Wang et al. 41, distinct from MRC's Bayesian concentration parameter $\lambda = 30$.)

The EWP weights $w_\tau = e^{-(t-\tau)/h}$ instantiate this geometric structure in continuous time, with the discrete ratio $e^{-1/h}$ playing the role of λ_{TD}. This means the EWP-estimated characteristic functions $v(S)$ (Eq. (14)) are, by construction, consistent with the Additivity axiom. The raw Shapley values themselves continue to rely on the standard four-axiom characterization proved in Appendix G. The remaining axioms are satisfied explicitly:

- **Efficiency:** $\sum_{i=1}^{3} \phi_i = v(\mathcal{N})$ is verified by direct computation in Appendix G.
- **Symmetry:** agents enter Eqs. (29)–(31) symmetrically; permuting agent indices permutes the Shapley values identically.
- **Dummy player:** Appendix G verifies that zero-marginal-contribution agents receive zero raw Shapley value.

Accordingly, by the uniqueness theorem of Shapley [34], the raw Shapley values used in MRC form the unique allocation satisfying all four axioms, providing a cooperative game-theoretic foundation for the credit-assignment layer. The deployable portfolio weights are derived from these scores after truncation and normalization.

Scope Note. The Shapley Machine of Wang et al. [41] operates on RL state-action values in open multi-agent systems, whereas MRC applies the same geometric discounting to financial coalition return histories. The mathematical connection (geometric weights $\leftrightarrow$ Additivity axiom) holds in both settings by the same basis-game argument. Proposition 1 above provides the additional cold-start path-dependence guarantee not addressed in Wang et al. 41, which focuses on stationary cooperative games.

G Three-Player Shapley Value Closed-Form Derivation

We derive the closed-form three-player Shapley values from the general formula and verify that they satisfy the four Shapley axioms.

Derivation via Permutation Enumeration

For $\mathcal{N} = \{1, 2, 3\}$, there are $|\mathcal{N}|! = 6$ orderings (permutations) of the players. The Shapley value of player i is the average marginal contribution of player i over all permutations. For player 1, the six permutations and the coalition S preceding player 1's entry are:

Order	S before 1	$S \cup \{1\}$	Marginal contribution
$(1,2,3)$	$\emptyset$	$\{1\}$	$v_1 - 0 = v_1$
$(1,3,2)$	$\emptyset$	$\{1\}$	v_1
$(2,1,3)$	$\{2\}$	$\{1,2\}$	$v_{12} - v_2$
$(3,1,2)$	$\{3\}$	$\{1,3\}$	$v_{13} - v_3$
$(2,3,1)$	$\{2,3\}$	$\{1,2,3\}$	$v_{123} - v_{23}$
$(3,2,1)$	$\{2,3\}$	$\{1,2,3\}$	$v_{123} - v_{23}$

Averaging over the six permutations (each with probability $1/6$):

$$\begin{aligned}
\phi_1 &= \frac{1}{6}\big[v_1 + v_1 + (v_{12} - v_2) + (v_{13} - v_3) + (v_{123} - v_{23}) + (v_{123} - v_{23})\big] \\
&= \frac{2v_1 + (v_{12} - v_2) + (v_{13} - v_3) + 2(v_{123} - v_{23})}{6} \\
&= \tfrac{1}{3}v_1 + \tfrac{1}{6}(v_{12} - v_2) + \tfrac{1}{6}(v_{13} - v_3) + \tfrac{1}{3}(v_{123} - v_{23}) \qquad (29)
\end{aligned}$$

By symmetry of the enumeration procedure, the formulas for ϕ_2 and ϕ_3 follow with indices permuted:

$$\phi_2 = \tfrac{1}{3}v_2 + \tfrac{1}{6}(v_{12} - v_1) + \tfrac{1}{6}(v_{23} - v_3) + \tfrac{1}{3}(v_{123} - v_{13}) \tag{30}$$

$$\phi_3 = \tfrac{1}{3}v_3 + \tfrac{1}{6}(v_{13} - v_1) + \tfrac{1}{6}(v_{23} - v_2) + \tfrac{1}{3}(v_{123} - v_{12}) \tag{31}$$

which are Eqs. (29)–(31).

An equivalent derivation via the coalition-weighting formula gives the same result: each coalition S of size $|S|$ contributes to player i's Shapley value with weight $|S|!(|\mathcal{N}| - |S| - 1)!/|\mathcal{N}|!$, which for $|\mathcal{N}| = 3$ equals $1/3$ for $|S| = 0$ (empty coalition preceding i's solo entry), $1/6$ for $|S| = 1$ (pairwise coalitions without i), and $1/3$ for $|S| = 2$ (grand coalition entry).

Axiom Verification

We verify that Eqs. (29)–(31) satisfy the four Shapley axioms:

Efficiency.

$$\begin{aligned}\phi_1 + \phi_2 + \phi_3 = {} & \tfrac{1}{3}(v_1 + v_2 + v_3) \\ & + \tfrac{1}{6}\big[(v_{12} - v_2) + (v_{12} - v_1) + (v_{13} - v_3) + (v_{13} - v_1) + (v_{23} - v_3) + (v_{23} - v_2)\big] \\ & + \tfrac{1}{3}\big[(v_{123} - v_{23}) + (v_{123} - v_{13}) + (v_{123} - v_{12})\big]\end{aligned} \tag{32}$$

Collecting terms: the individual terms contribute $\frac{1}{3}(v_1 + v_2 + v_3)$; each pairwise coalition v_{ij} appears with coefficient $\frac{1}{6} + \frac{1}{6} = \frac{1}{3}$ from the second group, and $-\frac{1}{3}$ from the third group (as the complement), canceling exactly; and v_{123} appears with coefficient $3 \times \frac{1}{3} = 1$. Thus $\phi_1 + \phi_2 + \phi_3 = v_{123}$, verifying efficiency.

Note on Efficiency under EWP. The efficiency axiom states $\sum_i \phi_i = v(\mathcal{N})$, where $v(\mathcal{N})$ is the EWP-estimated characteristic function of the grand coalition (Eq. (2)). This is not equal to the current running Sharpe estimate of the grand-coalition portfolio. The two quantities serve different roles: $v(\mathcal{N})$ is a discounted historical evaluator, whereas the current grand-coalition Sharpe is only a contemporaneous performance signal. This distinction is important for interpreting the Shapley values as historical attribution scores, not current-period performance contributions. When negative raw Shapley values are truncated via $\phi_i^+ = \max(\phi_i, 0)$ and weights are renormalized as $\bar{\omega}_i = \phi_i^+ / \sum_j \phi_j^+$ (the truncation-only weight, prior to Bayesian mixing in Eq. (4)), the efficiency axiom $\sum_i \phi_i = v(\mathcal{N})$ no longer holds for $\bar{\omega}_i$ [39, 3]. This is deliberate: raw Shapley values can be negative (when an agent is superfluous), and negative portfolio weights require short-selling, which is not available in MRC's long-only setting. The raw ϕ_i satisfy all four axioms, and the truncation-only weights $\bar{\omega}_i$ preserve the axiomatic ordering without strict efficiency.

Symmetry. If $v(S \cup \{i\}) = v(S \cup \{j\})$ for all $S \not\ni i, j$, then by inspection of Eqs. (29)–(31), interchanging all occurrences of i and j in the formula for ϕ_i yields the formula for ϕ_j, confirming $\phi_i = \phi_j$.

Dummy Player. Suppose player i is a dummy player: $v(S \cup \{i\}) = v(S)$ for all $S \not\ni i$. Then all marginal contributions of i are zero, and by the formula, $\phi_i = 0$.

Additivity. For two cooperative games $(v, \mathcal{N})$ and $(v', \mathcal{N})$ with characteristic functions v and v', the Shapley value of the sum game $(v + v', \mathcal{N})$ is $\phi_i^{v+v'} = \phi_i^v + \phi_i^{v'}$, which follows directly from the linearity of the Shapley formula in v.

Connection to Multi-Stage Architecture

In the MRC architecture, the three Shapley players correspond to A_1 (Market), A_2 (On-chain), and A_3 (Macro). The individual characteristic functions v_1, v_2, v_3 reflect the solo performance of each agent's portfolio, the pairwise values v_{12}, v_{13}, v_{23} reflect the debate-derived joint portfolios, and v_{123} reflects the grand-coalition output. Operationally, this grand-coalition output is instantiated by the Stage 3 readout $\mathbf{w}_{123}$ produced by A_4, which serves as a readout operator rather than as a

fourth player. Superadditivity ($v_{ij} \geq v_i + v_j$) is not assumed or required. In practice, some pairwise debates underperform their individual components during certain regimes, and the Bayesian adaptive weighting (Eq. (4)) accommodates this empirically.

H Additional Results

This section provides supplementary comparisons, sensitivity analyses, and backbone ablations referenced in the main text. Figures 5–8 compare MRC against each baseline group on cumulative returns. Figure 9 shows dynamic blend ratios and agent consensus. Tables 7 and 8 report temperature and burn-in sensitivity. Table 9 reports transaction-cost sensitivity. Tables 10–13 report backbone ablations across all three stages.



