Four Engines on the Metal-and-Energy Tape

A Per-Asset Volatility-Scaled, Order-Flow-Aware Architecture for Commodity-Spot Calibration — Q7 AEGIS AI S5

Haneef R. Haqq, MBA[1]

[1]Quant7 Alpha, LLC

Technical Whitepaper · v1.0 · Strategy S5 (Commodity Spot)

June 30, 2026

Abstract

We describe S5, the commodity-spot member of the Q7 AEGIS AI family, which trades four spot instruments—gold (XAU), silver (XAG), crude oil (WTI), and platinum (XPT)—across four intraday horizons (1, 3, 5, 15 minutes) on a 24/7 decentralized spot venue with continuous, fractional position sizing. S5 inherits the reference four-engine architecture of S1 but specializes it for the distinct microstructure of metals and energy: precious metals and oil differ from one another in volatility, liquidity, and jump character far more than the four digital assets of S1 do, so a single exit geometry calibrated on one asset misfires on the others. The load-bearing specialization is a *per-asset volatility adjustment* that bakes an exit-distance multiplier into each instrument—scaling every take-profit leg and the break-even trigger—with dedicated 1-minute cases where the noise floor dominates. Four concurrent engines target trend, breakout, mean-reversion, and liquidity-sweep regimes, arbitrated by a seven-component confluence score, a five-state volatility-regime router with ADX-conditioned exit gating, and a path-geometry operator. Leverage is a function of order-flow conviction capped per asset, exits run a four-leg ladder governed by a single unified stop on the residual quantity with a three-tier MFE ratchet and a post-first-leg tiered trail, and a per-configuration, train-equals-serve hybrid-AI gate—anchored on the model's own score distribution to avoid deny-all starvation—filters entries. Risk is enforced preventively by sizing to drawdown headroom under a hard ceiling. We state limitations explicitly and make no realized-performance claims.

1. Edge Thesis

S5 trades directional commodity-spot prices, where the four canonical Q7 engines each correspond to a real intraday regime that survives on metals and energy. **Trend (TRND)** fires when the moving-average stack and trend-strength index align—gold and platinum in particular exhibit clean, persistent intraday drift around macro and session opens. **Breakout (BRK)** fires on volatility expansion beyond the Bollinger envelope, which on commodities captures genuine expansions driven by inventory prints, geopolitical headlines, and dollar repricing. **Mean-Reversion (MR)** fires on intraday extremes inside ranges, where dealers fade exhaustion bursts. **Sweep (SWP)** targets prior-high/low liquidity sweeps with rejection wicks. The four engines partition the regime space; confluence and the volatility-regime layer arbitrate which fires. The structural contrast with S1 is not the engine set but the asset set: XAU, XAG, WTI, and XPT span a far wider band of volatility and liquidity than four perpetual majors, so the specialization weight in S5 falls on exit geometry and per-asset scaling rather than on the engine semantics themselves.

2. Architecture

S5 is organized in three independently testable layers.

Signal. Per (asset, timeframe), the four engines compute entry candidates each bar and write a candidate direction plus an engine tag. Arbitration selects the first eligible engine whose direction-aligned confluence clears its per-engine floor, subject to a per-engine cooldown.

Intelligence. Five subsystems run in parallel: the confluence aggregator (Section 3); the volatility-regime router with ADX-conditioned exit gating (Section 6); the path-geometry operator (Section 7); the four-leg adaptive exit with per-asset volatility scaling (Section 8); and the hybrid-AI overlay (Section 9).

Execution. Engine arbitration → confluence gate → cooldown → hybrid-AI gate → order-flow-scaled fractional sizing → entry. Exits run the four-leg ladder under a single unified stop on the remaining quantity, with a break-even arm, a three-tier MFE ratchet, and a post-first-leg tiered trail.

