CROSS-ASSET GRAPHS IN REINFORCEMENT-LEARNING TRADING: DERIVATION, VALIDATION, DECISION-RELEVANCE AND REPORTING

Anonymous authors
Paper under double-blind review

ABSTRACT

Reinforcement-learning traders increasingly condition on a graph over assets and report gains from doing so. Two questions go unasked: nobody applies multiple-testing control to the thousands of candidate edges a graph is selected from, and nobody reports how much the graph changes what the agent *does*. This paper addresses both. For a mean–variance-with-costs reward under an approximate factor model $\Sigma = \sigma_m^2 \beta\beta^\top + \Sigma_u$, it proves that the reward decomposes *exactly* into per-asset terms, pairwise terms supported precisely on $\mathrm{supp}_{\mathrm{off}}(\Sigma_u)$, and one rank-one global term, so the validated graph is not an architectural choice but the coupling structure the reward already has. It also bounds the value error by the false-discovery rate of edge selection and fixes the construction's scope by a third-order criterion. On 116 US large caps over 3,772 trading days that coupling changes a median of **0** of 116 actions, and forty pre-registered studies leave it intact, including two regimes built to revive it, a second architecture, and a published system whose released code was run. **The mechanism is measurable and does not depend on the agent**: the graph's information is *contemporaneous* at daily resolution, with out-of-sample R^2 of 0.0907 same-day against 0.0030 for a degree-matched random graph, but -0.0010 at one day's lag, for every structure tried. Same-day co-movement is a *risk* fact, and risk enters the decision marginal one order in position size below alpha and cost. That diagnosis names a remedy, which the paper tests: in an objective with *no return term*, the same edges as sparse precision support beat a density-matched random support out of sample by -10.42% in realised volatility. A pre-registered adversarial pass that nets out sector structure retains 48% of that effect and 27% on log-likelihood, and the retained figures are the ones claimed. This research recommends three numbers, which no graph-RL paper currently reports: the count of actions the graph changes, the fraction of decisions beyond *any* graph's reach, and the number of seeds the comparison can resolve (19,261 here against a cap of 100).

1 INTRODUCTION

Portfolio management is a sequential decision problem whose objects are not independent: two airlines move together, an airline and a fuel producer move apart. It is natural to give a reinforcement-learning agent a *graph* over assets, in which nodes are stocks and edges mark related pairs, and a growing literature does so (Wang et al., 2021; Zhang, 2026; Soleymani & Paquet, 2021; Kim et al., 2019), crediting it with better risk-adjusted returns. This work set out to strengthen that programme by placing a *statistically validated* graph in the value function rather than the feature encoder, and found two questions that the programme leaves unasked.

Gap 1: the edges are never validated. N assets offer $N(N-1)/2$ candidate pairs (6,670 here), so an uncorrected screen at the 5% level admits about 334 pairs by chance alone. Wang et al. (2021) compare four graph constructions and conclude a PC-algorithm causal structure (Spirtes et al., 2001) is best; a full-text search returns no occurrence of *FDR*, *p-value*, *multiple testing* or *significance*

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

level. The machinery exists (Benjamini & Hochberg, 1995; Benjamini & Yekutieli, 2001; Uematsu & Yamagata, 2024), but in a literature that graph-RL trading does not cite.

Gap 2: nobody reports how much the graph changes the policy. Evidence that a graph "works" is a performance ablation, which cannot separate *the graph matters* from *seed noise*, and the closest antecedent reports no seed count. The direct question needs no seeds, and nobody reports it.

What this paper does, and what it finds. The graph is *derived* rather than chosen: under an approximate factor model the reward decomposes exactly into per-asset terms, pairwise terms supported precisely on the residual covariance's support, and one rank-one global term (Prop. 1), so the validated graph *is* the reward's own coupling structure. The construction then does nothing, for a reason that is *measured* rather than asserted and that names its own remedy; most of the paper closes the rival explanations.

Contributions.

1. **An exact decomposition** (Prop. 1) putting validated edges where the reward says they belong, with an FDR bound on the value error (Prop. 4) and a scope fixed by a third-order criterion (Prop. 2).
2. **An edge pipeline assuming nothing about dependence**: Benjamini–Hochberg's positive-dependency condition is *measured* implausible here (15.7% of statistic pairs below -0.20), and Benjamini–Yekutieli costs 41% of the edges and no conclusion.
3. **A negative result with a measured mechanism**: the graph's information is *contemporaneous* at daily resolution, so it enters as risk, one order *in position size* below alpha (§6.3).
4. **The remedy this implies**: with no return term, the same edges as sparse precision support beat a density-matched random support out of sample, at a size deflated by the paper's own adversarial pass (§6.7).
5. **Forty pre-registered studies** (Table 6), plus a reliability audit finding learned edges *less* cross-seed consistent than untrained ones.

