

---

# Horizon-Dependent Neural GVAR for Interpretable Spillover Analysis in Cryptocurrency Markets

---

Shana Stämpfli<sup>1</sup> Cem Özkan<sup>1</sup> Lucas Poinsignon<sup>1</sup> Athanasios Kakafikas<sup>1</sup>

## 1. Introduction

Cryptocurrency markets exhibit rapid regime shifts, strong co-movement, and volatility cascades that challenge both forecasting and interpretability. Standard deep learning forecasters (e.g., LSTMs) can fit nonlinearities but do not directly expose directed, lagged spillovers across assets. Classical Granger causality offers interpretable directed dependence (1), but in its most common linear VAR form assumes time-invariant relationships and can miss nonlinearities or temporal variation. Our goal is a model that jointly predicts multivariate crypto returns and volatilities, yields an interpretable directed, lagged dependence network, and supports direct tests of domain hypotheses about leverage proxies and technical signals.

A key empirical lesson from our study is that spillover structure varies far more across forecast horizons than across within-horizon states. In particular, effects that are weak or unstable at  $h = 1$  hour can become more consistent at multi-hour or daily horizons, plausibly reflecting slower information diffusion, delayed portfolio rebalancing, and liquidation cascades. While our architecture allows state-conditioned gating, an explicit StaticGVAR ablation (no gating) achieves nearly identical forecasting performance, indicating that under our feature set and evaluation protocol, the dominant variation is horizon dependence rather than strong state dependence.

### Contributions.

- **Horizon-dependent neural GVAR with concept channels.** We train separate models for multiple horizons  $h \in \{1, 4, 12, 24\}$  and recover horizon-specific directed dependence graphs with explicit per-lag edges.
- **StaticGVAR ablation to quantify state dependence.** We implement a StaticGVAR variant (fixed adjacency; no state gating) and show it matches NeuralGVAR forecasting accuracy closely, supporting the interpretation that the model is best summarized as *horizon-dependent*.
- **Hypothesis-driven evaluation with falsification checks.** We test five hypotheses (H1–H4b) using (i) horizon-wise edge sign/stability summaries

(with block-bootstrap confidence intervals), (ii) out-of-sample event conditioning with matched controls and shift-placebos, and (iii) forecasting baselines (Last, VARX–LASSO, LSTM) for context.

**Related work.** Neural approaches to Granger-causal discovery extend linear VARs by learning nonlinear predictors with structured sparsity, enabling directed dependence estimation beyond classical tests (3; 4). In econometrics, time variation is commonly modeled through time-varying VARs (TV-VAR/TV-SVAR), emphasizing that dependencies can drift with regimes and policy or market states (5). Separately, the financial connectedness literature quantifies horizon- and direction-dependent spillovers using variance-decomposition networks (8; 9). Our work connects these threads by extracting interpretable, directed lagged networks from a neural VAR-style forecaster while showing empirically that, under our evaluation protocol, forecast horizon induces the dominant and most reproducible changes in spillover structure.

All code to reproduce the experiments and generate the tables is available at GitHub repository.

## 2. Models and Methods

### 2.1. Data and preprocessing

We use hourly BTCUSDT, ETHUSDT, and SOLUSDT perpetual futures data aligned to an hourly grid (UTC). Because SOL availability begins later (late 2020 on major venues), the analysis is run on backtest windows beginning in 2021. We construct:

- **Log-returns:**  $r_t = \log p_t - \log p_{t-1}$ .
- **Realized volatility:** rolling standard deviation of returns (24-hour window).
- **RSI divergence:** discrete signal  $\in \{-1, 0, +1\}$  indicating directional disagreement between price and RSI trends.
- **Funding rate:** perp funding, forward-filled from its native cadence to the hourly grid.

We define two prediction tasks: **RET** with  $y_t = [r_t^{BTC}, r_t^{ETH}, r_t^{SOL}]$  and **VOL** with  $y_t = [\sigma_t^{BTC}, \sigma_t^{ETH}, \sigma_t^{SOL}]$ . Concept/state inputs  $x_t$

concatenate all symbols' RSI divergence and funding features and are standardized using training data only.

**Rolling backtest windows and multi-horizon setup.** We evaluate three walk-forward windows: (i) train 2021–2022, test 2023; (ii) train 2021–2023, test 2024; (iii) train 2022–2024, test 2025 (to 2025-12-25 23:00). We train/evaluate separate models for each horizon  $h \in \{1, 4, 12, 24\}$  and aggregate across 3 backtest windows  $\times$  5 seeds  $\times$  3 sparsity settings ( $n = 45$  runs per task/horizon).