Asset	Horizons	Configs	Leverage cap
XAU (gold)	1·3·5·15m	4	150×
XAG (silver)	1·3·5·15m	4	100×
WTI (crude)	1·3·5·15m	4	100×
XPT (platinum)	1·3·5·15m	4	75×

The 16 configurations (4 assets × 4 horizons) are each calibrated independently; the leverage cap is a per-asset ceiling, not an operating point. Position sizing is *fractional and continuous*—notional is set as a fraction of collateral times the conviction-scaled effective leverage, with no integer-contract quantization—which is what lets a low-DD commodity configuration run capital-efficiently rather than being forced up to a one-contract minimum.

3. Confluence Scoring and Gating

Per bar, seven 0–100 component scores are computed—trend agreement, trend-strength (ADX) band, direction-specific oscillator band, volume ratio, multi-timeframe agreement, Bollinger position, and candle-pattern strength—and combined into direction-conditional confluence:

$$\mathrm{conf}_{dir} = \Sigma_k\, w_k \cdot \mathrm{score}_{k,dir}, \quad \Sigma_k\, w_k = 1.$$

Each engine carries a calibrated confluence floor; the effective gate applies a per-direction asymmetry adjustment under a fixed feasibility cap. A per-configuration confluence override and a per-configuration order-flow floor are evaluated jointly—a bar must clear both the effective confluence threshold and the minimum order-flow score to be entry-eligible. These per-config overrides are how thin-tape configurations (for example, the sparsest platinum horizon) receive a calibrated gate relaxation to maintain trade frequency without hand-setting any individual rule—the override is an output of calibration, not an exception to it.

4. The Four Engines and Regime Coverage

The engines are deliberately non-overlapping in their activation conditions, so that confluence arbitration selects among regimes rather than among redundant signals. TRND and BRK target directional persistence and expansion; MR and SWP target reversion and liquidity events. The coverage argument is empirical: each engine's activation maps to a documented intraday behavior of the commodity tape, and the union spans the observed regime distribution for gold and platinum especially, which show the cleanest four-regime structure.

Two structural carve-outs are applied where microstructure dictates narrower routing, and both are calibration findings rather than accommodations. First, **crude oil on the shortest horizons** is run mean-reversion-only with a tightened exit envelope (narrower take-profit cascade, tighter stop, earlier break-even arm): directional engines on WTI short horizons historically manufacture high-commission losing trades around inventory days, and restricting to the mean-reversion engine removes that directional sabotage while retaining the available edge. Second, the protected silver horizon is frozen at its validated baseline—its parameters are held fixed across cycles so that per-cycle re-search cannot drift a live, passing configuration off its optimum. Engines can additionally be disabled per configuration when structurally adverse for that (asset, horizon), which is how mean-reversion-dominant profiles emerge for the tighter-ranging tapes.

5. Order-Flow-Driven Leverage

S5 does not size with a fixed multiplier; per-trade leverage is a function of confluence and *order-flow* conviction, capped per asset and applied to a continuous fractional notional. This choice is grounded in the limit-order-book literature. Order-flow imbalance (OFI)—the net change in size at the best bid and ask, inclusive of cancellations—explains short-horizon mid-price drift far better than trade imbalance (reported $R^2 \approx 65\%$ vs. 32%) [1, 3]. The price-impact slope is *inversely proportional to market depth* [1], which dictates depth-normalized order-flow: leverage is attenuated when our own flow is large relative to local depth—materially relevant for silver and platinum, whose books are thinner than gold's and where a retail-scale order is a non-trivial fraction of top-of-book. The Markovian queueing treatment further yields an order-flow-implied volatility, $\sigma_{OF} \propto \sqrt{(\lambda/D)}$, forecastable from flow statistics without observing price [2]—used to open the breakout gate and scale trend leverage when σ_{OF}/σ_{ATR} expands, and to favor mean-reversion when it contracts. The same theory supplies a closed-form queue-imbalance direction probability with empirical first-lag *negative* autocorrelation [2], the structural fact the MR engine exploits.