2 Related Work

Graph-conditioned RL for trading. In every system examined here, the graph conditions a *representation*: a convolution or attention layer mixes an asset's features with its neighbours', and the result feeds a policy. Wang et al. (2021) is the closest antecedent; Sood et al. (2023) name a comparison flaw that this paper avoids: *"there is usually a discrepancy between the reward function used to train the RL agent, and the objective function used for MVO"*. Across the six systems tabulated in Appendix A, **none places the graph in a value function that takes an action argument, none applies multiple-testing control to edge selection, and none reports how many actions the graph changes**.

Statistically validated networks, and coordination graphs. Curme et al. (2015) filter lagged correlation networks with multiple-testing correction, but their synchronous variant validates *"well over 4,500 of the 4,950 possible links"*, which is too permissive to sparsify anything, and Bennett et al. (2022) supplies no multiple-testing procedure at all. Uematsu & Yamagata (2024) give FDR-controlled discovery in high-dimensional VARs; this paper adopts only the FDR framing, since lead–lag relations enter the conditional *mean* (Remark 1). On the RL side, Guestrin et al. (2001) give exact joint maximisation at a cost *"exponential in the induced width of the coordination graph"*; Böhmer et al. (2020) scale it with max-plus, whose convergence *"is only guaranteed for acyclic CGs"*; and Qu et al. (2020) prove exponential decay of Q in graph distance under assumptions that the present setting violates. A separate line decomposes the *reward*: van Seijen et al. (2017) split R into per-head components summed linearly, noting *"there are infinitely many different decompositions of a reward function possible"*, which makes the decomposition a design choice, whereas Proposition 2 asks when one of a given form exists at all.

3 PROBLEM SETUP

The decision problem. For N assets, $r_t \in \mathbb{R}^N$ holds period-t log returns, $\mu = \mathbb{E}[r_t]$, $\Sigma = \mathrm{Cov}(r_t)$; a portfolio is $w \in \mathbb{R}^N$ (w_i = fraction of capital in asset i, negative for a short), with *gross exposure* $g = \sum_i |w_i|$ and w_{prev} last period's holding. For symmetric M, $\mathrm{supp}_{\mathrm{off}}(M) = \{(i,j) : i < j,\ M_{ij} \neq 0\}$ is its off-diagonal support, a set of undirected edges. The MDP is finite-horizon; the state holds, per asset, technical indicators (MACD, RSI, CCI, ADX), recent returns, realised volatility and the current position. The action is *joint*: each a_i picks a level $w_i = L[a_i]$ from $L = \{-1, 0, +1\} \cdot g/N$, so a full-tilt book has gross exposure exactly g. Leverage is a free parameter, not a detail: $L = \{-1, 0, +1\}$ taken literally is 116× gross, a long book at 119% *daily* volatility, so g is swept.

The reward. The per-period reward is mean–variance (Markowitz, 1952) with proportional transaction costs, the covariance in approximate-factor form (Chamberlain & Rothschild, 1983; Fan et al., 2013):

$$r(w) = \underbrace{w^\top \mu}_{\text{alpha}} - \underbrace{\tfrac{\gamma}{2} w^\top \Sigma w}_{\text{risk}} - \underbrace{c\|w - w_{\mathrm{prev}}\|_1}_{\text{cost}}, \qquad \Sigma = \sigma_m^2 \beta\beta^\top + \Sigma_u, \tag{1}$$

with $\gamma > 0$ the risk aversion, c the proportional cost, β the market loadings and Σ_u the residual covariance; write $E^* = \mathrm{supp}_{\mathrm{off}}(\Sigma_u)$ and $\widehat{E}$ for its estimate. **The paper turns on Σ versus Σ_u**: Σ is dense because every equity loads on the market, while Σ_u is what remains once that driver is removed, and it is Σ_u that becomes a graph (Appendix D).

Assumption 1 (Standing assumptions, each used at one identified place)**. (A1) Factor structure, approximately sparse residual** in the POET sense (Fan et al., 2013), with a *single* pervasive factor, which is what collapses the global term to rank one; five nested factors change no conclusion (D1/E23). Proposition 1 needs only the *factor form*; sparsity makes the graph small enough to be useful. Factor–residual separation is *identified* only asymptotically in N (Chamberlain & Rothschild, 1983): Proposition 1 is an identity at any N, but which part of Σ is pervasive is pinned down only as N grows, so at $N = 116$ the loadings stay N-dependent. D1 and E23 vary the split; nothing moves. **(A2) Separable transaction cost**, which holds provided no cross-asset budget or leverage constraint binds; §6.4 builds both regimes in which it fails. **(A3) No dependence condition is assumed.** Benjamini–Hochberg (Benjamini & Hochberg, 1995) needs independence or positive regression dependency (Benjamini & Yekutieli, 2001). Neither is assumed here: the dependence among the test statistics is *measured* (§4.1), found to be materially negative, and handled by Benjamini–Yekutieli, valid under arbitrary dependence; p-value validity is a separate matter, measured in Gate F.

4 METHOD

The method has three parts, namely edge inference, a factored action-value function whose pairwise terms sit on the inferred graph, and a joint-action selector, drawn end to end in Figure 2 (Appendix N).