Figure 5: MRC vs. passive benchmarks (Single seed (T_{LLM}=0.7))

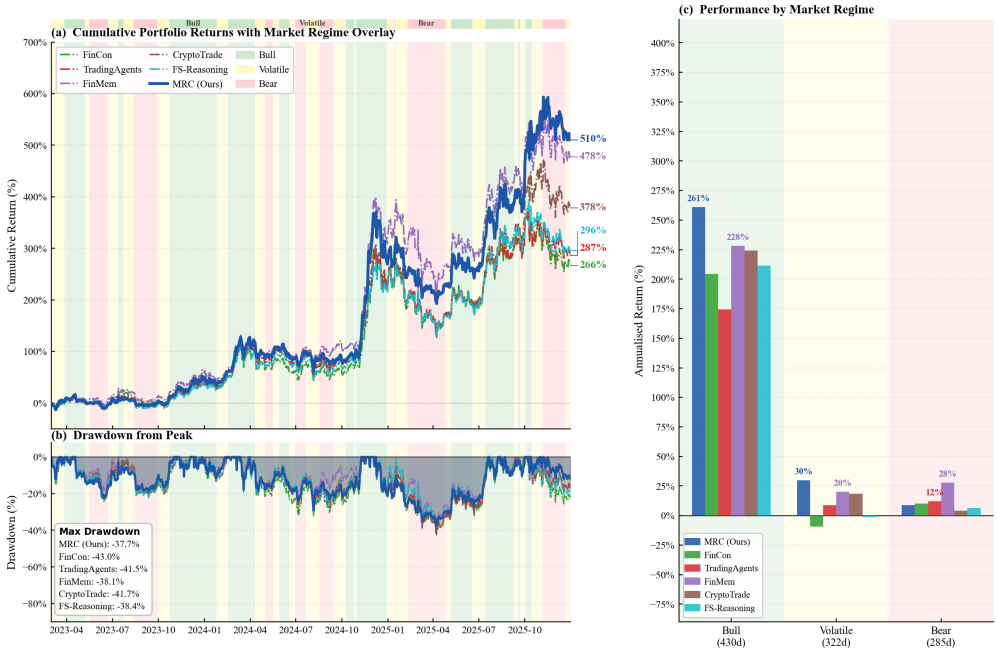

Figure 6: MRC vs. LLM multi-agent baselines (Single seed (T_{LLM}=0.7))

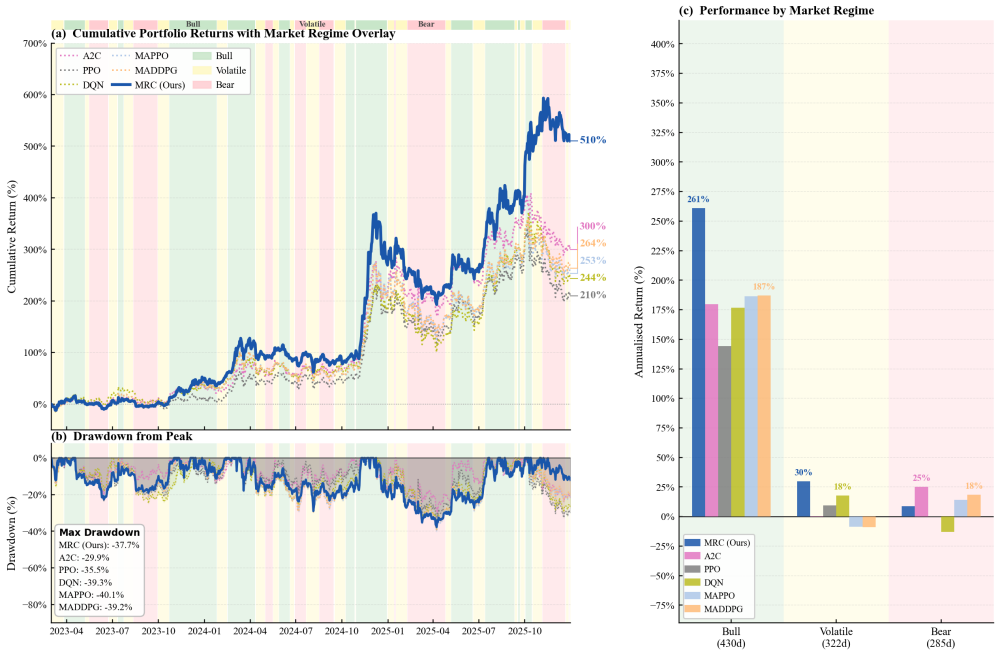

Figure 7: MRC vs. DRL baselines (Single seed (T_{LLM}=0.7))

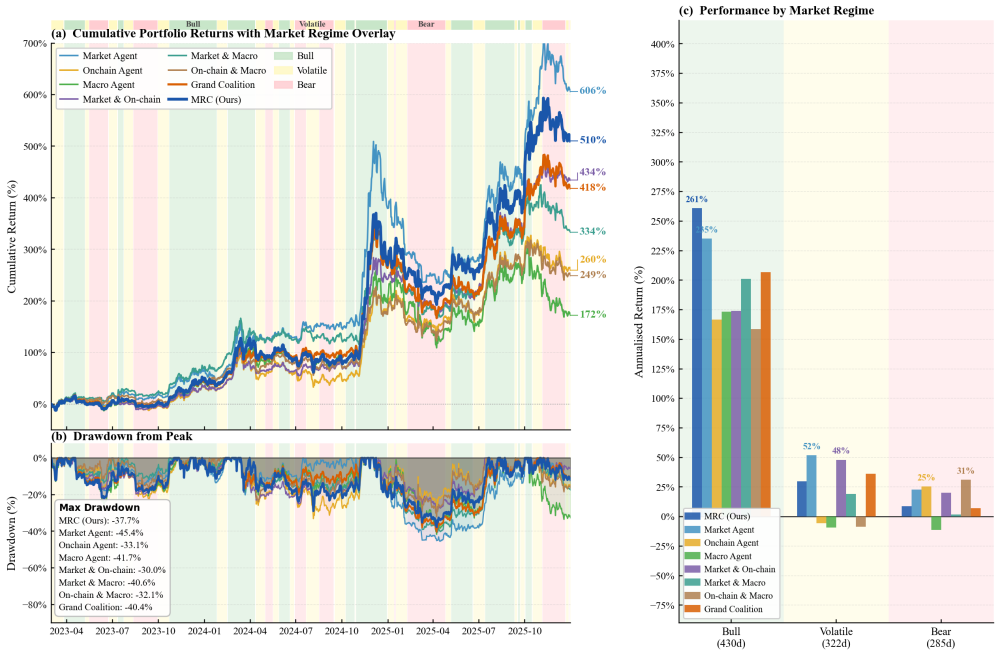

Figure 8: Shapley coalition ablation: individual agents, pairwise coalitions, grand coalition, and full MRC (Single seed (T_{LLM}=0.7))

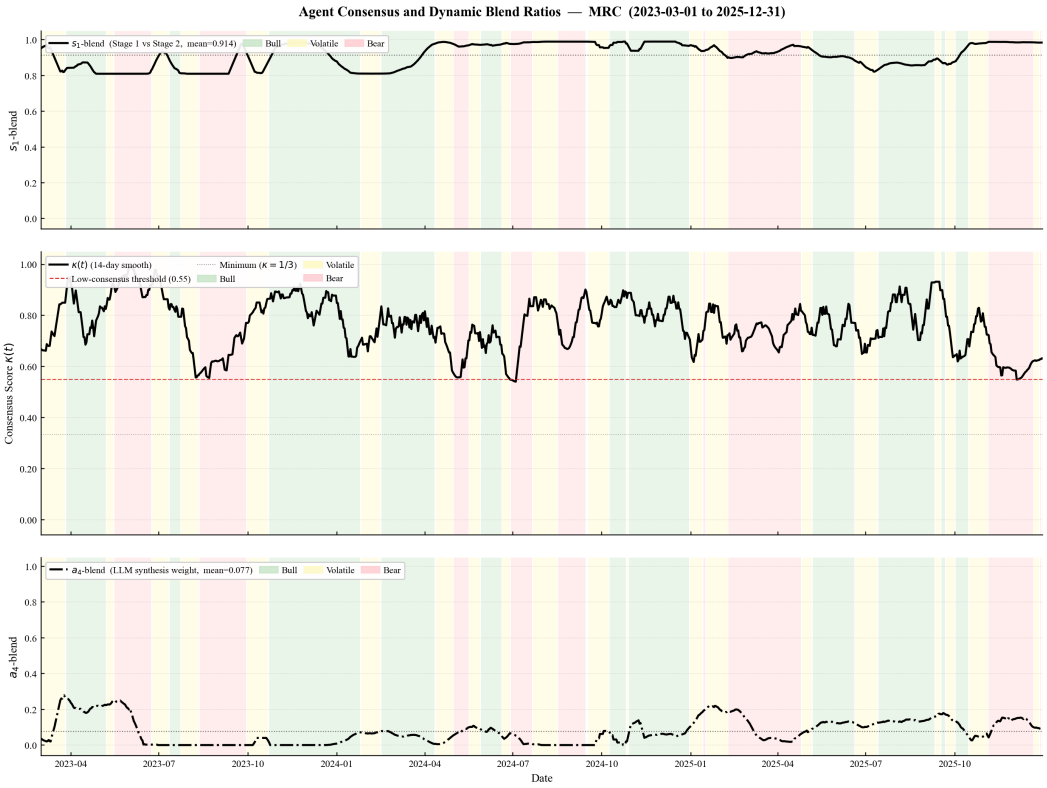

Figure 9: **Dynamic blend ratios and agent consensus (2023-03-01 to 2025-12-31).** Top: Stage-1 blend $\beta_{S1}^{(t)}$ confirms that the Stage-1 ensemble dominates Stage-2 debate outputs across all regimes. Middle: consensus $\kappa^{(t)}$ (dashed line at 0.55 shown for visual reference), where lower consensus continuously reduces the grand-coalition readout blend via Eq. (10). Bottom: grand-coalition readout blend $\beta_{\text{gc}}^{\text{final}}$.

