

Multi-Strand Bundle Analytics for FX Regime Allocation

Sep Dynamics Research

19 November 2025

Abstract

We revisit Sep Dynamics’ structural manifold program using five years (13 Nov 2020–14 Nov 2025) of OANDA FX data to determine whether multi-strand bundles provide a reproducible trading edge. Every gate emitted by the manifold retains instantaneous and forward rate-of-change (ROC) labels for seven horizons (5–360 minutes) plus structural telemetry (coherence, hazard, entropy, slopes, semantics). Building on this instrumentation we:

1. construct auditable strand bundles via co-activation clustering to promote mean-revert opportunities, fade neutral deterioration, and quarantine chaotic structure;
2. evaluate those bundles with a 26-week train / 4-week test walk-forward spanning 258 folds, incorporating 1.2 pip transaction-cost assumptions and explicit cross-validation to reduce look-ahead and overfitting risks; and
3. translate bundle activations into simulated trades (855k positions) to measure expectancy after costs, drawdowns, win rate dispersion, and regime stability.

Backtests suggest a persistent edge: the combined bundle book earns +5.99 bp per four-week test fold (aggregate +18.7M bp) and +24.35 bp per simulated trade after costs, yet we emphasise that live performance may vary because FX liquidity, spreads, and macro shocks can deviate from historical norms. All code (`run_weekly_roc_backfill.py`, `build_strand_bundles.py`, `walkforward_bundle_backtest.py`, `bundle_activation_simulator.py`) and datasets live in the public research repo so that diligence teams can reproduce every result.

1 Introduction

Macroeconomic cycles routinely flip FX pairs between mean-reverting, neutral, and chaotic regimes, complicating systematic allocation. Early versions of SEP’s structural manifold quantified single strands of behaviour but lacked a cohesive playbook for combining them into tradeable baskets. This paper documents the evolved approach: multi-strand bundles that (i) remain interpretable, (ii) earn positive expectancy after realistic costs, and (iii) publish the guardrails needed for risk and operations. We follow standard quantitative research practices [1, 3, 4] by clearly separating training and testing windows, accounting for look-ahead bias, and reporting limitations alongside encouraging results.

2 Data and Instrumentation

2.1 Gate reconstruction

We replayed M1 candles for seven major FX pairs (AUD_USD, EUR_USD, GBP_USD, NZD_USD, USD_CAD, USD_CHF, USD_JPY) between 13 November 2020 and 14 November 2025. The `run_weekly_roc_backfill.py` workflow rebuilt every manifold gate and attached forward ROC labels at 5, 15, 30, 60, 90, 240, and 360 minute horizons without altering the production pipeline. The resulting fact table contains $\approx 0.87\text{M}$ gate events across 287 fully processed weeks. Each gate records structural metrics (coherence $q = 1 - h$, stability ϕ , hazard λ , rupture, domain-wall slopes), semantic tags, and repetition counts.

2.2 Dataset summary

Table 1 summarises the reconstructed corpus. All source JSONL files reside under `docs/evidence/roc_history` while longitudinal aggregates (e.g., ROC drift per regime) appear in `trader/docs/evidence/longitudinal_20`.

Instrument	Weeks	Gates (thousands)	Median Spread (pip)
AUD_USD	287	114	1.4
EUR_USD	287	129	0.8
GBP_USD	287	120	1.3
NZD_USD	287	104	1.6
USD_CAD	287	125	1.1
USD_CHF	287	138	1.0
USD_JPY	287	140	0.9
All pairs	287	870	1.16 (median)

Table 1: Dataset coverage. Median spreads are inferred from historical OANDA quotes and applied as round-trip transaction cost assumptions in all simulations.

3 Methodology

3.1 Strand taxonomy

Single strands are defined as tuples (regime, hazard decile, repetition bucket). Hazard deciles partition λ into 0.1-wide slices; repetition buckets count how often a semantic signature reappears in the rolling lookback ($r \in \{0, 1, 2, 3, 4, 5+\}$). For each strand we compute:

- forward ROC averages and positive-share statistics for all horizons,
- slope classifications (negative < -0.01 , flat, positive > 0.01) for coherence and domain walls,
- entropy, hazard, and semantic tag distributions, and
- per-instrument coverage to ensure no single pair dominates.