## 2.2. Horizon-dependent neural GVAR and StaticGVAR

Let  $y_t \in \mathbb{R}^d$  with  $d = 3$ . For lags  $\mathcal{K} = \{1, 3, 6, 12\}$  (hours), we predict

$$\hat{y}_{t+h} = \sum_{k \in \mathcal{K}} A_k^{(h)}(x_t) y_{t-k} + C(x_t) + b. \quad (1)$$

We train separate models per forecast horizon to isolate horizon-specific interaction structures.

**NeuralGVAR (state-gated).** We parameterize adjacency as

$$A_k^{(h)}(x_t) = S_k^{(h)} \odot \sigma(g_k^{(h)}(x_t)), \quad (2)$$

where  $S_k^{(h)}$  is a learned base adjacency (interpretable backbone) and  $g_k^{(h)}$  is an MLP producing gates.

**StaticGVAR (no gating).** The StaticGVAR ablation removes state dependence by fixing

$$A_k^{(h)}(x_t) \equiv S_k^{(h)}. \quad (3)$$

This isolates the contribution of state gating to forecasting and interpretability.

**Training objective.** We minimize the prediction loss with regularization:

$$\begin{aligned} \mathcal{L} &= \text{MSE}(\hat{y}, y) + \lambda_{\text{sparse}} \mathcal{R}_{\text{sparse}} + \lambda_{\text{smooth}} \mathcal{R}_{\text{smooth}} \\ &\quad + \lambda_{\text{concept}} \|W_C\|_1, \\ \mathcal{R}_{\text{sparse}} &= \sum_{i,j} \sqrt{\sum_{k \in \mathcal{K}} \left(S_{k,ij}^{(h)}\right)^2}, \\ \mathcal{R}_{\text{smooth}} &= \mathbb{E}_t \left[ \left\| A^{(h)}(t) - A^{(h)}(t-1) \right\|_F^2 \right]. \end{aligned} \quad (4)$$

The group penalty encourages sparse directed edges (interpretable sparse VAR-like structure) (2), while the smoothness term stabilizes inferred temporal variation.

## 2.3. Concept head

The concept head is linear:

$$C(x_t) = W_C x_t, \quad (5)$$

so concept  $\rightarrow$  output effects are readable from  $W_C$  (sign and magnitude). We treat this primarily as an interpretability mechanism and validate hypotheses through out-of-sample conditioning and falsification checks.

## 2.4. Baselines

We compare to: (i) **Last** persistence ( $\hat{y}_{t+h} = y_t$ ), (ii) **VARX-LASSO** (linear VAR with  $\ell_1$  regularization and exogenous  $x_t$ ), (iii) **LSTM** trained on the same splits, and the ablation **StaticGVAR** (no state gating).

## 2.5. Hypotheses and diagnostics

We test:

- **H1:** RSI divergence  $\rightarrow$  future returns (mean reversion; expected negative effect).
- **H2:** BTC returns  $\rightarrow$  ETH/SOL returns (positive contagion).
- **H3:** volatility  $\rightarrow$  future volatility (positive clustering).
- **H4a:** funding  $\rightarrow$  future volatility (positive leverage-to-volatility).
- **H4b:** BTC volatility  $\rightarrow$  ETH/SOL volatility (positive spillover).

**Lag aggregation comparisons.** For compact reporting, we summarize a directed edge at time  $t$  using the *lag-aggregated* adjacency  $\bar{A}^{(h)}(t) = \sum_{k \in \mathcal{K}} A_k^{(h)}(t)$ .

**Edge hypotheses (H2/H3/H4b).** From  $\bar{A}^{(h)}(t)$  we compute, per run, (i) the fraction of test time where an edge is positive (sign fraction), and (ii) the median edge value. We estimate uncertainty using a moving block bootstrap over time (Appendix B).

**Concept hypotheses (H1/H4a).** We report: (i) concept head weights  $W_C$  (mean  $\pm$  std across runs), and (ii) **outcome conditioning:**  $\Delta = \mathbb{E}[y_{t+h} \mid \text{concept extreme}] - \mathbb{E}[y_{t+h}]$ .

**Event conditioning and placebos.** We define events as top 1% BTC return jumps (RET) or top 1% BTC volatility spikes (VOL), compute post-pre changes in edge strength, and compare to matched controls, a  $-6h$  shift placebo, and random events (Appendix C).

## 3. Results

### 3.1. Forecasting performance

Test MSE across horizons is shown in Figure 1 and 2. StaticGVAR and NeuralGVAR are extremely close, indicating limited incremental predictive value from state gating under our setup. Returns are hard to predict: all models perform similarly, except for Last, which performs substantially worse. Volatility is highly persistent at  $h = 1$ , favoring Last; at longer horizons, structured models become competitive.

All bars aggregate over 3 backtest windows  $\times$  5 seeds  $\times$  3 sparsity settings ( $n = 45$  runs per task/horizon).