Table 7: Ablation analysis of LLM temperature (T_{LLM}). Primary configuration is T_{LLM}=0.7, and other temperature values are shown for sensitivity. Red: best per column; Blue: second best. Single run per temperature setting.

Method	$T_{\text{LLM}} = 0.1$				$T_{\text{LLM}} = 0.3$				$T_{\text{LLM}} = 0.5$				$T_{\text{LLM}} = 0.7$				$T_{\text{LLM}} = 0.9$			
	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)
Market Benchmarks																				
EW	279.93	1.14	42.78	0.00	279.93	1.14	42.78	0.00	279.93	1.14	42.78	0.00	279.93	1.14	42.78	0.00	279.93	1.14	42.78	0.00
BTC	270.94	1.23	32.02	-0.07	270.94	1.23	32.02	-0.07	270.94	1.23	32.02	-0.07	270.94	1.23	32.02	-0.07	270.94	1.23	32.02	-0.07
ETH	78.45	0.63	63.75	-0.59	78.45	0.63	63.75	-0.59	78.45	0.63	63.75	-0.59	78.45	0.63	63.75	-0.59	78.45	0.63	63.75	-0.59
LLM-based Baselines																				
FinCon	245.50	1.11	40.40	-1.38	257.90	1.14	42.00	-1.03	262.70	1.13	41.80	-0.99	266.50	1.12	43.02	-1.13	286.60	1.17	40.90	0.06
TradingAgents	282.90	1.20	38.20	-0.20	292.20	1.22	39.50	-0.09	251.30	1.14	44.40	-0.47	286.52	1.21	41.55	-0.17	242.00	1.12	42.30	-0.59
FinMem	348.40	1.30	42.50	0.28	345.40	1.31	39.10	0.23	298.40	1.21	39.80	0.00	477.66	1.53	38.06	0.79	319.70	1.28	39.10	0.05
CryptoTrade	304.80	1.24	43.90	0.02	346.70	1.30	43.60	0.30	395.50	1.38	36.10	0.55	378.35	1.36	41.74	0.45	334.40	1.27	43.20	0.24
FSReason	240.60	1.11	40.70	-0.39	284.50	1.21	38.10	-0.12	309.60	1.25	36.60	0.03	295.74	1.23	38.38	-0.14	244.90	1.12	38.30	-0.43
DRL-based Baselines																				
A2C	299.71	1.33	29.88	-0.13	299.71	1.33	29.88	-0.13	299.71	1.33	29.88	-0.13	299.71	1.33	29.88	-0.13	299.71	1.33	29.88	-0.13
PPO	209.69	1.05	35.52	-0.56	209.69	1.05	35.52	-0.56	209.69	1.05	35.52	-0.56	209.69	1.05	35.52	-0.56	209.69	1.05	35.52	-0.56
DQN	243.50	1.18	39.30	-0.47	243.50	1.18	39.30	-0.47	243.50	1.18	39.30	-0.47	243.50	1.18	39.30	-0.47	243.50	1.18	39.30	-0.47
MAPPO	253.42	1.14	40.14	-1.17	253.42	1.14	40.14	-1.17	253.42	1.14	40.14	-1.17	253.42	1.14	40.14	-1.17	253.42	1.14	40.14	-1.17
MADDPG	263.56	1.16	39.22	-0.85	263.56	1.16	39.22	-0.85	263.56	1.16	39.22	-0.85	263.56	1.16	39.22	-0.85	263.56	1.16	39.22	-0.85
MRC (Ours)	381.20	1.44	34.00	0.26	375.40	1.42	32.40	0.23	375.50	1.42	31.30	0.24	510.18	1.60	37.71	0.74	307.10	1.31	34.20	-0.10

Table 8: Burn-in sensitivity analysis. The burn-in period is determined by inverting Eq. (3): $t = \lceil -\lambda \ln(1-\alpha^*) \rceil$, where $\alpha^* \in [0, 1)$ is the target mixing level and $\lambda = 30$. Evaluation windows share the same end date; start dates and sizes vary. Red: best per column; Blue: second best.

Method	$\alpha = 0$ (Full Window)				$\alpha = 0.25$ (Burn-in 29d)				$\alpha = 0.5$ (Burn-in 69d)				$\alpha = 0.75$ (139d)				$\alpha = 0.9$ (230d)			
	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)
Market Benchmarks																				
EW	279.93	1.14	42.78	0.00	265.18	1.14	42.78	0.00	287.60	1.21	42.78	0.00	227.96	1.17	42.78	0.00	272.26	1.35	42.78	0.00
BTC	270.94	1.23	32.02	-0.07	209.18	1.12	32.02	-0.31	216.78	1.18	32.02	-0.36	190.82	1.17	32.02	-0.27	207.53	1.30	32.02	-0.39
ETH	78.45	0.63	63.75	-0.59	65.73	0.60	63.75	-0.64	60.84	0.60	63.75	-0.76	55.49	0.60	63.75	-0.64	85.79	0.75	63.75	-0.62
LLM-based Baselines																				
FinCon	266.50	1.12	43.02	-1.13	251.82	1.12	43.02	-1.19	270.95	1.19	43.02	-1.45	214.21	1.14	43.02	-1.48	257.58	1.32	43.02	-1.48
TradingAgents	286.52	1.21	41.55	-0.17	272.03	1.22	41.55	-0.16	293.19	1.29	41.55	-0.18	249.65	1.27	41.55	0.05	291.35	1.46	41.55	0.00
FinMem	477.66	1.53	38.06	0.79	424.23	1.50	38.06	0.68	452.44	1.58	38.06	0.69	375.77	1.56	38.06	0.80	428.99	1.77	38.06	0.81
CryptoTrade	378.35	1.36	41.74	0.45	357.02	1.36	41.74	0.45	387.77	1.45	41.74	0.48	348.17	1.47	41.74	0.82	389.83	1.65	41.74	0.75
FSReason	295.74	1.23	38.38	-0.14	286.10	1.25	38.38	-0.03	301.79	1.31	38.38	-0.08	281.10	1.34	38.38	0.29	319.34	1.52	38.38	0.22
DRL-based Baselines																				
A2C	299.71	1.33	29.88	-0.13	288.25	1.35	29.88	-0.12	290.01	1.39	29.88	-0.20	255.66	1.39	29.88	-0.07	273.16	1.53	29.88	-0.21
PPO	209.69	1.05	35.52	-0.56	197.41	1.05	35.52	-0.58	201.69	1.09	35.52	-0.69	185.74	1.11	35.52	-0.45	229.62	1.31	35.52	-0.44
DQN	243.54	1.18	39.27	-0.47	215.55	1.14	39.27	-0.59	224.09	1.20	39.27	-0.68	171.15	1.12	39.27	-0.73	228.24	1.38	39.27	-0.60
MAPPO	253.42	1.14	40.14	-1.17	240.81	1.14	40.14	-1.16	259.84	1.21	40.14	-1.24	206.65	1.16	40.14	-1.23	245.17	1.34	40.14	-1.40
MADDPG	263.56	1.16	39.22	-0.85	250.54	1.16	39.22	-0.83	270.36	1.23	39.22	-0.89	215.28	1.18	39.22	-0.86	253.08	1.36	39.22	-1.05
MRC (Ours)	510.18	1.60	37.71	0.74	464.95	1.58	37.71	0.69	509.37	1.69	37.71	0.74	477.10	1.75	37.71	1.10	507.77	1.91	37.71	1.02

Table 9: Slippage sensitivity analysis. Red: best per column; Blue: second best. Slippage applied one-way per unit of portfolio turnover.

Method	0 bps (no cost)				5 bps				10 bps				30 bps			
	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)
Market Benchmarks																
EW	279.93	1.14	42.78	0.00	279.93	1.14	42.78	0.00	279.93	1.14	42.78	0.00	279.93	1.14	42.78	0.00
BTC	270.94	1.23	32.02	-0.07	270.94	1.23	32.02	-0.07	270.94	1.23	32.02	-0.07	270.94	1.23	32.02	-0.07
ETH	78.45	0.63	63.75	-0.59	78.45	0.63	63.75	-0.59	78.45	0.63	63.75	-0.59	78.45	0.63	63.75	-0.59
LLM-based Baselines																
FinCon	266.50	1.12	43.00	-1.13	263.90	1.12	43.10	-1.20	261.30	1.11	43.10	-1.27	251.20	1.10	43.30	-1.52
TradingAgents	286.50	1.21	41.50	-0.17	211.60	1.06	43.10	-0.91	151.20	0.90	44.50	-1.65	6.00	0.29	50.10	-4.60
FinMem	477.70	1.53	38.10	0.79	334.50	1.32	40.30	0.13	226.70	1.11	42.40	-0.54	4.40	0.27	50.30	-3.17
CryptoTrade	378.40	1.36	41.70	0.45	265.10	1.17	44.00	-0.26	178.70	0.98	46.20	-0.96	-5.50	0.21	54.10	-3.76
FSReason	295.70	1.23	38.40	-0.07	198.80	1.03	40.60	-0.78	125.60	0.83	42.70	-1.49	-26.80	0.02	50.50	-4.31
DRL-based Baselines																
A2C	299.70	1.33	29.90	-0.13	136.20	0.91	32.80	-0.92	39.50	0.49	36.50	-1.71	-83.10	-1.20	83.20	-4.86
PPO	209.70	1.05	35.50	-0.56	91.50	0.71	39.00	-1.54	18.40	0.37	42.20	-2.52	-82.80	-1.01	83.40	-6.44
DQN	243.50	1.18	39.30	-0.47	163.40	0.97	40.90	-1.05	102.00	0.77	42.50	-1.63	-30.30	-0.05	48.50	-3.90
MAPPO	253.40	1.14	40.10	-1.17	245.30	1.12	40.30	-1.34	237.30	1.11	40.40	-1.52	207.30	1.04	41.10	-2.21
MADDPG	263.60	1.16	39.20	-0.85	248.40	1.13	39.60	-1.18	233.80	1.10	39.90	-1.51	181.40	0.98	41.30	-2.83
MRC (Ours)	510.20	1.60	37.70	0.74	402.10	1.45	39.30	0.36	313.10	1.30	40.80	-0.02	89.20	0.71	46.50	-1.53