4.1 EDGE INFERENCE

Four steps, one graph per estimation window (Algorithm 2). *(1)* The market is removed by a per-asset *time-series* OLS of returns on the equal-weighted factor, leaving residuals u. *(2)* All $m = N(N-1)/2$ residual pairs are tested with Fisher's z, whose normality assumption real returns violate; Gate F (Appendix L) calibrates the realised false-discovery proportion under kurtosis-matched and GARCH errors rather than trusting the nominal level. *(3)* Multiplicity is corrected at level q. *(4)* Pairs discovered in at least 3 of the 4 most recent windows are kept. Steps (1) and (3) are load-bearing (Appendix E).

The residualization must be a time-series regression. Of the two natural ways to remove the market, only one is correct: the other **manufactures exactly the structure the pipeline is meant to detect**. Projecting each date's cross-section onto the orthogonal complement of β, with $P = I - \beta\beta^\top/(\beta^\top\beta)$ so that $P\beta = 0$ and $P = P^\top = P^2$, gives

$$\mathrm{Cov}(Pr_t) \;=\; P\Sigma P^\top \;=\; P\big(\sigma_m^2 \beta\beta^\top + \Sigma_u\big)P \;=\; P\,\Sigma_u\,P, \tag{2}$$

(2) kills the factor term but is *dense even when* Σ_u *is diagonal*, mixing every asset into every other through the shared β; these correlations are real in the projected data and spurious in the original. On a synthetic panel where $\mathrm{supp}_{\mathrm{off}}(\Sigma_u)$ is known by construction, that route gives a false-discovery proportion of **0.79** against a target of 0.10, versus 0.0877 for the time-series regression (T8b).

The correction is chosen by measuring dependence. Pairs are tested with Fisher's transform (Appendix E.1) and the $m = 6,670$ p-values corrected for multiplicity. *Which* correction is licensed depends on the dependence among the statistics, so it was measured: over 300 block resamples, **15.7%** of statistic pairs correlate below -0.20 (minimum **-0.717**). Material *negative* association is incompatible with positive regression dependency, so Benjamini–Hochberg's guarantee is unavailable here (A3), and Benjamini–Yekutieli is used at $q_{\mathrm{BY}} = q/\sum_{i=1}^{m} i^{-1} = 0.01066$, valid under *arbitrary* dependence (Benjamini & Yekutieli, 2001). That guarantee concerns *dependence* and presupposes valid p-values; heavy tails threaten the latter, which is Gate F's job. That correction is usually declined as too costly; here it keeps **59.0%** of Benjamini–Hochberg's edges and a non-empty graph in all 52 windows, so the cost is accepted, with BH kept at $q = 0.10$ as a sensitivity arm quoted at its *realized* FDR of ≈ 0.12.

4.2 The factored action-value function

Given $\widehat{E}$, the action-value function is parameterised as

$$Q(s,a) \;=\; \sum_{i=1}^{N} \varphi_i(s_i;a_i) \;+\; \sum_{(i,j)\in\widehat{E}} \psi_{ij}(s_i,s_j;a_i,a_j) \;+\; g_{\mathrm{mkt}}\big(\beta^\top w(a)\big), \tag{3}$$

with one network φ shared across nodes and one network ψ shared across edges, following Böhmer et al. (2020). (3) is the learned counterpart of the reward decomposition (4)–(5), term for term: that is why the rank-one market term is supplied *analytically* rather than learned, and why $\widehat{E}$ estimates a set the reward already determines rather than an architectural choice.

Joint-action selection. (3) must be maximised over $3^{116} \approx 10^{55}$ joint actions, once per decision and again inside every Bellman update, and both applicable published methods fail here (§6.1). The selector is multi-restart coordinate ascent, scoring candidates under the *full* objective (3). It carries **no optimality guarantee**, though on the one validated graph thin enough to solve exactly it returns the exact optimum in 100% of states (Appendix C).

5 Theory

Four results follow, in the order a reader needs them: the decomposition, its error under estimated edges, its exact scope, and why the coupling it produces is inert.

5.1 The decomposition, and what the graph is

Proposition 1 (Exact factored decomposition)**.** *Under A1–A2, the reward* (1) *satisfies*

$$r(w) \;=\; \sum_{i=1}^{N} \varphi_i(w_i) \;+\; \sum_{(i,j)\in E^*} \psi_{ij}(w_i,w_j) \;+\; g_{\mathrm{mkt}}(\beta^\top w), \tag{4}$$

with $E^* = \mathrm{supp}_{\mathrm{off}}(\Sigma_u)$ *and*

$$\varphi_i = w_i\mu_i - \tfrac{\gamma}{2}\Sigma_{u,ii}w_i^2 - c\,|w_i - w_{\mathrm{prev},i}|, \quad \psi_{ij} = -\gamma\,\Sigma_{u,ij}\,w_iw_j, \quad g_{\mathrm{mkt}}(z) = -\tfrac{\gamma}{2}\sigma_m^2 z^2. \tag{5}$$

Proof sketch (Appendix B). Expanding $w^\top\Sigma w$ under the factor form, the factor part collapses to $(\beta^\top w)^2$, one scalar however many assets there are, while the residual part splits into diagonal terms, which join φ_i, and off-diagonal terms that are exactly ψ_{ij}. Restricting to E^* is definitional, not an approximation: **the graph is read off the reward, not chosen**. Stating (4) on $\mathrm{supp}_{\mathrm{off}}(\Sigma)$ instead carries *no information*, Σ_{ij} being nonzero for essentially every pair (raw density 0.9836); and g_{mkt} reaches the action only through the scalar $\beta^\top w$, which makes it **rank one rather than an N-way interaction**.

