Crypto Multi-Factor Engine: Seven Trials, Capacity Decay, and No-Deploy Verdict

Arhan Canli
Founder, System Architect, and Quantitative Researcher
Canli Capital / AlphaC Algorithms

VERSION	0.1.0
RELEASED	23 August 2026
REGISTRY KEY	crypto_multifactor_engine
STATUS	Working paper preprint / not peer reviewed

> Research simulation and, where explicitly identified, Alpaca paper evidence. No funded performance, peer review, independent replication, external acceptance, or future return is claimed.

Short title: AlphaForge crypto multi-factor engine lineage
Author: Arhan Canli, Founder, System Architect, and Quantitative Researcher, Canli Capital
Research system: ALPHAC / AlphaForge
Family key: `crypto_multifactor_engine`
Status: public research record; not peer reviewed; not an investment solicitation
Evidence date: 2026-08-22

Abstract

AlphaForge ran a historical robustness matrix over crypto perpetual futures from 2021-01-01 through 2026-06-01. Seven charged identities belong to the broad multi-factor engine: a baseline rank allocator; ML, regime, and combined overlays; daily rebalancing; a tighter no-trade band; and mean-variance allocation. An eighth matrix identity explicitly selected carry and residual reversal and is accounted for in the separate `crypto_carry` family.

None of the seven engine identities cleared its artifact-era DSR gate. Annualized Sharpe ranged from -0.7901 to +0.0424 and maximum drawdown from 13.59% to 33.93%. The four baseline/overlay arms all produced the same Sharpe 0.0424 and DSR 0.2112. Cross-sectional combinatorially symmetric validation across all eight matrix trials reported PBO 0.8818 against a gate below 0.20 and no deflated winner. The matrix verdict was **NO-DEPLOY**.

A capital sweep on the regime-gated engine showed Sharpe 0.4009 at $100,000, 0.0424 at $1 million, and -0.3720 at $10 million under the committed impact model. This is useful capacity-failure evidence, not a promise that the $100,000 result is repeatable. The family has no admitted sleeve and no broker-reconciled forward record.

Claim boundary

This paper reconciles seven immutable return identities to one persisted matrix and publishes the positive, negative, capacity, and selection-adjusted evidence together. It does **not** establish a forward Sharpe, expected maximum drawdown, independent economic mechanism, live result, or future return. The artifact serializes `alpha_names: null`; therefore the exact expanded default feature list is not recoverable from the matrix configuration alone. No feature-level attribution is claimed.

Research question and overlap decision

The engine asked whether a broad cross-sectional blend becomes robust through nonlinear sizing, regime gating, cadence, turnover control, or covariance-aware allocation. Possible mechanisms come from the component families, carry, momentum, reversal, liquidity, and defensive effects, not from the word “multifactor.” Combining weak signals can reduce noise, but complexity can also increase estimation error and hidden selection.

Liu, Tsyvinski, and Wu construct crypto market, size, and momentum factors (Journal of Finance); Liu and Tsyvinski document crypto return predictors and differences from conventional asset classes (Review of Financial Studies). Bailey and coauthors develop the PBO framework used to test whether

strategy selection is likely to overfit (Journal of Computational Finance). These sources support the research questions and validation tools, not this implementation's return.

The seven configurations are one engine family, not seven sleeves. Carry-tilted `C_carry` remains in `crypto_carry`; named momentum trials remain in `crypto_momentum`. Shared data, instruments, construction, and latent market exposures prevent those labels from becoming automatic economic independence.

Complete trial lineage

Matrix identity	Change from baseline	Sharpe	Max DD	DSR	DSR gate
`A_blend`	rank baseline	0.0424	13.59%	0.2112	fail
`A_ml`	ML overlay	0.0424	13.59%	0.2112	fail
`A_regime`	regime overlay	0.0424	13.59%	0.2112	fail
`A_ml_regime`	ML plus regime	0.0424	13.59%	0.2112	fail
`C_rebal24`	daily rebalance	-0.7901	33.93%	0.0049	fail
`C_band10`	10-basis-point band	-0.1994	13.89%	0.0935	fail
`C_mvo`	MVO allocator	-0.0946	15.94%	0.1373	fail

The public packet maps each row to its exact hypothesis key, config hash, immutable first ledger record, source file, SHA-256 hash, 1,612 observations, skew, and kurtosis. All seven remain charged to the current 228-identity union. The matrix's own DSR values use its complete eight-trial local selection set and are labeled artifact-era evidence rather than current-union restatements.