Table 10: LLM backbone ablation — Stage 1: Individual Agents and Full MRC (2023-03-01–2025-12-31, 1 037 days). Each specialist agent operates independently without cross-agent communication. A_1: Market Observer; A_2: On-chain Observer; A_3: Macro Observer. See Table 11 for Stage 2 and Table 12 for Stage 3 results. Red: best per column; Blue: second best.

Backbone	**Stage 1 — Independent Agent Performance**												**Full MRC System**			
	A_1 (Market Observer)				A_2 (On-chain Observer)				A_3 (Macro Observer)				MRC (Full)			
	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)
Market Benchmarks																
EW							*n/a*						279.9	1.14	42.8	0.00
BTC							*n/a*						270.9	1.23	32.0	−0.07
ETH							*n/a*						78.5	0.63	63.8	-0.59
LLM Backbone (Qwen3-VL variants)																
`qwen3-vl-8b-instruct`	271.4	1.42	28.8	−0.29	249.0	1.21	31.1	−0.55	195.7	1.02	40.2	−0.97	336.8	1.42	30.3	−0.01
`qwen3-vl-8b-thinking`	256.9	1.27	41.3	−0.35	215.0	1.15	36.5	−0.66	188.2	1.02	39.9	−0.98	299.1	1.30	36.4	−0.17
`qwen3-vl-32b-instruct`	343.3	1.44	45.9	0.09	286.3	1.43	35.2	−0.15	189.6	1.03	41.6	−0.77	387.4	1.45	36.9	0.25
`qwen3-vl-30b-a3b-instruct`	462.3	1.48	47.5	0.25	269.8	1.22	39.7	−0.73	195.2	1.01	42.1	−1.51	418.2	1.48	39.1	0.46
`qwen3-vl-30b-a3b-thinking`	589.0	1.60	48.5	0.66	332.5	1.36	33.9	0.00	156.9	0.96	33.7	−1.00	521.6	1.57	38.0	0.69
`qwen3-vl-235b-a22b-instruct`	341.1	1.36	48.4	0.23	368.5	1.49	31.1	0.43	189.3	1.04	41.4	−0.53	444.3	1.52	39.6	0.51
`qwen3-vl-235b-a22b-thinking`	451.0	1.52	43.9	0.22	341.7	1.52	33.3	−0.21	173.0	1.01	38.9	−1.18	431.9	1.49	37.5	0.44

Table 11: LLM backbone ablation — Stage 2: Pairwise Coalitions and Full MRC (2023-03-01–2025-12-31, 1 037 days). P_{12}: Market⊕On-chain Socratic debate; P_{13}: Market⊕Macro debate; P_{23}: On-chain⊕Macro debate. Market benchmarks have no pairwise-coalition counterpart (*n/a*). See Table 10 for Stage 1 and Table 12 for Stage 3 results. Red: best per column; Blue: second best.

Backbone	Stage 2 — Pairwise Coalitions												Full MRC System			
	P_{12} (Mkt.⊕Chain)				P_{13} (Mkt.⊕Macro)				P_{23} (Chain⊕Macro)				MRC (Full)			
	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)
Market Benchmarks																
EW							*n/a*						279.9	1.14	42.8	0.00
BTC							*n/a*						270.9	1.23	32.0	−0.07
ETH							*n/a*						78.5	0.63	63.8	-0.59
LLM Backbone (Qwen3-VL variants)																
`qwen3-vl-8b-instruct`	300.2	1.62	20.0	−0.28	218.7	1.37	28.1	−0.48	210.8	1.16	29.5	−0.79	336.8	1.42	30.3	−0.01
`qwen3-vl-8b-thinking`	174.5	1.16	38.9	−0.79	172.8	1.12	38.9	−0.83	165.8	1.08	34.1	−0.93	299.1	1.30	36.4	−0.17
`qwen3-vl-32b-instruct`	231.0	1.33	42.0	−0.31	276.4	1.32	47.2	−0.08	295.8	1.44	31.4	−0.11	387.4	1.45	36.9	0.25
`qwen3-vl-30b-a3b-instruct`	335.5	1.42	39.5	−0.33	363.2	1.42	39.1	−0.20	197.0	1.07	38.4	−1.29	418.2	1.48	39.1	0.46
`qwen3-vl-30b-a3b-thinking`	189.7	1.13	44.6	−0.48	227.5	1.25	37.4	−0.44	180.2	1.15	31.5	−0.72	521.6	1.57	38.0	0.69
`qwen3-vl-235b-a22b-instruct`	347.2	1.57	39.3	0.13	272.6	1.38	42.3	−0.06	291.3	1.38	33.6	0.03	444.3	1.52	39.6	0.51
`qwen3-vl-235b-a22b-thinking`	193.2	1.30	38.0	−0.67	285.1	1.51	38.5	−0.34	269.3	1.50	28.1	−0.47	431.9	1.49	37.5	0.44

Table 12: LLM backbone ablation — Stage 3: Grand Coalition and Full MRC (2023-03-01–2025-12-31, 1 037 days). The Stage 3 grand-coalition readout $\mathbf{w}_{123}$ is produced by A_4 from the full coalition context. Market benchmarks have no grand-coalition counterpart (*n/a*). See Table 10 for Stage 1 and Table 11 for Stage 2 results. Red: best per column; Blue: second best.

Backbone	Stage 3				Full MRC System			
	Grand Coalition Readout ($\mathbf{w}_{123}$)				MRC (Full)			
	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)	CR % (↑)	SR (↑)	MDD % (↓)	IR (↑)
Market Benchmarks								
EW			*n/a*		279.9	1.14	42.8	0.00
BTC			*n/a*		270.9	1.23	32.0	−0.07
ETH			*n/a*		78.5	0.63	63.8	-0.59
LLM Backbone (Qwen3-VL variants)								
`qwen3-vl-8b-instruct`	201.1	1.18	31.3	−0.71	336.8	1.42	30.3	−0.01
`qwen3-vl-8b-thinking`	201.2	1.12	40.2	−0.85	299.1	1.30	36.4	−0.17
`qwen3-vl-32b-instruct`	248.9	1.30	39.4	−0.28	387.4	1.45	36.9	0.25
`qwen3-vl-30b-a3b-instruct`	349.6	1.34	44.1	−0.23	418.2	1.48	39.1	0.46
`qwen3-vl-30b-a3b-thinking`	384.8	1.43	43.0	0.22	521.6	1.57	38.0	0.69
`qwen3-vl-235b-a22b-instruct`	299.1	1.34	40.2	0.11	444.3	1.52	39.6	0.51
`qwen3-vl-235b-a22b-thinking`	303.8	1.38	38.4	−0.33	431.9	1.49	37.5	0.44

Table 13: LLM backbone ablation — Inference cost and parse failure rate (2023-03-01–2025-12-31, 1 037 days). **Parse failure** occurs when an LLM response cannot be decoded into a valid portfolio weight vector; the system falls back to the previous period's weights for that step. Rates are computed over all 1 037 rebalancing steps. The end-to-end Full MRC failure rate (all three stages fail simultaneously) is 0.10% uniformly across all backbones. **Inference time** reports mean ± std. per individual LLM call: Stage 1 = one specialist-agent call; Stage 2 = one pairwise-debate turn; Stage 3 = grand-coalition synthesis call; Full MRC = total wall-clock time per rebalancing period. Red: best (lowest) per column; Blue: second best. ↓ lower is better. See Tables 10–12 for portfolio performance metrics.