5.2 Error under imperfect edge discovery

In practice only $\widehat{E}$ is available, not E^*. Proposition 4 (Appendix B) connects them, and is why edge validation is not cosmetic: if $\widehat{E}$ comes from a procedure controlling the false discovery rate at level q, then the expected *fraction* of pairwise coupling terms in $\widehat{Q}$ that are spurious is at most q, and $|\widehat{Q} - Q|$ is bounded by γ times three sums: a *variance* term over spurious edges $\widehat{E} \setminus E^*$, a *bias* term over omitted mass $E^* \setminus \widehat{E}$, and an *estimation* term over edges correctly found but with magnitudes estimated. Two wordings are load-bearing: FDR bounds the expectation of a *ratio*, so it governs the expected fraction of spurious terms and never their *number*; and the third group cannot be dropped, since a correctly found edge still has an estimated magnitude.

5.3 Exact scope: which objectives admit this at all

Proposition 2 (Decomposability criterion)**.** *A three times continuously differentiable $f : \mathbb{R}^N \to \mathbb{R}$ admits a representation $f(w) = \sum_i \varphi_i(w_i) + \sum_{i<j} \psi_{ij}(w_i, w_j)$* ***if and only if***

$$\frac{\partial^3 f}{\partial w_i \, \partial w_j \, \partial w_k} \;=\; 0 \qquad \textit{for all distinct } i, j, k. \tag{6}$$

Differentiating over three distinct indices annihilates every one- and two-variable term; the converse integrates $\partial^2 f / \partial w_i \partial w_j$ (Appendix B). Measured on the real covariance this is not a marginal call: mean–variance with costs returns only roundoff, falling as h^{-3} with the differencing step, while risk-parity dispersion (Maillard et al., 2010) returns a step-*independent* 65 and Differential Sharpe (Moody et al., 1998) depends on w only through $w^\top r_t$, so deleting *every* edge changes it by exactly zero. The construction is as broad as the mean–variance quadratic and no broader, which is why the negative results here **cannot be read as refuting published systems**.

5.4 Why the coupling cannot move a discrete decision

The construction is well-posed; this subsection, which carries the paper's central claim, shows that it is inert. Differentiating (3) in w_i, with other positions fixed, gives the marginal value the agent sees (the coupling sum running over the discovered $\widehat{E}$, not E^*):

$$\frac{\partial \widehat{Q}}{\partial w_i} \;=\; \underbrace{\mu_i}_{O(1)} \;-\; \underbrace{c\,\operatorname{sgn}(w_i - w_{\mathrm{prev},i})}_{O(1)} \;-\; \underbrace{\gamma \Sigma_{u,ii} w_i}_{O(w)} \;-\; \underbrace{\gamma \sum_{j:(i,j)\in\widehat{E}} \Sigma_{u,ij} w_j}_{O(w),\ \textbf{the graph}} \;-\; \underbrace{\gamma \sigma_m^2 \beta_i (\beta^\top w)}_{O(w)}. \tag{7}$$

Two liberties, neither load-bearing: $\operatorname{sgn}(0)$ is any element of $[-1, 1]$, so (7) is an inclusion at the kink (the no-trade band D2 tests), and it is the *myopic* gradient, whose continuation term adds a further $O(1)$ cost subgradient that widens the gap the graph must close. Ordered by *position size*: alpha and cost are $O(1)$, since they do not shrink as the book is spread thinner, while every risk term, the graph's included, is $O(w)$ *per term*. The *aggregate* coupling does not vanish with N: degree grows as the weights shrink and Proposition 3 shows the two cancel exactly. What makes the graph inert is therefore its absolute size at realistic leverage, not asymptotic decay. Since the discrete action is essentially the *sign* of (7), at $w_j \sim g/N$ the graph perturbs a quantity whose sign alpha and cost have already fixed. **That is the whole mechanism**, and it is about the objective, not the graph.

Proposition 3 (Scale invariance of the coupling ratio)**.** *If residual off-diagonals are bounded, degrees in $\widehat{E}$ grow as $O(\delta N)$ at density δ, positions obey $|w_j| = O(g/N)$ and $|\mu_i|$ is bounded below, then the coupling term of* (7) *is $O(\delta g)$ and its ratio to own-alpha is* ***constant in*** *N (constants in Appendix B).*

The proof, with constants, is in Appendix B: degree growth and weight shrinkage cancel. The claim is falsifiable and §6.2 tests it: *a more concentrated portfolio cannot rescue the mechanism*. What *would* rescue it is leverage g, or an objective promoting the coupling to $O(w)$; §6.4 builds both.