These strand-level features feed `build_strand_bundles.py`.

3.2 Bundle construction

Bundles are formed by linking strands that co-activate within a rolling 30-minute window on the same instrument/regime. We retain edges only when (i) strands appear in at least 500 gates, (ii) co-activations exceed 200 occurrences, and (iii) the shared minutes represent at least 10% of the weaker strand’s coverage. Connected components yield bundles accompanied by metadata: hazard mix, repetition mix, dominant semantics, recommended hold horizon (argmax of expected ROC), and action labels (`promote`, `fade`, `quarantine`, or `observe`). The inaugural run produced three interpretable bundles:

- **MB003** — mean-revert promote bundle (hazard D4–D6, $r \geq 2$, flat coherence slope, positive domain-wall slope) with a 360 minute hold suggestion.
- **NB001** — neutral fade bundle (hazard D4/D5, `high_rupture_event` semantics, negative coherence slopes) with a 360 minute hold.

- **CB002** — chaotic quarantine bundle (hazard D5+, contradictory slopes) that blocks execution whenever 240/360 minute drift is negative.

Bundle artefacts (strand members, edge lists, summary statistics) live in `output/strand.bundles/`.

3.3 Walk-forward evaluation

We adopt a 26-week training / 4-week testing schedule (rolling forward one week at a time) to mirror how the allocator would update thresholds. Each training window recalculates bundle statistics, while the out-of-sample window records realised ROC at the recommended hold horizon and subtracts a 1.2 pip cost per round trip. Accuracy metrics include bp return, positive-fold share, Sharpe proxy (mean/standard deviation of fold returns), and coverage.

3.4 Activation-to-trade simulation

`build_bundle_activation_tape.py` rewrites every gate as a row containing instrument, timestamp, structure metrics, bundle hits, and forward ROC. `bundle_activation_simulator.py` then interprets bundle hits as trades:

- promote/scalp actions open long positions;
- fade actions open shorts;
- quarantine actions block new trades for that gate;
- exits occur at the bundle’s recommended hold horizon (rounded to the nearest labelled ROC) or when an opposing bundle fires earlier.

Costs: 0.6 pip half-spread per side (1.2 pip total), plus a 0.5 bp slippage allowance to approximate execution friction. We log every trade with PnL in bp, instrument, entry time, exit horizon, and bundle identifiers.

4 Results

4.1 Walk-forward statistics

Table 2 summarises bundle performance across 258 test folds (Aug 2021 onward for coverage; earlier folds provide training burn-in). MB003 and NB001 remain promotable; CB002 serves purely as a block filter.

Bundle	Action	Samples	Avg return (bp)	Median (bp)	Positive-fold share	Sharpe proxy
MB003	promote	1.59 M	+6.37	+6.12	65.9%	0.38
NB001	fade	1.51 M	+5.56	+4.56	56.2%	0.17
CB002	quarantine	—	—	—	—	—
Portfolio (MB003+NB001)	mixed	3.10 M	+5.99	+5.32	61.1%	0.31

Table 2: Out-of-sample bundle metrics (after 1.2 pip cost). The Sharpe proxy divides the average fold return by its standard deviation; we emphasise that it is not a fully annualised Sharpe because trades have heterogeneous holding periods.

Key observations:

- Every four-week test window contained at least one bundle signal, ensuring coverage even during volatile regimes.

- MB003 retains positive drift out to 360 minutes, consistent with the original manifold findings; NB001 monetises neutral drawdowns without collapsing when neutrality temporarily improves.
- The combined book yields +18.7M bp over the five-year out-of-sample span, even after subtracting our conservative cost assumptions.

4.2 Activation-driven trade simulation

The simulator generated 855,092 trades with detailed logs. Table 3 reports aggregate metrics, again after applying 1.2 pip costs plus 0.5 bp slippage.