Backbone	Parse Fail. Rate (%) (↓)	Mean Inference Time per LLM Call (s) (↓)			
		Stage 1 (Agent)	Stage 2 (Debate)	Stage 3 (GC)	Full MRC
`qwen3-vl-8b-instruct`	0.39	$24.5_{\pm 8.2}$	$12.0_{\pm 4.0}$	$5.8_{\pm 2.4}$	$42.5_{\pm 10.4}$
`qwen3-vl-8b-thinking`	0.19	$14.6_{\pm 2.3}$	$7.1_{\pm 1.8}$	$3.5_{\pm 1.8}$	$25.5_{\pm 3.6}$
`qwen3-vl-32b-instruct`	2.80	$22.0_{\pm 2.1}$	$11.7_{\pm 6.2}$	$6.1_{\pm 2.8}$	$40.1_{\pm 7.2}$
`qwen3-vl-30b-a3b-instruct`	0.96	$26.9_{\pm 8.7}$	$12.9_{\pm 4.5}$	$6.2_{\pm 3.9}$	$46.3_{\pm 11.3}$
`qwen3-vl-30b-a3b-thinking`	0.68	$135.5_{\pm 38.1}$	$258.6_{\pm 105.4}$	$63.8_{\pm 42.4}$	$458.1_{\pm 130.4}$
`qwen3-vl-235b-a22b-instruct`	4.53	$42.2_{\pm 18.9}$	$22.2_{\pm 11.2}$	$10.5_{\pm 7.2}$	$75.1_{\pm 26.9}$
`qwen3-vl-235b-a22b-thinking`	3.47	$42.2_{\pm 15.6}$	$21.1_{\pm 8.1}$	$9.6_{\pm 6.2}$	$73.1_{\pm 19.8}$

I Source of the Gain and Robustness

Tables 14, 15 and 16 come from an offline replay of the five logged runs. Each arm reuses the logged portfolios of the seven coalitions and changes only the parts named in its row. The grand-coalition portfolio enters every arm as logged, after its blend with the Shapley-weighted ensemble (Eq. (7)), at a fixed weight of 0.10, which keeps the LLM readout identical across arms. MRC scores SR 1.512 in the replay against 1.51 in Table 2. Differences are paired by seed and computed before rounding. The Seeds columns give the number of the five seeds that move in the stated direction, and $\pm$ marks the standard deviation over seeds.

Aggregation Rules and Leave-One-Out. Table 14 separates the credit mechanism (Shapley credit, the Bayesian mixture, the Regime-Aware Multiplier and Selective WTA) from the risk-control overlays for RQ2 in Section 5. The left side changes only the aggregation rule over the same coalition portfolios, with the overlays on or off. Its worst window is the lowest SR over 190 rolling 252-day windows (38 per seed), and its 10 bps cost also counts changes in the cash weight. The right side removes one part of the credit mechanism or of the risk-control overlays at a time and gives the change in SR at 0 bps. With the overlays on, the credit mechanism reaches SR 1.512 against 1.374 for uniform weighting and beats it in all five seeds ($+0.137$, first row of Table 15). It is ahead of the three alternative rules with and without the overlays, at a 10 bps trading cost and on the worst rolling 252-day window. Removing any part listed on the right lowers SR, the cash target most of all.

Table 14: Source of the gain (offline replay, 5 seeds). Left: aggregation rules on the same coalition portfolios. Red: best per column; Blue: second best. Right: SR change when one part is removed from MRC.

	SR, overlays		SR at	Worst
Aggregation rule	on	off	10 bps	window
Uniform $1/N$	1.374	1.258	1.156	0.574
Inverse volatility	1.393	1.276	1.170	0.572
Hedge [9]	1.455	1.306	1.168	0.412
MRC (Ours)	1.512	1.321	1.235	0.608

	Change at 0 bps	
Removed from MRC	ΔSR	Seeds
Credit mechanism		
Shapley credit in the mixture	-0.057	4/5
Selective WTA	-0.041	5/5
Regime-Aware Multiplier	-0.022	4/5
Risk-control overlays		
Cash target	-0.157	5/5
Momentum Overlay	-0.063	5/5
Transition buffer	-0.005	5/5

Constraint Sensitivity. Table 15 adds caps after the overlays. A final projection caps each asset weight at $w_{\max}$ and the cash weight at $c_{\max}$, and weight above a cap moves to the other positions in proportion to their weights. At zero cost the credit mechanism beats uniform weighting under every cap in all five seeds, by $+0.101$ to $+0.137$. At 10 bps its lead is $+0.047$ to $+0.088$ and holds in at least four of five seeds.

Table 15: Constraint sensitivity analysis (offline replay, 5 seeds, mean $\pm$ std). ΔSR: MRC minus uniform $1/N$ under the same caps.

	SR ($\uparrow$)		ΔSR at 0 bps		ΔSR at 10 bps	
Feasible set	MRC	Uniform	Mean	Seeds	Mean	Seeds
As run	1.512	1.374	$+0.137_{\pm 0.029}$	5/5	$+0.079$	5/5
$w_{\max} = 0.25$	1.510	1.377	$+0.133_{\pm 0.028}$	5/5	$+0.081$	5/5
$w_{\max} = 0.20$	1.514	1.377	$+0.137_{\pm 0.027}$	5/5	$+0.088$	5/5
$w_{\max} = 0.15$	1.479	1.374	$+0.106_{\pm 0.048}$	5/5	$+0.064$	4/5
$c_{\max} = 0.20$	1.468	1.362	$+0.106_{\pm 0.027}$	5/5	$+0.049$	4/5
$c_{\max} = 0.10$	1.436	1.330	$+0.106_{\pm 0.026}$	5/5	$+0.047$	4/5
$w_{\max} = 0.20$, $c_{\max} = 0.10$	1.433	1.332	$+0.101_{\pm 0.016}$	5/5	$+0.058$	5/5

Sensitivity to the Constants. Table 16 perturbs 25 scalar constants of the credit mechanism and blending, the smoothing and overlays, and the regime definition. One at a time, each constant is

multiplied by 0.7, 0.8, 1.2 and 1.3. The table gives the lowest and highest SR over these four settings (mean over seeds) and the worst change against the reported configuration (SR 1.512), paired by seed. No single constant costs more than 0.031 in SR. Jointly, all 25 constants are multiplied by independent draws from $U(0.8, 1.2)$ (200 draws, last row of the table). SR then ranges from 1.465 to 1.589, above Equal-Weight, BTC and CryptoTrade (Table 2) in every draw. The response is broad rather than peaked at the reported values, which is consistent with constants that were not tuned on the evaluation window.

Table 16: Sensitivity analysis of 25 scalar constants (offline replay, 5 seeds).

	SR range		Worst ΔSR	
Constant	Min	Max	Mean	Seeds
Credit mechanism and blending				
WTA rolling window n_{win}	1.481	1.552	-0.031	5/5
WTA dominance threshold θ_{WTA}	1.488	1.527	-0.024	5/5
WTA weight ω_{wta}	1.491	1.541	-0.021	5/5
Characteristic-function weight γ_ρ	1.507	1.515	-0.005	5/5
Regime multiplier spread (scale on $\psi_i - 1$)	1.508	1.518	-0.004	4/5
EWP e-fold decay h	1.510	1.518	-0.002	3/5
Bayesian concentration λ	1.511	1.512	-0.001	3/5
Stage 1/2 blend bandwidth τ_{S1}	1.512	1.512	0.000	4/5
Grand-coalition weight in the replay	1.512	1.512	0.000	3/5
Smoothing and overlays				
EMA smoothing rate	1.486	1.524	-0.026	5/5
Momentum Overlay slope	1.502	1.522	-0.010	5/5
BTC floor	1.504	1.528	-0.008	3/5
Cash target floor	1.505	1.520	-0.007	5/5
Momentum Overlay base	1.505	1.519	-0.007	5/5
Per-asset cap in the Dominance Overlay	1.507	1.512	-0.005	4/5
Transition buffer depth	1.510	1.514	-0.002	5/5
Drawdown bandwidth τ_{dd}	1.510	1.514	-0.002	5/5
Drawdown reduction ν_{dd}	1.511	1.513	-0.001	4/5
Cash target bandwidth	1.511	1.513	-0.001	5/5
Transition buffer bandwidth	1.511	1.514	-0.001	5/5
Cash target amplitude	1.511	1.513	0.000	4/5
Dominance coefficient in volatile regimes	1.512	1.512	0.000	3/5
Dominance bandwidth τ_d	1.512	1.513	0.000	3/5
Regime definition				
Bear threshold $\lvert\xi_-\rvert$	1.507	1.509	-0.005	4/5
Bull threshold ξ_+	1.519	1.540	$+0.007$	4/5
All 25 constants jointly				
200 draws from $U(0.8, 1.2)$	1.465	1.589		

Rule Council. To test whether the credit mechanism needs LLM specialists, we replace the three LLM agents by three rule specialists that read one representative signal per data family: cross-sectional momentum for A_1, on-chain activity growth for A_2, and a risk-appetite score from VIX and the Fear & Greed Index for A_3.

They emit weight vectors of the same form, fill the same seven coalitions and pass through the same credit mechanism and overlays. Every rule constant is fitted on the 789 days that end on 2023-02-28, before the evaluation window, and then frozen, and the softmax temperatures are set so that each rule specialist's median portfolio breadth matches its LLM counterpart. Switching from uniform weighting to the full credit mechanism is worth $+0.065$ in SR on the rule specialists and $+0.137$ on the LLM specialists, so the mechanism pays on both sides and more on the LLM side.

Regimes and Agent Weights over Five Seeds. Table 17 repeats the regime decomposition of Table 4 for all 14 methods of Table 2, each over its five runs. The regime labels are the same in every run. Mean rank is the mean SR rank over the three regimes among the 14 methods. Over the same

five seeds, the mean regime-adjusted weights $\tilde{\omega}_i^{(t)}$ of A_1, A_2 and A_3 are 0.34 ± 0.10, 0.24 ± 0.09 and 0.42 ± 0.16 in the first 100 periods and 0.50 ± 0.07, 0.35 ± 0.07 and 0.15 ± 0.05 after day 100.

Table 17: Regime-conditional SR over five seeds. Bull / Bear / Volatile identified by $\xi^{(t)}$. Non-benchmark methods: mean $\pm$ std. Benchmarks: deterministic. Red: best mean per column among non-benchmark methods; Blue: second best.