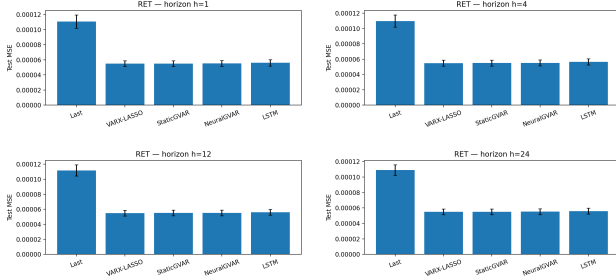


Figure 1: Test MSE for returns (RET) across horizons.

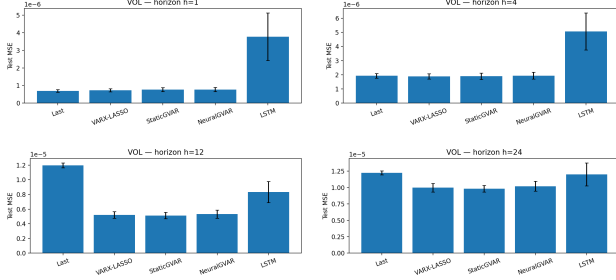


Figure 2: Test MSE for volatility (VOL) across horizons.

### 3.2. Hypothesis tests (H1–H4b)

#### H1: RSI divergence $\rightarrow$ returns (mean reversion).

Across horizons, RSI-divergence conditioning effects are small and not consistently negative, indicating weak evidence for robust mean reversion under this pipeline.

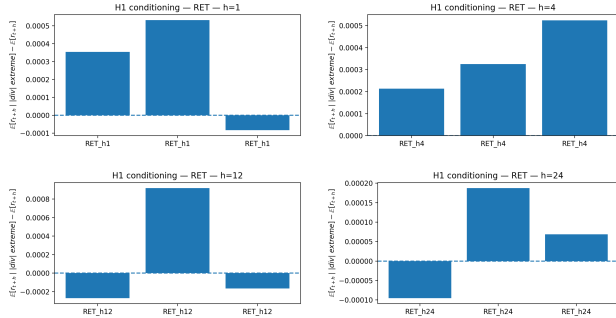


Figure 3: H1: outcome conditioning for RSI divergence across horizons:  $\mathbb{E}[r_{t+h} \mid |div| \text{ extreme}] - \mathbb{E}[r_{t+h}]$ .

#### H4a: funding $\rightarrow$ future volatility (leverage proxy).

Funding conditioning effects are consistently positive across assets and horizons, with particularly strong increases for SOL in our experiments. Figure 4 summarizes the horizon dependence; all bars aggregate over 3 backtest windows  $\times$  5 seeds  $\times$  3 sparsity settings ( $n = 45$  runs per task/horizon).

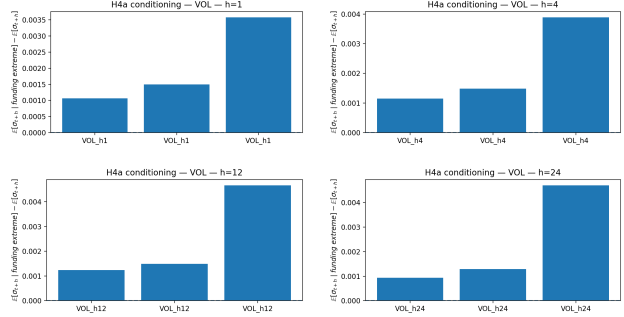


Figure 4: H4a: outcome conditioning under extreme funding across horizons:  $\mathbb{E}[\sigma_{t+h} \mid \text{funding extreme}] - \mathbb{E}[\sigma_{t+h}]$ .

**H2/H3/H4b (edge-based hypotheses).** Using the lag-aggregated adjacency  $\bar{A}^{(h)}(t)$ , we summarize horizon-wise sign stability for key spillover edges in Figure 5. All lines aggregate over  $n = 45$  runs per horizon. Full horizon-wise edge tables with block-bootstrap confidence intervals are in Appendix B.

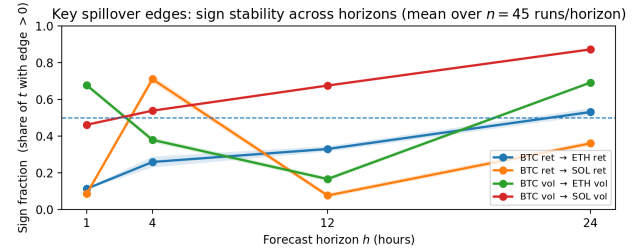


Figure 5: Sign stability of key spillover edges across horizons. Lines show mean sign fraction over  $n = 45$  runs per horizon; the dashed line marks 0.5 (equal time positive vs. negative).