What the unchanged overlays mean

The four Block A arms returned byte-for-byte equivalent summary moments. That is not evidence that ML and regimes work equally well. The persisted verdict reports that the regime gate was inactive in early out-of-sample legs while its expanding fit window lacked 730 daily BTC observations and fell back to the blend. The matrix also does not persist enough feature-level diagnostics to prove where the ML path altered positions. The defensible conclusion is that these overlays did not improve the measured return, not that they were validated.

Daily rebalancing was the strongest falsifier: turnover rose to 80.58x, Sharpe fell to -0.7901, and maximum drawdown reached 33.93%. MVO lowered turnover to 15.19x but did not create positive Sharpe. A tighter band also remained negative. Portfolio machinery did not rescue the edge.

Selection and capacity evidence

The eight-column matrix, including the separately classified carry arm, used 5,000 CSCV combinations over 1,612 daily observations. PBO was 0.8818; the required gate was below 0.20. No configuration cleared DSR 0.95 and no deflated winner was selected.

Initial capital	Sharpe	DSR	Max DD	Turnover	Final equity
$100,000	0.4009	0.4803	13.14%	44.53x	$118,509.86
$1,000,000	0.0424	0.2112	13.59%	39.74x	$993,219.39
$10,000,000	-0.3720	0.0460	22.69%	44.31x	$8,043,578.96

The sweep varies capital for the regime-gated engine without adding return identities. Under its notional/ADV impact model, the already thin edge decays materially with size. This does not prove true exchange capacity, the model still needs external calibration, but it decisively forbids a claim of scalable backtested edge.

Live and sleeve decision

The family is not an Alpaca paper sleeve, not an AlphaForge live sleeve, and not included in published broker returns. There is no continuous broker-reconciled forward record for these seven configurations. Historical capacity rows cannot substitute for live slippage, funding, fills, venue outages, or counterparty risk.

Decision: FAIL / NO DEPLOY. The family contributes zero sleeves. Any future attempt must freeze the expanded feature set rather than serialize `null`, preregister one configuration, establish point-in-time universe lineage, use current-union deflation, calibrate impact from observed fills, and accumulate forward evidence. It would be a new charged identity.

Reproduction and authorship

- `scripts/grand_backtest.py` and `src/alphaforge/analytics/grand_matrix.py`: matrix driver and validation logic.
- `artifacts/grand_backtest/20260616T143620Z/matrix.json`: persisted measurements and capacity.
- `artifacts/grand_backtest/20260616T143620Z/verdict.md`: contemporaneous no-deploy decision.
- `artifacts/grand_backtest/20260616T143620Z/experiments.jsonl`: immutable dedicated ledger.
- `scripts/audit_crypto_multifactor_family.py`: deterministic seven-identity audit.
- `/glassbox/crypto_multifactor_family.json`: public machine-readable family packet.
- `/glassbox/trial_packet_manifest.json`: global identity join and incomplete identity-level packet ledger.

Reproduction requires the pinned AlphaForge environment and underlying point-in-time market-data lake. Public hashes make the persisted evidence tamper-evident; they do not make private market data publicly reproducible or the work independently peer reviewed.

This record and the ALPHAC system were authored and directed by **Arhan Canli**, Founder of Canli Capital. Credit is explicit and verifiable in the source and public corpus; scientific claims remain bounded by the evidence.

Conclusion

The multi-factor engine is valuable because it records a failed robustness test cleanly. Complex overlays did not improve the baseline, faster trading made outcomes worse, PBO was high, and modeled capacity decayed. The institutional decision is zero sleeves and continued evidence collection, not a renamed backtest winner.

References

1. YUKUN LIU, ALEH TSYVINSKI, XI WU (2022). *Common Risk Factors in Cryptocurrency*. The Journal of Finance. https://doi.org/10.1111/jofi.13119
2. Yukun Liu, Aleh Tsyvinski (2021). *Risks and Returns of Cryptocurrency*. The Review of Financial Studies. https://doi.org/10.1093/rfs/hhaa113
3. David Bailey, Jonathan Borwein, Marcos López de Prado, Qiji Jim Zhu (2016). *The probability of backtest overfitting*. The Journal of Computational Finance. https://doi.org/10.21314/jcf.2016.322