	SR (↑)			Mean rank
Method	Bull	Bear	Volatile	(↓)
Market Benchmarks				
EW	2.26	0.50	0.04	7.33
BTC	1.77	0.57	1.11	5.67
ETH	2.08	−0.79	−0.19	13.33
DRL-based Baselines				
A2C	$2.13_{\pm0.30}$	$0.23_{\pm0.49}$	$0.06_{\pm0.11}$	11.00
PPO	$2.08_{\pm0.21}$	$0.29_{\pm0.14}$	$0.17_{\pm0.24}$	10.67
DQN	$2.31_{\pm0.28}$	$0.30_{\pm0.43}$	$0.32_{\pm0.21}$	7.33
MAPPO	$2.26_{\pm0.01}$	$0.51_{\pm0.01}$	$0.03_{\pm0.01}$	8.00
MADDPG	$2.25_{\pm0.04}$	$0.50_{\pm0.05}$	$0.03_{\pm0.03}$	9.33
LLM-based Baselines				
FinCon	$2.27_{\pm0.04}$	$0.46_{\pm0.03}$	$0.03_{\pm0.02}$	8.00
TradingAgents	$2.34_{\pm0.09}$	$0.40_{\pm0.18}$	$0.24_{\pm0.23}$	6.00
FinMem	$2.38_{\pm0.26}$	$0.36_{\pm0.26}$	$0.40_{\pm0.64}$	5.67
CryptoTrade	$2.46_{\pm0.09}$	$0.37_{\pm0.25}$	$0.44_{\pm0.16}$	4.33
FSReason	$2.42_{\pm0.11}$	$0.32_{\pm0.13}$	$0.12_{\pm0.12}$	7.00
MRC (Ours)	$2.55_{\pm0.05}$	$0.63_{\pm0.27}$	$0.45_{\pm0.22}$	1.33

Turnover and Cost. Table 18 reports turnover and costed SR for the same 14 methods. Turnover is the yearly sum of $|\Delta w|$ over the 13 risky weights. Cost is charged one way per unit of turnover, and 10 bps is the reference rate. Unlike Table 9, which shows a single run, this table averages the five logged runs and charges Equal-Weight for its daily rebalancing. Break-even is the cost at which SR falls below Equal-Weight charged the same rate. A break-even of 0 means that SR is not above Equal-Weight even at zero cost.

Scalability. Table 19 separates the two costs of exact credit. Exact Shapley values need all 2^N-1 coalition portfolios per period, and the LLM cost per period grows in proportion to this count. The Shapley arithmetic itself, measured as wall-clock time per period, stays below 0.05 s up to $N=12$. Sampling error is the largest difference between sampled and exact ϕ_i, relative to $\max_i |\phi_i|$, when permutation sampling [4] with 500 permutations replaces exact values. It is measured on 20 games per N whose coalition values follow our logged runs and stays below 0.10 up to $N=12$. Exact values are therefore practical up to about $N=5$ (31 portfolios per period).

Table 18: Turnover and transaction cost. Non-benchmark methods: mean $\pm$ std (5 seeds). Benchmarks: deterministic. Red: best SR and break-even among non-benchmark methods; Blue: second best.

	Turnover	SR (↑)		Break-even
Method	per year	0 bps	10 bps	bps (↑)
Market Benchmarks				
EW (charged)	6.2	1.14	1.13	*n/a*
BTC	0.0	1.23	1.23	*n/a*
ETH	0.0	0.63	0.63	*n/a*
DRL-based Baselines				
A2C	$386.3_{\pm 10.4}$	$1.03_{\pm 0.18}$	$0.22_{\pm 0.17}$	0.0
PPO	$344.4_{\pm 3.9}$	$1.04_{\pm 0.13}$	$0.36_{\pm 0.13}$	0.0
DQN	$210.5_{\pm 28.4}$	$1.16_{\pm 0.18}$	$0.71_{\pm 0.19}$	0.3
MAPPO	$17.0_{\pm 0.5}$	$1.14_{\pm 0.01}$	$1.11_{\pm 0.01}$	0.0
MADDPG	$28.5_{\pm 2.1}$	$1.14_{\pm 0.03}$	$1.08_{\pm 0.03}$	0.0
LLM-based Baselines				
FinCon	$14.3_{\pm 2.4}$	$1.14_{\pm 0.01}$	$1.11_{\pm 0.01}$	0.0
TradingAgents	$147.2_{\pm 4.7}$	$1.19_{\pm 0.03}$	$0.89_{\pm 0.02}$	1.6
FinMem	$192.8_{\pm 7.5}$	$1.25_{\pm 0.20}$	$0.87_{\pm 0.19}$	3.0
CryptoTrade	$181.4_{\pm 12.2}$	$1.31_{\pm 0.06}$	$0.95_{\pm 0.05}$	4.8
FSReason	$188.6_{\pm 12.7}$	$1.18_{\pm 0.06}$	$0.80_{\pm 0.06}$	1.1
MRC (Ours)	$128.7_{\pm 5.5}$	$1.51_{\pm 0.07}$	$1.23_{\pm 0.06}$	13.8

Table 19: Cost and error of credit assignment as the number of agents N grows.

	Exact Shapley values		Sampling
N	Coalitions (2^N-1)	Time per period (s)	error (↓)
3 (MRC)	7	<0.001	0.04
4	15	<0.001	0.06
5	31	<0.001	0.07
6	63	<0.001	0.07
8	255	0.001	0.08
10	1,023	0.008	0.09
12	4,095	0.047	0.09

J Discussion

Why Debate Underperforms. The most striking finding from the coalition ablation is that Stage 2 Socratic debate consistently reduces Sharpe ratio relative to the best individual specialist, a result that stands in apparent tension with evidence that structured LLM debate improves reasoning quality [8]. The resolution lies in a distinction between reasoning tasks and portfolio construction tasks. In open-ended reasoning, debate surfaces inconsistencies between agents' reasoning over a shared evidence base. In portfolio construction, the three specialist agents process fundamentally different data modalities across separate information channels, and when their views diverge, the divergence is often epistemically valid rather than a correctable error. Forcing pairwise reconciliation attenuates signal rather than eliminating noise. Within MRC, Stage 2 serves a different purpose: it generates structured disagreement signals that modulate β_{gc} via the divergence discount (Eq. 10), reducing reliance on the grand-coalition readout precisely when inter-agent consensus is low. The full system recovers the efficient frontier not by improving debate quality, but by Shapley-weighting all coalition outputs so that empirically inferior coalitions contribute proportionally less to the final allocation.

Mathematical Integration versus LLM synthesis. A consistent pattern emerges across all ablation configurations: Shapley-weighted mathematical integration outperforms LLM synthesis at every stage. The dynamic blend ratios $\beta_{S1}^{(t)}$ and $\beta_{\mathrm{gc}}^{(t)}$ reflect this empirically, with $\beta_{\mathrm{gc}}^{(t)}$ remaining near its lower bound throughout most of the evaluation window. This finding has a broader implication for multi-agent LLM systems in financial applications: the primary value of LLM agents may lie not in their capacity to aggregate information better than quantitative rules, but in their ability to translate heterogeneous raw inputs (candlestick images, on-chain tables, macroeconomic time series) into a common representational format amenable to principled mathematical integration. The Shapley system provides the bridge between these two layers: it quantifies each modality's marginal contribution and weights the final allocation accordingly, separating the question of which signals matter from the question of how to combine them.

Regime-conditional Performance Pattern. Over five seeds (Table 17), MRC keeps a positive annualised Sharpe in all three market conditions (bull, volatile, and bear) and ranks first in bull and bear markets among the 14 methods, with the advantage over passive benchmarks most pronounced in bull markets.

Convergence and Regret Analysis. The Exponentially Weighted Characteristic Value (EW-CV) characteristic function $v(S;\mathcal{H}_t)$ with e-fold decay period h=252 (discrete decay ratio $e^{-1/h}$, distinct from the Bayesian concentration parameter $\lambda = 30$ in Eq. (3)) connects MRC to discounted online learning [57], bounding the effective memory to approximately h recent trading periods regardless of the evaluation horizon T.

The Bayesian adaptive mixture $\alpha(t) = 1 - \exp(-t/\lambda)$ with λ=30 converges from the uniform prior to the Shapley-dominated distribution at an exponential rate, yielding a warm-up period of $\lceil \lambda \ln(2/\varepsilon) \rceil$ periods to reach ε-proximity. For λ=30 and ε=0.05, this is approximately 110 trading days, close to the 100-day cold-start window used in RQ4. This connection motivates the use of exponential discounting and the Bayesian warm-up schedule in MRC. We do not claim a formal regret bound for the full deployed system, since regime multipliers, blend rules, and overlay layers introduce additional nonlinearities beyond the standard expert-aggregation setting.

Cold-start Correctness and EWP behavior. Over five seeds (Appendix I) no agent dominates the first 100 periods (mean regime-adjusted weights $\tilde{\omega}_i^{(t)}$ of 0.34, 0.24 and 0.42 for A_1, A_2 and A_3), and the weights then diverge as Shapley evidence accumulates, averaging 0.50, 0.35 and 0.15 after day 100. This validates the joint design of the EWP decay and the Bayesian adaptive mixture. The decline of A_3 (Macro Agent) to a mean weight of 0.15 is consistent with theoretical prediction: in a risk-on, momentum-driven environment, macroeconomic indicators have genuinely low marginal contribution to portfolio returns, and the EWP correctly discounts early periods in which macro signals appeared more relevant. This weight is expected to recover in a sustained macro-driven bear regime, and the geometric discounting ensures the recovery is continuous rather than abrupt. The three-phase learning schedule operates without any external reset mechanism, a property practically important in live deployment, where retraining is infeasible.