#### H2: BTC returns $\rightarrow$ ETH/SOL returns (positive contagion).

Return contagion is not an immediate ( $h = 1$ ) effect in this dataset: BTC $\rightarrow$ ETH and BTC $\rightarrow$ SOL are rarely positive at  $h = 1$  and have negative medians. However, the effect becomes horizon-dependent. BTC $\rightarrow$ SOL is frequently positive at  $h = 4$  (sign fraction  $\approx 0.71$  with a positive median), whereas BTC $\rightarrow$ ETH becomes closer to balanced only at  $h = 24$  (sign fraction  $\approx 0.53$  with a slightly positive median). Two interpretations are consistent with these patterns: (i) cross-asset return co-movement may require time to materialize through flows and delayed positioning, or (ii) the hourly horizon is dominated by microstructure noise that washes out weak directed effects, while multi-hour horizons reveal more stable co-movement.

**H3: volatility  $\rightarrow$  future volatility (clustering).** H3 is strongly supported. Own-volatility edges are positive essentially always in the exports (sign fractions near 1.0 across horizons), consistent with classical volatility clustering (6; 7). This is also reflected in the forecasting results: Last is highly competitive at  $h = 1$  for VOL, where persis-

tence captures much of the signal.

#### H4b: BTC volatility $\rightarrow$ ETH/SOL volatility spillovers.

Volatility spillovers show the clearest horizon dependence. BTC $\rightarrow$ SOL volatility spillover grows steadily with horizon, becoming strongly positive by  $h = 24$  (high sign fraction and positive median). BTC $\rightarrow$ ETH is more nuanced: it is positive at  $h = 1$ , turns negative at intermediate horizons ( $h = 4, 12$ ), and returns toward balance at  $h = 24$ . This non-monotonicity suggests that ETH may absorb BTC volatility shocks quickly (short horizon), while intermediate horizons reflect mean reversion or hedging/liquidity effects; in contrast, smaller-cap SOL shows stronger delayed propagation consistent with slower liquidity-driven transmission.

**Event conditioning and placebos.** Event-based diagnostics highlight why placebos are necessary: shift-placebo effects can be comparable to real-event effects, cautioning against causal timing interpretations from short-horizon post-event changes alone. We therefore emphasize horizon-wise stability summaries as the primary result; event-conditioning details are provided in Appendix C.

## 4. Discussion and implications

### 4.1. What the StaticGVAR result means

The StaticGVAR ablation provides a clear constraint on interpretation. Because StaticGVAR matches NeuralGVAR forecasting performance closely across tasks and horizons, state-gated adjacency does not yield robust out-of-sample gains under our feature set and evaluation protocol. This implies either that true state dependence is weak at the hourly scale for these assets, or that it is not well captured by our current state variables (RSI divergence, funding) or gating parameterization.

Importantly, this is not a negative result for interpretability. The most reproducible signal is the horizon-specific structure: edges that remain stable across runs at a given horizon. State gating may still be useful for local or regime-specific explanation, but such effects should be treated as secondary unless they survive stricter stability tests.

### 4.2. Horizon-specific graphs as a practical representation

A common assumption in Granger-style analysis is that a single directed graph summarizes the system. Our results instead support a family of graphs indexed by forecast horizon  $h$ , which aligns naturally with how market participants operate. Three horizon-dependent patterns emerge:

- (i) Return spillovers are weak and unstable at  $h = 1$ , while propagation and clustering become more consistent at longer horizons;
- (ii) the narrative “BTC leads alts” does not appear as an immediate return effect, but emerges at intermediate horizons

(notably BTC $\rightarrow$ SOL at  $h = 4$ );

- (iii) BTC $\rightarrow$ SOL volatility spillovers strengthen toward  $h = 24$ , consistent with delayed transmission via liquidity and leverage channels.

This horizon-indexed view discourages over-interpretation of time-varying edges at a single timescale and emphasizes robustness at the relevant horizon.

### 4.3. Concept signals: funding vs. RSI divergence

Funding extremes show consistent positive conditioning for future volatility, supporting a leverage-based interpretation: crowded positioning increases vulnerability to volatility cascades. That conditioning effects are robust even when  $W_C$  weights are mixed highlights why interpretability should not be reduced to reading linear coefficients alone.

In contrast, RSI divergence in the discretized form used here does not yield robust mean-reversion evidence out of sample. This may reflect discretization loss, missing regime information, or rapid arbitrage of simple technical signals at the hourly scale.

### 4.4. Limitations and future work

This study is intentionally narrow (three assets, hourly data, limited state features). Several extensions could strengthen structural interpretation:

- (i) **Per-lag attribution** to identify propagation delays;
- (ii) **Richer state variables** (e.g., open interest, liquidation proxies, liquidity measures) to better test state dependence;
- (iii) **Broader asset universes** to assess heterogeneity and delayed spillovers; and
- (iv) **Joint multi-horizon training** with horizon-specific heads to share information while preserving interpretability.