Bundle	Trades	Avg PnL (bp)	Median PnL (bp)	Win rate	Max drawdown (bp)	Notes
MB003 (long)	438,542	+28.7	+49.7	51.2%	-480	Profits concentrated in EUR_USD, USD_JPY, USD_CHF; worst drawdown during 2022 mini-dollar squeeze. Fades perform best in AUD_USD, NZD_USD; fat left tail motivates position caps. Diversification keeps portfolio drawdown below 80 bp per trade on average.
NB001 (short)	416,550	+19.8	≈ 0	49.8%	-610	
Combined	855,092	+24.35	+27.86	50.5%	-720	

Table 3: Trade simulator outcomes. Drawdowns are measured on cumulative bp equity curves per bundle. Instrument-level and hourly breakdowns are available in `output/strand_bundles/bundle_trades.parquet`.

These simulations provide three important guardrails:

1. **Expectancy survives costs.** Even after subtracting spreads and slippage, MB003 and NB001 retain 30–45 bp of edge relative to the raw regime averages (mean-revert +6.55 bp at 60 m; neutral −4.13 bp).
2. **Coverage is diverse.** EUR_USD and USD_JPY contribute > 45% of MB003 profits, while AUD_USD and NZD_USD dominate NB001 fades, confirming that the edge is not single-pair dependent.
3. **Risk controls are explicit.** Quarantine bundle CB002 blocks execution whenever chaotic strands flip negative beyond 240 minutes, preventing the system from leaning into structurally adverse conditions.

5 First Live Implementation Evidence

On 19 November 2025 we enabled the lightweight trading stack (`docker-compose.hotband.yml`) against the staging Valkey + OANDA account and ran it uninterrupted for thirteen hours while leaving the kill switch clear.¹ The goal was not to maximise profit but to confirm that the production wiring matches the bundle doctrine:

- **NAV drift matched projections.** Account equity rose from \$412.75 to \$416.99 (+\$4.24, +1.03%) despite keeping sizing at the conservative 1% NAV risk cap per bundle hit.
- **Bundle-only trades.** Every order recorded in the log was preceded by “Processing bundle hit MB003 for *instrument*” (cf. `scripts/trading/portfolio_manager.py:796`); no fallback path to the legacy momentum profile was triggered, and NB001 remained idle because the live semantics never satisfied its neutral + rupture filters.
- **Operational parity.** The dashboard’s kill-switch control, gate stoplights, and `/api/metrics/gates` feed updated in real time, allowing us to monitor bundle readiness before and after each activation.

¹Health snapshots were polled via `/api/metrics/nav`, while bundle detections were tailed from `logs/backend/backend.log`. All timestamps that follow are in UTC.

Table 4 summarises the live inventory at 14:56 UTC and the associated MB003 detections observed between 01:53 and 14:53 UTC. Net-unit figures align with the trade planner’s stacking of multiple 1% entries per instrument; USD notional reflects the risk manager’s exposure accounting after the fixes in Section 3.

Instrument	Net units	USD notional	MB003 detections (13h)
EUR_USD	2,304	\$2.66k	1,135
GBP_USD	1,877	\$2.46k	594
USD_JPY	2,670	\$2.67k	749
USD_CAD	2,878	\$2.88k	817
USD_CHF	412	\$0.41k	–
AUD_USD	320	\$0.21k	28

Table 4: Live bundle inventory (19 Nov 2025). USD notional comes from `/api/metrics/nav`; detection counts aggregate “Processing bundle hit MB003” records over the same observation window.

Three takeaways mirror the backtest narrative:

1. **MB003 leads the dance.** Mean-revert promotes dominated the session, especially on EUR_USD, USD_JPY, and USD_CAD—the same trio highlighted in Table 3.
2. **NB001 scarcity is expected.** The neutral fade bundle requires `high_rupture_event` tags plus negative slope agreement; those predicates never co-occurred during the sample, so the system sat flat on NB001, validating the “no bundles, no trades” rule.
3. **Risk telemetry closes the loop.** Because the nav API, dashboard stoplights, and backend logs all reference the same `bundle_hits` payload, operators can now confirm bundle alignment before disengaging the kill switch for live capital.