K Individual Token, Strategy Return Profiles, and Additional Details of Agent Settings

Figures 10 and 11 provide full per-asset and per-strategy visualisations over the entire backtest window (2023-03-01 to 2025-12-31, $T = 1{,}037$ trading days).

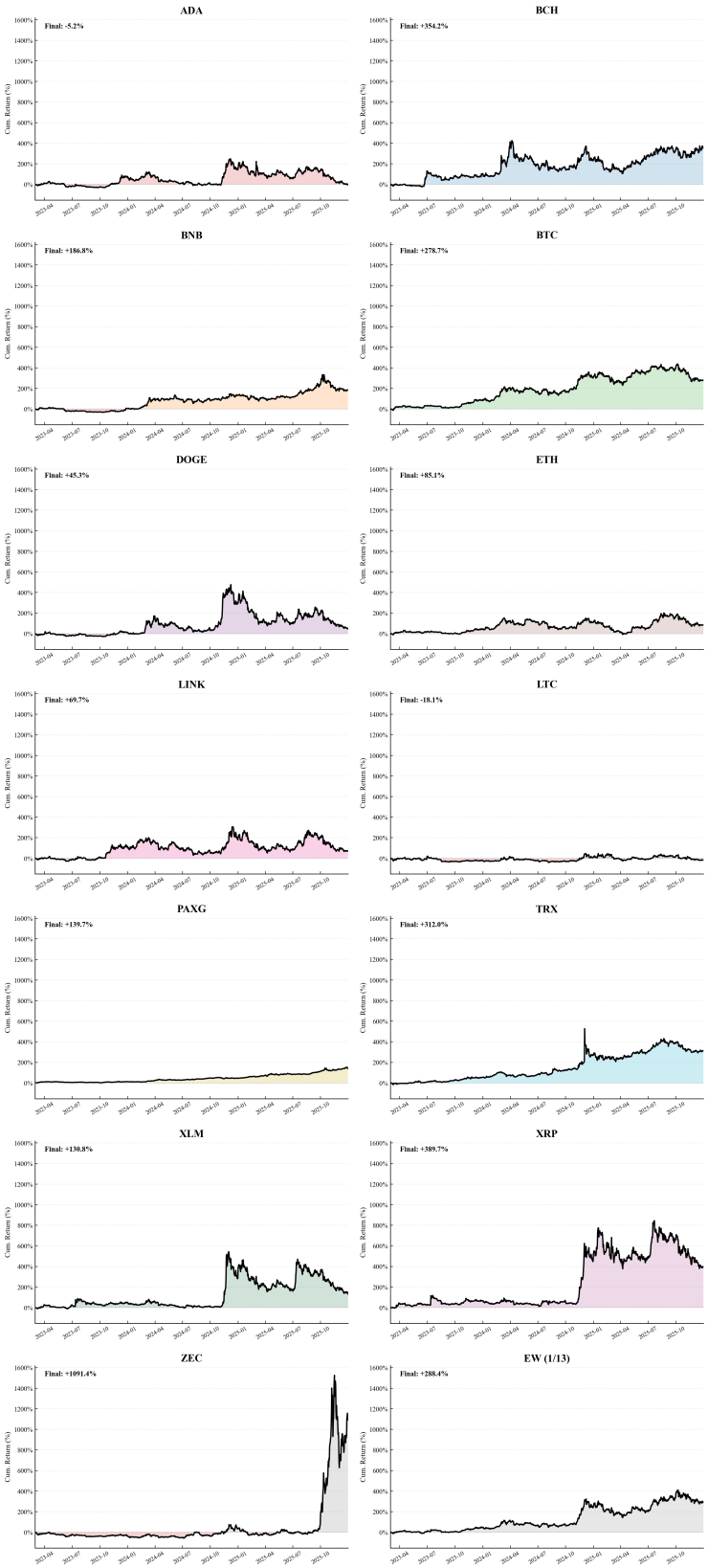

Figure 10: **Cumulative Hold returns of the 13 tokens and EW benchmark (2023-03-01 to 2025-12-31).** Each panel: single-asset cumulative return; final value annotated top-left.

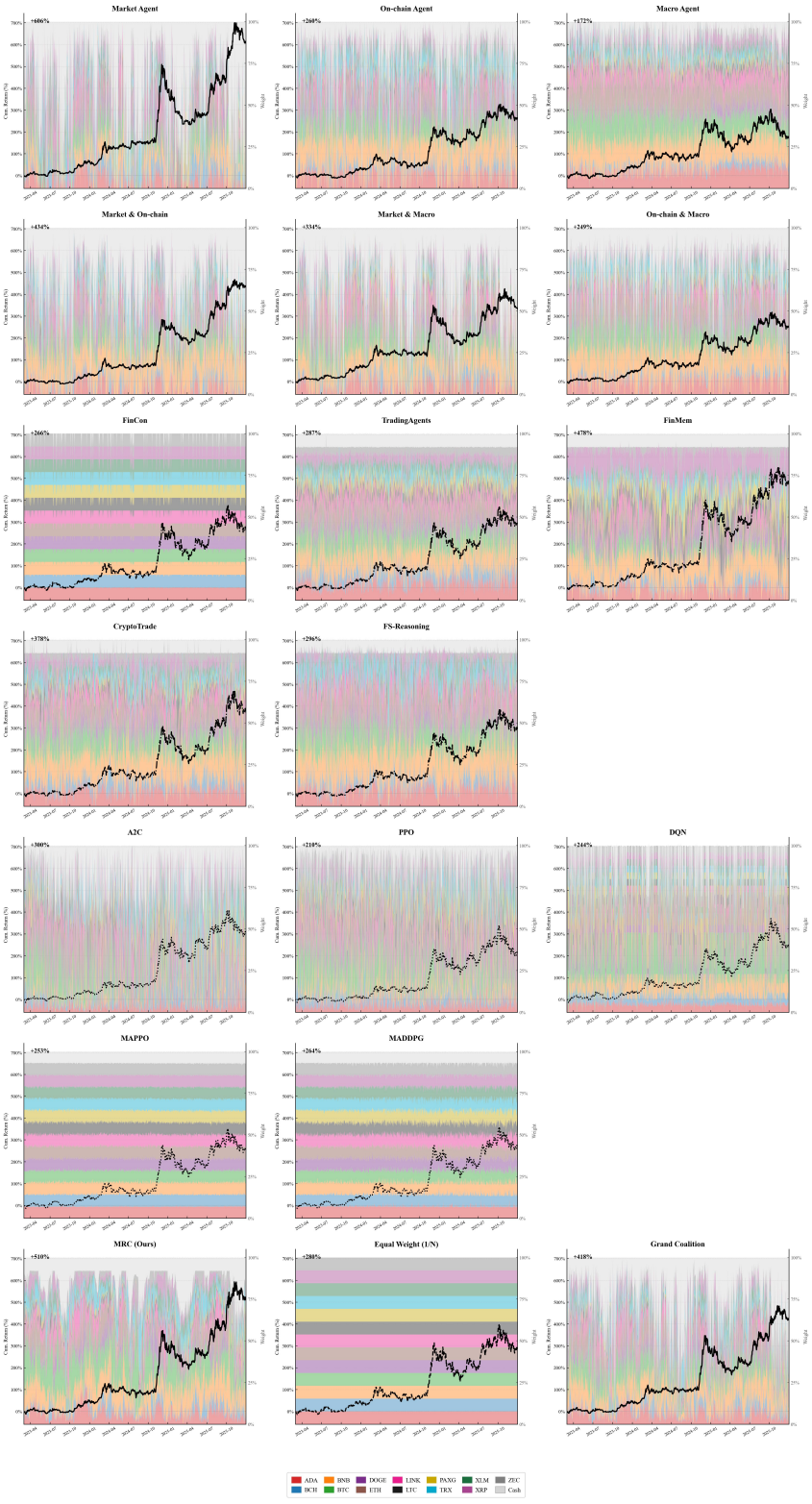

Figure 11: Cumulative returns and portfolio weights for all 17 strategies.