## 5. Summary

We evaluated a horizon-dependent neural GVAR with interpretable concept channels (RSI divergence, funding) on hourly BTC/ETH/SOL perpetual futures using rolling 2023–2025 out-of-sample windows and repeated-run aggregation ( $n = 45$  per task/horizon). Across tasks, the dominant systematic variation is across forecast horizons, and a StaticGVAR ablation indicates limited incremental predictive value from state gating under this feature set and protocol. Funding extremes reliably precede higher volatility (especially for SOL), while RSI divergence does not yield robust mean-reversion effects in returns. Edge summaries suggest spillovers are horizon-specific: return contagion is weak at  $h = 1$  but can emerge at multi-hour horizons, and volatility spillovers (especially BTC $\rightarrow$ SOL) strengthen toward  $h = 24$ . Overall, the approach provides an interpretable framework for multi-horizon spillover analysis in crypto that complements standard forecasting baselines.

## A. Full forecasting metrics tables

In the Table 1, SignHit denotes the fraction of correctly predicted signs, averaged over the three target components (BTC/ETH/SOL) and all test timestamps for the given task and horizon:  $\frac{1}{d} \sum_{i=1}^d \mathbb{E}_t [\mathbb{1}\{\text{sign}(\hat{y}_{t+h,i}) = \text{sign}(y_{t+h,i})\}]$  with  $d = 3$ .

task	h	model	MSE_mean	MSE_std	MAE_mean	MAE_std	SignHit_mean	SignHit_std	n
RET_h1	1	VARX-LASSO	5.4943e-05	3.76484e-06	0.00463461	0.000384148	0.507089	0.0048293	45
RET_h1	1	StaticGVAR	5.50261e-05	3.79098e-06	0.00464199	0.000383277	0.501226	0.00469808	45
RET_h1	1	NeuralGVAR	5.52753e-05	3.80134e-06	0.00465447	0.00038175	0.501542	0.00244376	45
RET_h1	1	LSTM	5.60274e-05	4.10451e-06	0.00473181	0.000417927	0.499789	0.00642153	45
RET_h1	1	Last	0.000110712	8.72952e-06	0.00684806	0.000534155	0.481683	0.00274746	45
RET_h12	12	VARX-LASSO	5.49227e-05	3.76794e-06	0.00463736	0.000383857	0.502132	0.0034622	45
RET_h12	12	StaticGVAR	5.50843e-05	3.80069e-06	0.00464744	0.000383711	0.497047	0.00458597	45
RET_h12	12	NeuralGVAR	5.51589e-05	3.78683e-06	0.00465466	0.000383632	0.498808	0.00400369	45
RET_h12	12	LSTM	5.60339e-05	3.88593e-06	0.00473387	0.000364987	0.50057	0.00460378	45
RET_h12	12	Last	0.000111773	7.45878e-06	0.00687894	0.000521119	0.495704	0.00300708	45
RET_h24	24	VARX-LASSO	5.49723e-05	3.76094e-06	0.00464033	0.000378856	0.488774	0.0125942	45