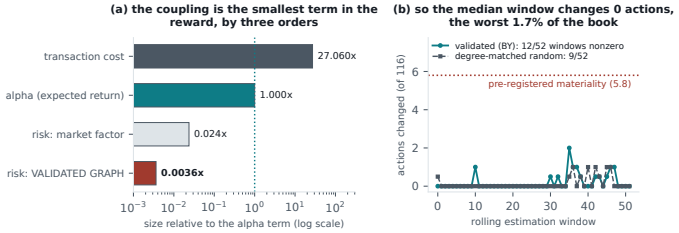


Figure 1: **the paper in one figure.** *(a)* mean absolute size of each reward term (1) on random books at gross $1\times$, as a multiple of alpha: the coupling is $0.0036\times$ alpha and $7{,}463\times$ smaller than cost, so it cannot flip the sign alpha and cost fix. *(b)* the consequence over all 52 windows: the pairwise terms change a median of 0 of 116 actions, and the count-matched random control behaves the same way.

6 Experiments

Setup. 116 US large-cap equities, daily log returns, 2010-01-05 to 2024-12-30, 3,772 trading days; 52 rolling windows of 504 days stepped by 63. The panel is present-day large caps, so **no absolute figure here is an achievable strategy result**. Six arms differ *only* in which pairs carry a ψ_{ij} term; the important one is **G1**, random edges count-matched to the validated graph **G4**, the control that rules out "more edges helped". Metrics appear beside chance baselines, with intervals *on the difference* (Appendices E.4–M). Every study was pre-registered, with its predictions, decision rules and the outcome that would count against it fixed before the run, and Table 6 catalogues all 40. Five of them were **voided** by their own validity checks, covering eleven runs in total, and two verdicts were **withdrawn** after the fact (Appendix L). Every result below ran under both BY and BH, and **no conclusion differs** (Appendix J).

6.1 The graph, and what it costs

Raw correlation is all but complete: density **0.9668** at the primary BY level and 0.9836 under BH, against a permutation-null chance of 2.13×10^{-4}. Residualizing and correcting gives **0.2552** and 0.4062, paired differences *within each arm* of 0.7115 and 0.5774, both intervals excluding zero (Table 5).

Neither published exact method survives contact with this graph, which is why the selector of §4.2 is heuristic: a minor-min-width *lower* bound on the induced width is median 44 over all 52 windows, so exact elimination needs $\geq 3^{45}$ table entries per $\arg\max$, and a bound from *below* means no ordering escapes it. Max-plus (Böhmer et al., 2020) matched the optimum on 56.7% of real 12-asset subgraphs against 80.0% for coordinate ascent. A validated *materiality floor* restores exactness (Appendix L).

6.2 The main result: how many actions does the graph change?

The measurement. For each state the joint $\arg\max$ of (3) is taken with and without the pairwise terms, and the differing coordinates are counted. The comparison is *within-state and paired*, so no training seed enters. Sweeping every axis on which §5.4 makes a prediction (universe size, gross exposure, risk aversion, cost and action granularity; Appendix M) over all 52 windows, the coupling changes **no actions** at every setting; across the windows the median changes **0** of 116 and the worst 1.7%, below the 5% materiality threshold fixed in advance. The reason is arithmetic: the coupling is $0.0036\times$ alpha, where cost is $27.06\times$.

Four rival explanations, closed (Appendix L). *The action grid?* No (D2): in the continuous relaxation the graph reallocates only 1.32% of the book. *A cost-induced no-trade band?* No (D2): with costs set to zero the no-graph policy moves 76 of 116 names off their previous positions, so the book is fully active, and the graph still changes nothing. *States the agent never visits?* No (D3, E1): on the occupancy of eight independently trained agents it changes a mean of **0.196** actions $[0.151, 0.244]$. *Edges too weak?* No (D4): the p-value ordering *is* the magnitude ordering, and the 253 strongest

residual correlations alone change at most 1 of 116. Proposition 3 therefore holds on real data: invariant in N, visible only at $20\times$ leverage. And because the arms differ only in a quantity that does not change behaviour, the natural performance hypothesis is **untestable rather than null**, because there is no effect to power for.

6.3 WHY: THE GRAPH'S INFORMATION IS SAME-DAY ONLY

§5.4 argues the coupling cannot move the argmax because it enters one order *in position size* below alpha and cost (Proposition 3). That is algebra; what the edges *contain* is measurable. Fitting on $t < 1,600$ and scoring on the 2,172 days that follow (one long split rather than the rolling windows used elsewhere, because this measures a property of the data and not of a decision rule), each asset's market residual is regressed on the mean residual of its validated neighbours. Same day the graph is unmistakably real: out-of-sample R^2 **0.0907** against **0.0030** for a degree-matched random graph. Lagged one day it is **-0.0010**, and so is every alternative: sector labels and hierarchical clusters at $K = 5$–40 all land between -0.0010 and -0.0005 at lag 1 despite reaching 0.2001 same-day, and residual-volatility terciles give six ordered lead–lag pairs, **every one negative**.