6. Volatility-Regime Router (Λ-Vol) and ADX Exit Gating

A deterministic five-state classifier—trend-calm, trend-active, mean-revert, stress-high, transition—maps the short/long volatility ratio and drift (in standard-deviation units) to regime-conditional stop and target multipliers. It is a rule-based state machine, not a learned model, chosen for interpretability and stability; it routes exit geometry and flags engine downsizing under stress. The exit machinery is additionally *ADX-gated*: the trend-strength regime conditions how aggressively the ladder trails and when the runner is permitted to extend, so that a strong-trend regime stretches targets and a weak/ranging regime compresses them. The motivation for an explicit regime layer is the documented tendency of high-volume regimes to shift return distributions from unimodal to multi-modal [8], which a single static exit geometry serves poorly; the same realized-vol-surface decomposition that motivates a three-factor level/slope/convexity reading of volatility dynamics [7] is the principled extension of this layer, and is treated as a roadmap item rather than a present claim.

7. Path-Geometry Operators (ζ-Field)

The ζ-Field supplies a curvature signal to the direction-aligned confluence channel via three operators on the log-price path: velocity (first derivative), jerk (third derivative, via a

Savitzky–Golay coefficient), and curvature (discrete circulation). The three combine into a single signed score that boosts the aligned side additively, leaving the gate structure unchanged, and degrade to a no-op if the operator is unavailable. The rationale is that the *order* of moves and cross-channel geometry carry information that vanishes under bag-of-features representations—the same premise that motivates the signature method in machine learning [6]; the ζ-Field is a lightweight, deterministic instance of that geometric coverage, and its confluence-boost weight is calibratable, with a guard that prevents it from suppressing the underlying confluence signal below the gate floor.

8. Four-Leg Exit, Unified Stop, and Per-Asset Volatility Scaling

Each position scales out across four take-profit legs plus a runner, with engine-specific leg allocations and target multipliers (calibrated per configuration). The residual quantity is governed by a *single unified stop*: rather than maintaining independent stops per leg, S5 carries one global stop on the remaining position that is advanced by the exit state machine—to break-even after the first leg arms (or once favorable excursion exceeds a threshold), then ratcheted in three tiers as the maximum favorable excursion advances, with a post-first-leg tiered trail that tightens the anchor as successive legs fill. This keeps the exit coherent: there is always exactly one risk boundary on the open quantity, which is what makes the leg machinery composable with the hard-risk path. Any hard-risk event (circuit breaker, rolling halt, liquidation guard, macro gate) closes all remaining legs immediately—the leg-and-ratchet machinery operates only on the normal exit path.

8.1 Per-asset volatility adjustment (assetVolAdj)

The distinctive S5 control is a *per-asset volatility adjustment*: each instrument carries a baked exit-distance multiplier that scales every take-profit leg and the break-even trigger. This is load-bearing because the four S5 assets are not interchangeable in volatility or liquidity—silver and platinum are materially more volatile and thinner than gold, and crude carries inventory- and geopolitics-driven jump risk that neither metal shares. A single exit geometry calibrated on gold's volatility would place silver and crude targets too close (premature scale-outs that forfeit the move) and platinum stops too tight (stopped out on ordinary noise). The per-asset multiplier stretches or compresses the entire ladder—and the break-even arm with it—to the instrument's own volatility, so the *shape* of the exit is asset-appropriate while the *logic* remains shared across the family. The 1-minute horizons carry dedicated multiplier cases, because at one-minute resolution the microstructure noise floor dominates the signal and a separate, wider scaling is required to keep the break-even trigger from arming on noise alone. The adjustment composes multiplicatively with the Λ-Vol regime multipliers (Section 6): the regime layer adapts the geometry within an asset over time, while the per-asset multiplier sets the baseline scale across assets.

9. The Hybrid-AI Gate (Train = Serve)

S5 filters rule-based entries through a hybrid-AI overlay that fuses three models—a supervised directional learner, a financial-language sentiment model, and a market-regime classifier—in the spirit of the multi-modal architecture of [9], adapted to per-configuration, intraday, train-equals-serve operation. The motivation is the one that paper documents empirically: trend and mean-reversion logic fail through regime transitions and are blind to news shocks, and a fusion of structured ML with sentiment and regime awareness is materially more robust than any isolated predictor. The overlay is calibratable on/off per configuration; if the optimizer selects it off, the rule-based stack carries the decision.