RET_h24	24	StaticGVAR	5.50341e-05	3.74137e-06	0.00464525	0.000376587	0.497085	0.00418272	45
RET_h24	24	NeuralGVAR	5.52226e-05	3.74448e-06	0.00465557	0.000376491	0.498086	0.00312926	45
RET_h24	24	LSTM	5.60057e-05	3.8666e-06	0.00470879	0.000377495	0.499913	0.00444993	45
RET_h24	24	Last	0.000109034	6.79706e-06	0.00681353	0.000525563	0.506564	0.00830891	45
RET_h4	4	VARX-LASSO	5.49018e-05	3.80918e-06	0.00463492	0.000385018	0.501736	0.00223215	45
RET_h4	4	StaticGVAR	5.49995e-05	3.83498e-06	0.00464126	0.000384662	0.500962	0.0032427	45
RET_h4	4	NeuralGVAR	5.51806e-05	3.78251e-06	0.00465338	0.000382051	0.500161	0.00261766	45
RET_h4	4	LSTM	5.64762e-05	3.93909e-06	0.00475209	0.000375323	0.500136	0.00636262	45
RET_h4	4	Last	0.00010979	7.94018e-06	0.00686403	0.000527287	0.491735	0.0016819	45
VOL_h1	1	VARX-LASSO	7.26637e-07	8.68815e-08	0.000453892	2.15923e-05	1	0	45
VOL_h1	1	StaticGVAR	7.68076e-07	1.12018e-07	0.000456991	1.90179e-05	1	0	45
VOL_h1	1	NeuralGVAR	7.73841e-07	1.15633e-07	0.000456164	2.09149e-05	0.999939	0.000123142	45
VOL_h1	1	LSTM	3.77098e-06	1.34708e-06	0.00140074	0.000286267	0.999964	4.95202e-05	45
VOL_h1	1	Last	6.95514e-07	7.69459e-08	0.000375807	1.08693e-05	1	0	45
VOL_h12	12	VARX-LASSO	5.21691e-06	4.46338e-07	0.00159026	7.93267e-05	1	0	45
VOL_h12	12	StaticGVAR	5.129e-06	4.40682e-07	0.00154563	5.1144e-05	1	0	45
VOL_h12	12	NeuralGVAR	5.3206e-06	5.5449e-07	0.00158109	9.05574e-05	0.999864	0.000249059	45
VOL_h12	12	LSTM	8.31757e-06	1.43722e-06	0.00210624	0.000260042	0.99997	7.03462e-05	45
VOL_h12	12	Last	1.19516e-05	3.32353e-07	0.00240238	5.48698e-05	1	0	45
VOL_h24	24	VARX-LASSO	9.9807e-06	6.54738e-07	0.00232941	0.000136339	1	0	45
VOL_h24	24	StaticGVAR	9.80933e-06	5.03664e-07	0.00228329	8.06189e-05	1	0	45
VOL_h24	24	NeuralGVAR	1.01959e-05	7.62273e-07	0.00231629	0.000141137	0.99962	0.000689695	45
VOL_h24	24	LSTM	1.19961e-05	1.74342e-06	0.0026132	0.000307517	1	0	45
VOL_h24	24	Last	1.22213e-05	3.0521e-07	0.00243177	5.82396e-05	1	0	45
VOL_h4	4	VARX-LASSO	1.89005e-06	1.91484e-07	0.000829107	2.49333e-05	1	0	45
VOL_h4	4	StaticGVAR	1.90101e-06	2.24469e-07	0.000816685	2.56801e-05	1	0	45
VOL_h4	4	NeuralGVAR	1.94044e-06	2.3738e-07	0.000824306	3.1462e-05	0.999925	0.000149917	45
VOL_h4	4	LSTM	5.06276e-06	1.30417e-06	0.00161624	0.000250637	0.999916	0.000108863	45
VOL_h4	4	Last	1.93309e-06	1.63417e-07	0.000758922	1.75966e-05	1	0	45