Contemporaneous co-movement is a risk fact, not a signal: by the time a neighbour's move is observable the move one would trade has happened. Risk enters the marginal at $O(w)$ *per term*, alpha at $O(1)$, so *what can be validated at daily frequency is exactly the kind of structure that cannot move an argmax*: the inertness of §6.2 belongs to daily equity data, not to the agent. The conditional-mean channel confirms this. A lagged relation enters μ_i, the $O(1)$ channel, and the identical pipeline pointed at the mean gives an effect ratio of **0.67** from $1\times$ to $20\times$ gross, so it does *not* scale with leverage where the coupling's does. But after persistence that graph holds **2 edges of 13,340** ordered pairs. **Both channels are inert**, one by its order and the other because its graph is empty, which identifies not a different objective but a different frequency: Curme et al. (2015) validate one lead–lag link at 130 minutes and hundreds below 30, and a lead–lag graph is a conditional-mean object this paper's result does not cover (Remark 1).

And architecture does not rescue it. Putting the graph in the *representation* instead, a trained GAT changes 11–39% of actions where the coupling changes none, but a count-matched *random* graph changes just as many, the validated arm scoring **-0.0181** $[-0.0247, -0.0109]$ against it. Pooling is variance reduction, and variance reduction does not care which neighbours. Finally, 41.4% of validated edges are same-sector against 12.1% at chance, and a leave-one-out sector factor removes **78%** of them. Most of the graph is therefore sector structure that a single market factor leaves behind, which is reported here rather than adopted because sector labels are not point-in-time (Appendix H).

6.4 THE TWO REGIMES WHERE THE GRAPH SHOULD MATTER

If §5.4 is right, promoting the coupling to the dominant order should make it matter, so both were built, each with two controls (Appendix L). Under *quadratic cross-impact costs*, which make cost $O(w^2)$, an effect first appeared and proved to be an artifact: both impact matrices were *indefinite*, a cost with negative-cost directions; repaired on the same supports (E22), **every arm changes 0 actions in every window**. Under a *binding volatility budget* the number of altered actions reaches **16.5** of 116 (E3), but that is the *level*; the *contrast* at the true discrete optimum is **+2.8 [+0.9, +4.7]** names (E24b) at the *tightest* of five budgets, the one most favourable to the graph (-0.2 when slack), yet still below the 5.8-name materiality line and inside the selector's restart noise. **In neither regime do validated edges beat a count-matched random control materially.**

Is the inertness a property of the graph, or of the objective? With the graph held fixed and only the objective changed to risk-parity dispersion (Maillard et al., 2010), the identical edges change **18** of 116 actions against 4 random, or $4.5\times$ its control. **The inertness belongs to the objective, not the graph**, as Proposition 2 predicts. Trained there over 8 seeds the mechanism *survives and amplifies*, with a median **65.6%** of names differing between arms, yet out of sample they are indistinguishable, because the cross-seed standard deviation of the paired *relative* performance difference is 4.957, so resolving a 10% relative effect needs **19,261** seeds $[12,380, 34,044]$ for 80% power at a 10% effect against a cap of 100. **Decision-relevance does not imply measurable performance**: the hypothesis is *underpowered in advance*, a noise floor no paper in this literature reports. §6.7 follows the objective thread to one with no return term at all.

6.5 ARE LEARNED RELATIONAL EDGES RELIABLE ACROSS SEEDS?

If a model *learns* which assets are related, independent runs should learn the same relations. They do not. Over 20 seeds, chance-corrected agreement of the top 146 learned edges looks weakly reliable until a *parameter-randomization control* (Adebayo et al., 2018) is added, at which point **untrained** networks agree *more* (0.1895) than trained (0.1106); the same holds on a multi-head GAT once raw attention is degree-corrected (-0.0394 $[-0.0543, -0.0239]$). **A chance baseline shows that a score is above random; it does not show that training produced it.** The deficit is one of *variance*, not absence: averaging sixteen seeds buys a half-stable edge set (E30, Appendix F).

6.6 THE DIAGNOSTIC ON A PUBLISHED SYSTEM

Every result above concerns the construction studied here; this one does not. Of six published graph systems in scope, one releases runnable code: HATS (Kim et al., 2019). It was trained on all *four* test phases with its own code and hyperparameters behind a pre-registered gate: a run counts only if its test macro-F1 lands within 0.02 of their Table 2. **Two of four clear it** (0.3098 vs. a published 0.3100; 0.3289 vs. 0.3267); the others are void and nothing is drawn from them. On both, deleting the graph leaves the book identical on 95% and 95% of test days, and the graph does **not** beat a degree-matched random one in any of **eight** trained models (Appendix G); deleting the relational pathway moves average daily return by [-0.0055, +0.0000] pp, inside the 0.0350 pp margin their Table 3 claims over its strongest baseline.