9.1 Directional model — gradient-boosted trees (XGBoost)

The directional learner is an XGBoost classifier that predicts whether the forward move is favorable. The reference design [9] frames this as a binary next-period-direction problem (label 1 if the next-period return is positive) over a compact technical feature set, with ~200 boosting rounds, tree depth 6, and a 0.05 learning rate on a 70/30 split. S5 adapts that recipe to its short-horizon, per-(asset, timeframe) setting: labels are the forward 10-bar return thresholded at ±0.5% (favorable/unfavorable), and the feature set is *direction-conditional*—separate long and short framings (roughly a dozen technical and microstructure features plus an is-long indicator and the ζ-Field velocity/jerk terms) are computed so one model scores both sides. Mechanically, XGBoost is an additive ensemble of shallow regression trees fit by stagewise gradient boosting on a regularized logistic objective: each new tree fits the negative gradient of the loss given the current ensemble, scaled by a shrinkage rate, with row/column subsampling and L1/L2 leaf-weight penalties; the ensemble outputs a calibrated probability. Two disciplines distinguish S5's use from the reference: the model is trained *in-cycle* on the current cycle's own accumulated trades (no stale prior-cycle model), and it is admitted only if it clears a per-direction discriminability floor (AUC ≥ 0.55 long *and* short) and an adversarial leakage bound (adversarial AUC < 0.85).

Score-anchored entry threshold (mlMinScore). The entry threshold is itself a calibratable parameter, but it is *anchored on the model's own score distribution* rather than fixed at a nominal probability. A hard-coded threshold (for example, 0.55) silently becomes a deny-all gate when an in-cycle model's probability ceiling sits below it—the gate then vetoes essentially every entry and the configuration starves to a blocked, sub-threshold trade count for reasons that have nothing to do with edge. S5 instead derives the operating threshold from the realized score distribution of the admitted model, so the gate filters the weaker tail of the model's *own* signal rather than rejecting against an unreachable absolute. This anchoring is what keeps low-signal commodity horizons from being gated to zero, and it preserves train=serve: the threshold that was calibrated is the threshold that is deployed.

9.2 Sentiment model — FinBERT

The sentiment channel is FinBERT [10], a BERT-family transformer fine-tuned on financial text. Each news item is tokenized and contextualized through stacked self-attention layers and mapped by a classification head to probabilities over {positive, neutral, negative}; domain fine-tuning is what lets it parse negation, hedging, and finance-specific jargon that generic sentiment models miss—and commodities are an especially news-driven tape, where OPEC decisions, inventory surprises, central-bank gold flows, and geopolitical shocks move price discontinuously. Per asset per day the article-level scores are aggregated into a normalized polarity,

$$S_t = (1/N)\ \Sigma_i\ (\ P_{positive}^{(i)} - P_{negative}^{(i)}\ ),\quad S_t \in [-1, 1].$$

Following [9], S5 consumes S_t as a *risk filter and soft threshold skew, never as a standalone entry signal*: a sufficiently negative S_t (below a calibrated floor) suspends new entries and can flatten exposure, shielding core positions against adverse-news shocks that purely technical logic is blind to. *Honesty note:* in a historical backtest without a time-aligned, point-in-time commodity news feed for the evaluated window, this channel is run neutral under a signed no-feed exemption (honest-empty) rather than fabricating sentiment; the exemption is recorded in the cycle dossier, and train=serve is preserved by deploying the same neutral configuration that was calibrated.

9.3 The tandem and the regime channel

The third model is the market-regime classifier—the Λ-Vol router (Section 6) at the strategy level, and a regime label at the overlay level—which conditions both the gate and the exit geometry. The three combine as a tandem: the directional model proposes a probability, the sentiment channel can veto or attenuate it, and the regime label selects the operating envelope. This is the multi-modal fusion that [9] shows outperforms any single isolated predictor across changing regimes. The decisive engineering choice is that all three are computed in-cycle from the same data the system trades—the overlay calibrated is the overlay deployed, so the gate cannot be a flattering in-sample artifact of a model the live system never runs.