Table 1: Full forecasting metrics across tasks, horizons, and models (generated by the notebook). All entries aggregate over 3 backtest windows  $\times$  5 seeds  $\times$  3 sparsity settings ( $n = 45$  runs per task/horizon).

## B. Edge summary tables and uncertainty

We estimate uncertainty for sign fractions and medians using a moving block bootstrap over the test window to preserve short-range dependence.

## Horizon-Dependent Neural GVAR for Interpretable Spillover Analysis in Cryptocurrency Markets

task	h	hyp	sign_frac_mean	sign_frac_std	sign_ci_lo_mean	sign_ci_hi_mean	median_mean	med_ci_lo_mean	med_ci_hi_mean	n
RET_h1	1	H2 BTCret→ETHUSDret (+)	0.114096	0.299686	0.111019	0.117347	-0.0778514	-0.0789101	-0.0766728	45
RET_h1	1	H2 BTCret→SOLUSDret (+)	0.0864893	0.212284	0.0777635	0.0959422	-0.0748356	-0.0765113	-0.0732329	45
RET_h12	12	H2 BTCret→ETHUSDret (+)	0.330277	0.409137	0.315245	0.345613	-0.0171803	-0.0193523	-0.0152129	45
RET_h12	12	H2 BTCret→SOLUSDret (+)	0.0772459	0.163987	0.0647611	0.0893665	-0.032749	-0.0350237	-0.0306843	45
RET_h24	24	H2 BTCret→ETHUSDret (+)	0.53109	0.436477	0.511785	0.550794	0.00207765	0.000967726	0.00319324	45
RET_h24	24	H2 BTCret→SOLUSDret (+)	0.360626	0.430695	0.34618	0.375278	-0.00436621	-0.00492286	-0.00382797	45
RET_h4	4	H2 BTCret→ETHUSDret (+)	0.25957	0.287267	0.229593	0.288792	-0.011742	-0.0133784	-0.0105857	45
RET_h4	4	H2 BTCret→SOLUSDret (+)	0.710046	0.374511	0.689917	0.730166	0.00802441	0.00699064	0.00878523	45
VOL_h1	1	H3 BTCUSDtvol→BTCUSDtvol (+)	1	0	1	1	0.917695	0.917488	0.91793	45
VOL_h1	1	H3 ETHUSDtvol→ETHUSDtvol (+)	1	0	1	1	0.934896	0.934703	0.935126	45
VOL_h1	1	H3 SOLUSDtvol→SOLUSDtvol (+)	1	0	1	1	0.951694	0.951323	0.951974	45
VOL_h1	1	H4b BTCvol→ETHUSDtvol (+)	0.676986	0.430217	0.666127	0.687766	0.0136991	0.0135038	0.0138941	45
VOL_h1	1	H4b BTCvol→SOLUSDtvol (+)	0.462086	0.489374	0.457223	0.467218	-0.0064321	-0.00686715	-0.00614116	45
VOL_h12	12	H3 BTCUSDtvol→BTCUSDtvol (+)	0.99967	0.00083128	0.999444	0.999851	0.491234	0.48924	0.493766	45
VOL_h12	12	H3 ETHUSDtvol→ETHUSDtvol (+)	1	0	1	1	0.764158	0.763314	0.765353	45
VOL_h12	12	H3 SOLUSDtvol→SOLUSDtvol (+)	1	0	1	1	0.734632	0.734051	0.73519	45
VOL_h12	12	H4b BTCvol→ETHUSDtvol (+)	0.166778	0.33266	0.155717	0.178339	-0.059289	-0.0612836	-0.0569678	45
VOL_h12	12	H4b BTCvol→SOLUSDtvol (+)	0.675209	0.464757	0.671452	0.678954	-0.0153359	-0.0160904	-0.0147431	45
VOL_h24	24	H3 BTCUSDtvol→BTCUSDtvol (+)	1	0	1	1	0.223437	0.221675	0.225811	45
VOL_h24	24	H3 ETHUSDtvol→ETHUSDtvol (+)	1	0	1	1	0.563	0.560083	0.566639	45
VOL_h24	24	H3 SOLUSDtvol→SOLUSDtvol (+)	1	0	1	1	0.544307	0.542909	0.545447	45
VOL_h24	24	H4b BTCvol→ETHUSDtvol (+)	0.691135	0.434977	0.683016	0.699156	-0.000228527	-0.00102324	0.000397367	45
VOL_h24	24	H4b BTCvol→SOLUSDtvol (+)	0.871746	0.330847	0.869671	0.873666	0.00214281	0.00147848	0.00242895	45
VOL_h4	4	H3 BTCUSDtvol→BTCUSDtvol (+)	1	0	1	1	0.819016	0.818424	0.819582	45
VOL_h4	4	H3 ETHUSDtvol→ETHUSDtvol (+)	1	0	1	1	0.887813	0.887606	0.888166	45
VOL_h4	4	H3 SOLUSDtvol→SOLUSDtvol (+)	1	0	1	1	0.896443	0.895989	0.896854	45
VOL_h4	4	H4b BTCvol→ETHUSDtvol (+)	0.379786	0.440528	0.366035	0.393544	-0.00387722	-0.00435032	-0.00348933	45
VOL_h4	4	H4b BTCvol→SOLUSDtvol (+)	0.538453	0.498277	0.5354	0.542182	-0.00903136	-0.00948535	-0.00876462	45

Table 2: H2/H3/H4b edge summaries by horizon (generated by the notebook). All entries aggregate over 3 backtest windows  $\times 5$  seeds  $\times 3$  sparsity settings ( $n = 45$  runs per task/horizon).