And no graph could have mattered there. The relational module enters through one residual into one dense layer, so no graph can change a class margin by more than $\|w_c - w_{c'}\| \, \|R_i\|$, for readout weights w_c of class c and company i's relational-pathway output R_i. On the trained weights that puts 98.2–99.7% of decisions beyond any graph's reach across the eight models, with zero flips inside that set at every pathway scale up to $2048\times$ (E28). That is a property of the *weights*, so the objection that the system was given the wrong graph has nothing to attach to (Appendix G).

6.7 WHERE THE GRAPH DOES PAY

Every negative above has the same shape: the coupling is crowded out by terms entering at $O(1)$. The remedy is **an objective with no return term**, which consumes exactly the same-day covariance information §6.3 shows the graph carries. Removing that term reduces the problem to *minimum-variance allocation*, so what is tested here is the covariance model and not the agent; the baseline is the textbook sample-covariance minimum-variance portfolio, which the validated support beats by **-18.66%**. For each support, the fitted covariance is the maximum-likelihood one whose *precision* matrix vanishes off it: positive definite by construction, no repair, and a likelihood, so the arms are nested models rather than a volatility race. Figures quoted cover the **90 walk-forward windows after** $t = 1,862$, which no other experiment here uses; construction and every arm are in Appendix I.

Against a density-matched random support at the same per-node degree the validated support wins on both endpoints, **+12.86** $[+10.75, +14.59]$ log-likelihood per held-out day and **-10.42%** $[-17.04, -5.48]$ in realised volatility. The covariance-estimation window W, in days, was a pre-registered axis and the contrast holds at both points (-12.75% at $W = 750$, intervals excluding zero).

An attempt was then made to take it away. Since most of the graph is sector structure, the win might be the sectors. The design was re-run on market + sector residuals under a rule fixed in writing beforehand: retain at least half the effect on both endpoints. It retains **27.0%** and **47.9%**: below threshold, so by the paper's own rule the first measurement was substantially sector structure. What survives is smaller and real, both intervals still excluding zero: **+3.47** $[+1.49, +4.73]$ and **-4.99%** $[-9.72, -1.21]$. **That deflated figure is the one claimed**. Validated-structure clustering never beat raw-correlation clustering to begin with, so this is a result about *sparse precision support*, not clustering (Appendix I).

7 DISCUSSION

Three numbers, and what each rules out. *(i) How many of an agent's actions does the graph change?* This number is cheap, mechanism-independent and absent from every graph-RL trading paper examined, but §6.4 shows it is **necessary and not sufficient**: under the risk-parity objective, where no return term crowds the coupling out, the graph changed 66% of the policy and still bought no detectable performance. In a continuous action space a strict count is noise-dominated, and the usable forms are the fraction of gross exposure reallocated or a pre-declared materiality threshold, both reported here. *(ii) What fraction of decisions could **no** graph change?* Where the relational pathway enters additively before a linear readout, the bound of §6.6 answers it from the trained weights alone, putting 98.2–99.7% of a published system's decisions out of reach, which a performance ablation would credit to the graph. The bound tracks the realised effect at Spearman 0.98 (E28). **The generalisation does not carry**: under a nonlinear readout the Lipschitz form is sound but vacuous on 6 of the GAT models trained here (E29), so the number rules a graph *out* and never in. *(iii) What seed count can the comparison resolve?* Here 19,261 against a cap of 100, and more backtest is no substitute: 74% of the variance is between-seed.

Scope and limitations. The result concerns one construction, the natural formalisation of "put the graph in the value function", and **no verified incumbent does this**: the published reward family is Sharpe-type (Sood et al., 2023; Moody et al., 1998) or drawdown-based (Wang et al., 2021), with no covariance at all. This is *not* a demonstration that graphs cannot help RL traders. What carries beyond it is the ablation question, the noise floor, and the timing measurement of §6.3, a statement about daily equity data rather than about any agent. Caveats: the panel is present-day large caps and daily; the selector is a heuristic whose restart noise reaches 62 names at the tightest risk budget; the cross-impact regime was tested on four supports, not all; the graph-attention result uses one GAT configuration; and §6.7 is an allocation result, not a rescue of the agent.

8 CONCLUSION

This paper asked what a graph does to a reinforcement-learning trader's *decisions* rather than its reported returns. Placed where the reward says they belong, as the pairwise coupling of a factored action-value function, validated edges change a median of 0 of 116 actions, and forty pre-registered studies leave that intact. The reason is not an empty graph: **its information is contemporaneous at daily resolution** (-0.0010 at one day's lag), and same-day co-movement is a risk fact, entering the decision marginal one order *in position size* below alpha and cost. *What can be validated at daily frequency is exactly the kind of structure that cannot move an argmax.* That diagnosis names its remedy, which the paper tests: in an objective with no return term the same edges, used as sparse precision support, beat a density-matched random support out of sample. The paper's own adversarial pass then deflates that effect to 48% of the first measurement, and the deflated figure is the one claimed. The practical recommendation is to report *how many actions the graph changes*, *what fraction no graph could change*, and *the seed count the comparison can resolve*; and, where the graph must earn its place, to use it in an objective in which covariance does not compete with a first-order return term.