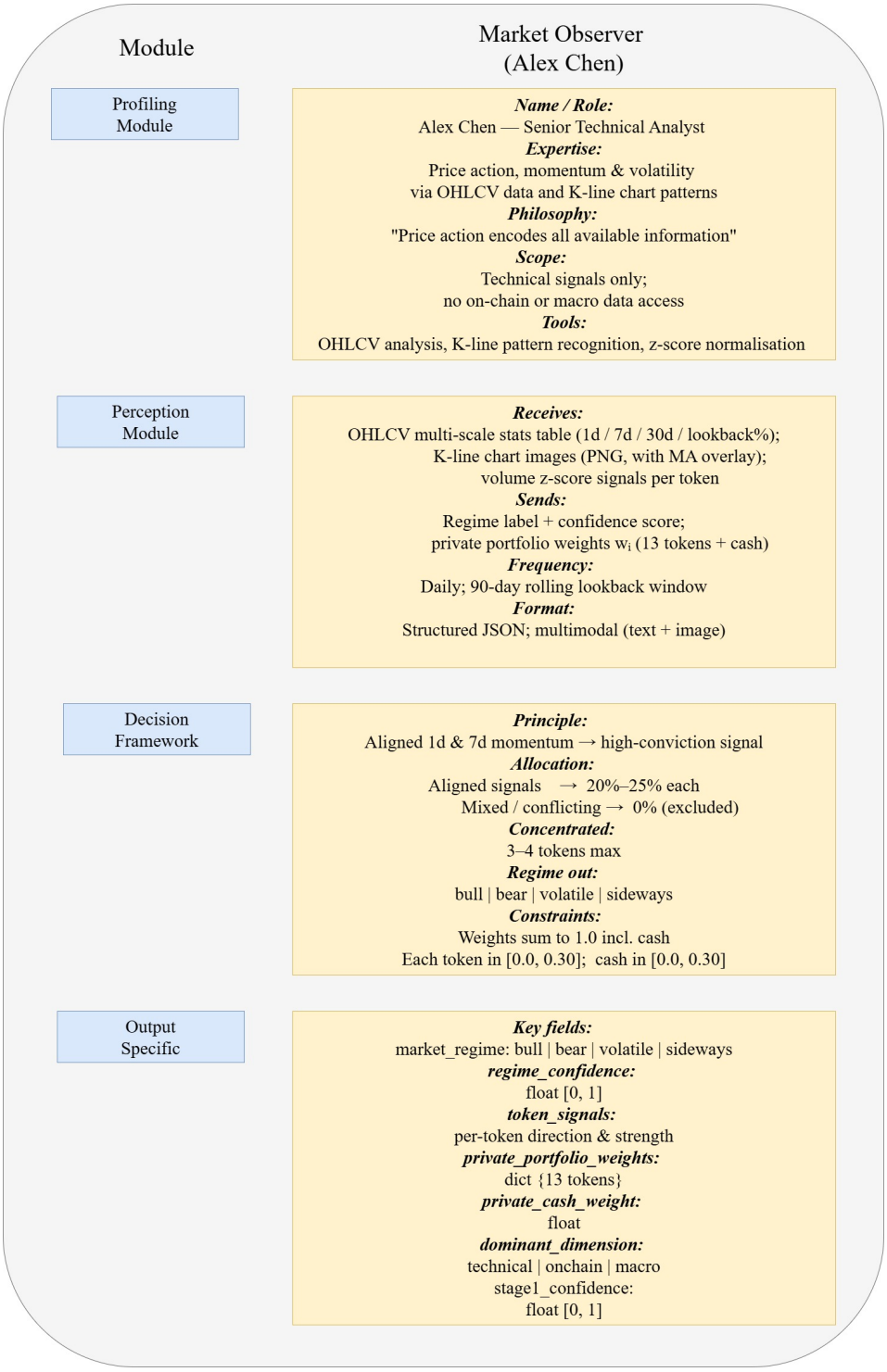

Figure 12: Full specification of A_1 Market Observer (Alex Chen).

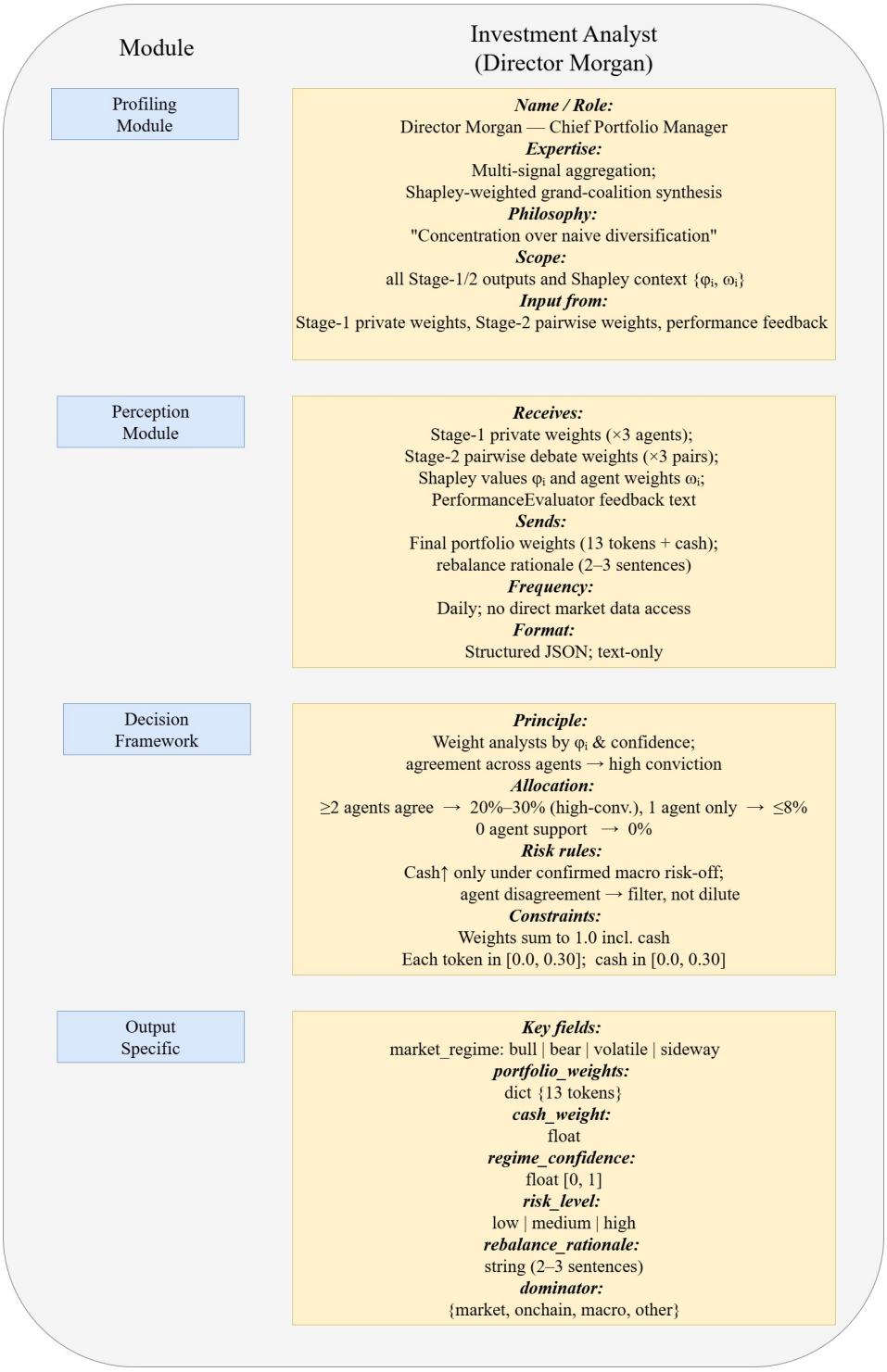

Figure 13: Full specification of A_4 Investment Analyst (Director Morgan).

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: All claims in the abstract and introduction are supported by experimental results in the paper.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: Limitations are discussed in the Conclusion and the Discussion appendix, including the restriction to cryptocurrency assets, the fixed three-agent design, the zero-cost assumption, and the bear-regime performance gap.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes]

Justification: All assumptions are stated in the Preliminaries. Formal proofs are provided in the appendix, including a proposition on EWP cold-start attenuation and a full derivation and axiom verification for the three-player Shapley values.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [Yes]

Justification: All hyperparameters, dataset details, backbone configuration, and evaluation protocol are fully specified in the Implementation Details appendix. Code and demo data are released at `https://github.com/PEIYUNHUA/Market_Regime_Council`.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes]

   Justification: Code is released at `https://github.com/PEIYUNHUA/Market_Regime_Council` under the MIT license, with data samples and a description of each data source and its format. All data sources are described in the dataset inventory table in the appendix.

   Guidelines:

   - The answer [N/A] means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: The dataset, baselines, metrics, and hyperparameters are described in the Experiments section and the Implementation Details appendix.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: Results are reported as mean and standard deviation over five independent seeds. Significance of MRC's advantage over the second-best method is assessed with Welch's two-sample t-test, with significance levels annotated in the table captions.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.

- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: Mean and standard deviation of wall-clock inference time per LLM call and per full rebalancing period are reported for all seven backbone variants in the backbone ablation cost table in the appendix. All inference is performed via the OpenRouter API.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: The paper uses publicly available financial data and does not involve human subjects, private data, or any ethical concerns.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

Justification: The five-layer causal trace improves transparency of algorithmic trading decisions. Potential risks include signal manipulation and correlated positioning. The released code is for research reproducibility only.

Guidelines:

- The answer [N/A] means that there is no societal impact of the work performed.
- If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

Answer: [N/A]

Justification: The paper does not release new models or high-risk datasets. The supplementary code is a backtesting framework using existing public APIs.

Guidelines:

- The answer [N/A] means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [Yes]

Justification: All data sources and models are identified in the dataset inventory table in the appendix. They are accessed through public APIs under standard terms of use.

Guidelines:

- The answer [N/A] means that the paper does not use existing assets.

- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [Yes]

    Justification: The MRC codebase is released at `https://github.com/PEIYUNHUA/Market_Regime_Council` under the MIT license, with a README covering installation, data layout and commands, and with data samples.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: This paper does not involve crowdsourcing or human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

Justification: This paper does not involve human subjects research.

Guidelines:

- The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

Answer: [Yes]

Justification: LLMs are a core component of the method. Three specialist agents and one readout agent use the Qwen3-VL model family at temperature $T_{\mathrm{LLM}}{=}0.7$. Prompt templates are described in the Implementation Details appendix.

Guidelines:

- The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.