## C. Event conditioning protocol and additional results

model	task	h	edge	event_mode	events_n_mean	real_mean	real_std	ctrl_mean	ctrl_std	shift_mean	shift_std	rand_mean	rand_std	n
NeuralGVAR	RET_h1	1	BTCUSDt→ETHUSDt	jump	87.6667	0.000737723	0.0030201	-0.00143643	0.00198682	0.00145281	0.00283042	-0.000392244	0.00172207	45
StaticGVAR	RET_h1	1	BTCUSDt→ETHUSDt	jump	87.6667	0	0	-1.88146e-12	1.26212e-11	0	0	0	0	45
NeuralGVAR	RET_h12	12	BTCUSDt→ETHUSDt	jump	87.6667	-0.000348033	0.00340875	4.66658e-05	0.00245194	0.000788509	0.00351623	0.00048298	0.00284097	45
StaticGVAR	RET_h12	12	BTCUSDt→ETHUSDt	jump	87.6667	0	0	-1.76387e-13	8.7388e-13	0	0	0	9.12913e-12	45
NeuralGVAR	RET_h24	24	BTCUSDt→ETHUSDt	jump	87.3333	-0.000120667	0.00188511	-0.000610284	0.00237923	0.00148587	0.00218588	0.000586713	0.00105095	45
StaticGVAR	RET_h24	24	BTCUSDt→ETHUSDt	jump	87.3333	0	0	4.70365e-13	3.1553e-12	0	0	0	0	45
NeuralGVAR	RET_h4	4	BTCUSDt→ETHUSDt	jump	87.6667	0.000140709	0.0015454	-0.000491607	0.00290424	0.00313783	0.00321621	0.000911299	0.00238558	45
StaticGVAR	RET_h4	4	BTCUSDt→ETHUSDt	jump	87.6667	0	0	0	4.5127e-12	0	0	-4.75771e-13	2.23099e-12	45
NeuralGVAR	VOL_h1	1	BTCUSDt→ETHUSDt	vol	87.6667	-0.000217396	0.000969202	6.5062e-05	0.000857546	-0.000325381	0.0017434	-0.000104981	0.000618744	45
StaticGVAR	VOL_h1	1	BTCUSDt→ETHUSDt	vol	87.6667	0	0	0	0	0	0	0	0	45
NeuralGVAR	VOL_h1	1	BTCUSDt→SOLUSDt	vol	87.6667	-0.000270329	0.000703715	-5.24643e-05	0.000300548	-0.00034688	0.000904551	-2.392e-05	0.000323875	45
StaticGVAR	VOL_h1	1	BTCUSDt→SOLUSDt	vol	87.6667	0	0	0	0	0	0	0	0	45
NeuralGVAR	VOL_h12	12	BTCUSDt→ETHUSDt	vol	87.6667	-0.00281359	0.00366897	0.00175775	0.0028583	-0.00403123	0.00583831	0.000149587	0.00223301	45
StaticGVAR	VOL_h12	12	BTCUSDt→ETHUSDt	vol	87.6667	0	0	0	0	0	0	1.90309e-12	1.27663e-11	45
NeuralGVAR	VOL_h12	12	BTCUSDt→SOLUSDt	vol	87.6667	-0.00190001	0.00313248	-0.000899943	0.00254321	-0.00139973	0.00387565	-0.00040848	0.00139048	45
StaticGVAR	VOL_h12	12	BTCUSDt→SOLUSDt	vol	87.6667	0	0	0	0	0	0	1.90309e-12	1.27663e-11	45
NeuralGVAR	VOL_h24	24	BTCUSDt→ETHUSDt	vol	87.3333	-0.0021743	0.00316711	-0.00119699	0.00243664	-0.000526106	0.00563246	-0.000373948	0.00191932	45
StaticGVAR	VOL_h24	24	BTCUSDt→ETHUSDt	vol	87.3333	0	0	0	0	0	0	0	0	45
NeuralGVAR	VOL_h24	24	BTCUSDt→SOLUSDt	vol	87.3333	-0.00290121	0.00483032	-0.00161687	0.00474036	-0.00105084	0.00371443	-0.00107279	0.00276967	45
StaticGVAR	VOL_h24	24	BTCUSDt→SOLUSDt	vol	87.3333	0	0	0	0	0	0	0	0	45
NeuralGVAR	VOL_h4	4	BTCUSDt→ETHUSDt	vol	87.6667	-0.00101761	0.00143103	0.000653866	0.00175054	-0.00126461	0.00293627	4.91262e-05	0.00128824	45
StaticGVAR	VOL_h4	4	BTCUSDt→ETHUSDt	vol	87.6667	0	0	0	0	0	0	-9.51543e-13	6.38314e-12	45
NeuralGVAR	VOL_h4	4	BTCUSDt→SOLUSDt	vol	87.6667	-0.000813219	0.000987322	-0.00012213	0.000906912	-0.00100163	0.00145118	-3.23067e-05	0.000701053	45
StaticGVAR	VOL_h4	4	BTCUSDt→SOLUSDt	vol	87.6667	0	0	0	0	0	0	0	0	45

Table 3: Event conditioning summary with matched controls, shift placebo, and random events (generated by the notebook). All entries aggregate over 3 backtest windows  $\times 5$  seeds  $\times 3$  sparsity settings ( $n = 45$  runs per task/horizon).

## References

- [1] C. W. J. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3), 1969.
- [2] W. B. Nicholson, I. Wilms, J. Bien, and D. S. Matteson. High dimensional forecasting via interpretable vector autoregression. *Journal of Machine Learning Research*, 21(166):1–52, 2020.
- [3] A. Tank, I. Covert, N. Foti, A. Shojaie, and E. Fox. Neural Granger causality for nonlinear time series. *arXiv:1802.05842*, 2018.
- [4] R. Marcinkevičs and J. Vogt. Interpretable and robust neural Granger causality. *arXiv:2106.16087*, 2021.
- [5] G. E. Primiceri. Time varying structural vector autoregressions and monetary policy. *The Review of Economic Studies*, 72(3), 2005.
- [6] R. F. Engle. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4), 1982.
- [7] T. Bollerslev. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31(3), 1986.
- [8] F. X. Diebold and K. Yilmaz. Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*, 28(1), 2012.
- [9] F. X. Diebold and K. Yilmaz. On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, 182(1), 2014.