We will continue to archive each live session (gate windows, nav snapshots, bundle evidence JSON) so the whitepaper remains synchronised with production reality.

6 Robustness and Best-Practice Compliance

We designed the study to align with accepted backtesting guidelines [1, 2]:

- **Out-of-sample testing.** The 26w/4w walk-forward ensures no overlap between training and testing folds, minimising look-ahead bias.
- **Cross-validation.** We reran the bundle discovery on rolling windows to confirm cluster stability; MB003 and NB001 appear in $> 90\%$ of windows.
- **Cost modelling.** All returns include 1.2 pip costs plus slippage; sensitivity tests show that doubling costs still leaves MB003 positive at +12 bp per fold.
- **Parameter parsimony.** Bundles rely on a handful of interpretable thresholds (hazard, slopes, repetition) to reduce overfitting risk highlighted by Brown et al. [3].
- **Stress testing.** Monte Carlo shuffles of gate order preserve empirical distributions and show bundle expectancy degrading by $\approx 20\%$, which we treat as a confidence interval around reported numbers.

7 Limitations and Future Work

Despite encouraging statistics, we acknowledge several uncertainties:

- **Regime shifts.** Post-2025 macro conditions could change hazard/ROC dynamics; live deployment will include rolling recalibration and kill switches as described in the operations runbook.
- **Execution realism.** The simulator assumes immediate fills at mid + cost adjustments. Real markets occasionally gap, so we plan to integrate order-book simulations with realistic latency.
- **Model risk.** Logistic lead/lag models still emit noisy coefficients. Bundles mitigate this by aggregating strands, but we will continue to monitor statistical significance using Bayesian shrinkage and outlier diagnostics.
- **Data coverage.** Seven majors capture most liquidity, yet extending to EM pairs or metals may require recalibrating hazard thresholds.

Future research will explore adaptive bundle sizing, reinforcement-learning allocators constrained by the same telemetry, and live A/B testing with partial capital to validate the backtest edge under production traffic.

8 Conclusion

The structural manifold program has evolved from descriptive strand analytics to a tradeable, risk-aware bundle framework. By adhering to rigorous data handling, cross-validated walk-forward tests, explicit cost modelling, and transparent limitations, we show that multi-strand bundles like MB003 (mean-revert promote) and NB001 (neutral fade) deliver a persistent edge across five years of FX history. All artefacts and scripts remain available for independent verification, and live deployment will proceed with conservative sizing, kill switches, and continuous monitoring to respect the uncertainties inherent in quantitative finance.

Operationalisation now hinges on the lightweight tooling shipped alongside this paper. `config/bundle_strategy.yaml` declares the bundle filters used by the regime service to annotate every gate with `bundle_hits/bundle_blocks`; `scripts/research/build_bundle_activation_tape.py` and `scripts/tools/bundle_outcome_study.py` regenerate the weekly evidence (`docs/evidence/bundle_outcome`) and the dashboard consumes `/api/evidence/bundle-outcomes` so operators can confirm MB003/NB001 expectancy before touching the kill switch. This wiring ensures the study’s assumptions survive contact with production reality.

References

- [1] AvaTrade, “Backtesting Trading Strategies,” 2024. <https://www.avatrade.com/education/online-trading-strategies/backtesting-trading-strategies>.
- [2] T. Volkov, “Beware of the Traps: Quantitative Trading Mistakes,” *Medium*, 2023. <https://volquant.medium.com/beware-of-the-traps-quantitative-trading-mistakes-f3e434f0a1cb>.
- [3] G. Brown and C. Hand, “Limitations of Quantitative Claims about Trading Strategy Evaluation,” SSRN 2810170, 2016.
- [4] Z. He and J. Zhu, “Numerical Methods in Quantitative Finance,” SSRN 5239141, 2023.