10. Calibration and Acceptance

Each configuration is accepted into a cycle only on a simultaneous gate—win rate, gross profit factor, trade frequency, drawdown (a soft, logged criterion), and a commission-to-gross-profit bound (a hard block)—evaluated on the hybrid-AI-gated system, with the hybrid-AI discriminability bounds enforced separately. Commodity configurations are particularly sensitive to the commission bound at low trade frequency, so the per-config sizing, order-flow floor, and confluence override are tuned jointly to keep commission within budget. Cycle eligibility additionally requires full 16/16 coverage with no blocked configuration, positive aggregate result, and in-cycle model freshness. The protected silver horizon is held at its validated baseline and exempt from re-search. Promotion toward live deployment requires consecutive consistent cycles, specification-to-implementation parity within tolerance, and an independent re-derivation—the governance described in the program's pipeline whitepaper.

11. Risk and Venue

Risk is enforced preventively at the sizing layer: positions are sized so that a worst-case adverse move, measured from the monotonically ratcheting equity peak, cannot carry the account past the hard drawdown ceiling—a structural constant of 24% that is identical across the specification, the calibrator, and the executor. The ceiling is enforced by sizing to *headroom*, not by a reactive kill: a reactive intrabar stop cannot prevent an overnight or weekend gap from breaching the ceiling, whereas sizing each position to the distance remaining to the ceiling bounds the breach preventively. Order-flow-aware leverage attenuation (Section 5) further reduces self-impact in the thinner metal books. The execution venue is a 24/7 decentralized commodity-spot venue (gTrade on Arbitrum) with USDC collateral; commissions are a percentage of notional per side under a conservative taker-on-both-sides round-trip assumption, and continuous fractional sizing avoids any integer-contract minimum that would otherwise inflate small-account exposure.

12. Limitations

The order-flow results that motivate the leverage map are strongest with a live depth feed; under a bar-derived order-flow proxy the depth normalization is approximate, and the architecture treats the live-depth path as an enhancement, not a precondition—this matters more on silver and platinum than on gold. The per-asset volatility multipliers are calibrated to historical volatility regimes and can lag a structural break in an asset's volatility character; the Λ-Vol classifier is deterministic and can saturate; and the realized-vol-surface and cross-asset network-momentum extensions that the metals cluster invites [2, 7] are roadmap items, not present claims. The hybrid-AI gate constrains but does not certify causal or durable structure, and the score-anchored threshold prevents starvation without guaranteeing the underlying signal is economically meaningful. Execution, liquidity, leverage, venue, and regulatory risks are not eliminated by architectural integrity. As with every Q7 deliverable, no realized-performance representation is made, and the acceptance criteria above are calibration gates, not outcomes.

References

[1] R. Cont, A. Kukanov, S. Stoikov. *The Price Impact of Order Book Events.* 2011.

[2] R. Cont, A. de Larrard. *Price Dynamics in a Markovian Limit Order Market.* 2011.

[3] M. Gould et al. *Limit Order Books* (survey). Quantitative Finance, 2013.

[4] *Follow the Leader: Enhancing Systematic Trend-Following Using Network Momentum.*

[5] *Network Momentum across Asset Classes.*

[6] *A Primer on the Signature Method in Machine Learning.*

[7] R. Cont, J. da Fonseca. *Dynamics of Implied Volatility Surfaces.* Quantitative Finance, 2002.

[8] *Quantum-Probability-Theoretic Asset Return Modeling* (multi-modal return distributions at high-volume regimes). 2024–2025.

[9] V. N. K. Pillai, A. Ajith, S. K. J. *Generating Alpha: A Hybrid AI-Driven Trading System Integrating Technical Analysis, Machine Learning and Financial Sentiment for Regime-Adaptive Equity Strategies.* ComSIA 2026 (Springer LNNS).

[10] D. Araci. *FinBERT: Financial Sentiment Analysis with Pre-trained Language Models.* 2019.

Methodology disclaimer. This document describes a research and trading-system architecture. It is not investment advice, an offer, or a solicitation. Nothing herein represents realized trading results or guarantees future performance; calibration acceptance criteria are gates, not outcomes, and any backtested figures are research artifacts subject to the limitations inherent in historical simulation. Material intended for public distribution should be reviewed by qualified counsel prior to publication.