AI USE STATEMENT

The authors used generative AI tools for coding assistance and for drafting and editing manuscript text. Research ideas, hypotheses and decision rules were fixed by the authors in pre-registrations written before each run; all results come from code the authors reviewed; and AI-assisted prose was reviewed and edited by the authors. AI was not a substitute for reading the literature (Appendix O). The authors take full responsibility for this work.

ETHICS STATEMENT

This work reports no deployable system, introduces no new capability for market manipulation and releases no model; the data are public daily prices of listed equities, with no human subjects. One point concerns other researchers' work: the paper argues that a class of published performance claims cannot separate signal from seed noise, and *quantifies* that floor rather than asserting it. **That is a claim about what is reported, not an allegation of error or misconduct**: those systems were not designed to answer this question, every quotation is verbatim, and for the one with released code only its behaviour at the commit run is reported.

REPRODUCIBILITY STATEMENT

Proofs and 16 numerical checks on synthetic ground truth are in Appendix B, with every object's data origin in Appendix D; data, arms, hyperparameters, execution order and pre-registrations are in §6 and Appendices M and L. Every number is read from a saved artifact at build time rather than transcribed, and the text names the artifact behind each one, so a reader can check any figure against the file that produced it. The code and artifacts are not included with this submission; they will be released upon publication.

REFERENCES

Julius Adebayo, Justin Gilmer, Michael Muelly, Ian Goodfellow, Moritz Hardt, and Been Kim. Sanity checks for saliency maps. In *Advances in Neural Information Processing Systems*, 2018.

Bharathan Balaji, Petros Christodoulou, Xiaoyu Lu, Byungsoo Jeon, and Jordan Bell-Masterson. FactoredRL: Leveraging factored graphs for deep reinforcement learning. In *NeurIPS Deep Reinforcement Learning Workshop*, 2020.

Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*, 57(1):289–300, 1995.

Yoav Benjamini and Daniel Yekutieli. The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, 29(4):1165–1188, 2001.

Stefanos Bennett, Mihai Cucuringu, and Gesine Reinert. Lead-lag detection and network clustering for multivariate time series with an application to the US equity market. *arXiv preprint arXiv:2201.08283*, 2022.

Peter J. Bickel and Elizaveta Levina. Covariance regularization by thresholding. *The Annals of Statistics*, 36(6):2577–2604, 2008.

Wendelin Böhmer, Vitaly Kurin, and Shimon Whiteson. Deep coordination graphs. In *Proceedings of the 37th International Conference on Machine Learning*, 2020.

Gary Chamberlain and Michael Rothschild. Arbitrage, factor structure, and mean-variance analysis on large asset markets. *Econometrica*, 51(5):1281–1304, 1983.

Yves Choueifaty and Yves Coignard. Toward maximum diversification. *Journal of Portfolio Management*, 35(1):40–51, 2008.

Chester Curme, Michele Tumminello, Rosario N. Mantegna, H. Eugene Stanley, and Dror Y. Kenett. Emergence of statistically validated financial intraday lead-lag relationships. *Quantitative Finance*, 15(8):1375–1386, 2015.

A. P. Dempster. Covariance selection. *Biometrics*, 28(1):157–175, 1972. doi: 10.2307/2528966.

Jianqing Fan, Yuan Liao, and Martina Mincheva. Large covariance estimation by thresholding principal orthogonal complements. *Journal of the Royal Statistical Society, Series B*, 75(4):603–680, 2013.

Carlos Guestrin, Daphne Koller, and Ronald Parr. Multiagent planning with factored MDPs. In *Advances in Neural Information Processing Systems*, 2001.

Kung-Cheng Ho, Shih-Cheng Lee, and Jiun-Lin Chen. Book-to-market equity and asset correlations — an international study. *International Review of Economics & Finance*, 2022. doi: 10.1016/j.iref.2022.02.014.

Raehyun Kim, Chan Ho So, Minbyul Jeong, Sanghoon Lee, Jinkyu Kim, and Jaewoo Kang. HATS: A hierarchical graph attention network for stock movement prediction. *arXiv preprint arXiv:1908.07999*, 2019.

Olivier Ledoit and Michael Wolf. Honey, i shrunk the sample covariance matrix. *The Journal of Portfolio Management*, 30(4):110–119, 2004.

Sébastien Maillard, Thierry Roncalli, and Jérôme Teïletche. On the properties of equally-weighted risk contribution portfolios. *Journal of Portfolio Management*, 36(4):60–70, 2010.

Harry Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952.

John Moody, Lizhong Wu, Yuansong Liao, and Matthew Saffell. Performance functions and reinforcement learning for trading systems and portfolios. *Journal of Forecasting*, 17(5–6):441–470, 1998.

Guannan Qu, Adam Wierman, and Na Li. Scalable reinforcement learning for multi-agent networked systems. *arXiv preprint arXiv:1912.02906*, 2020.

Farzan Soleymani and Eric Paquet. Deep graph convolutional reinforcement learning for financial portfolio management – DeepPocket. *Expert Systems with Applications*, 182:115127, 2021.

Srijan Sood, Kassiani Papasotiriou, Marius Vaiciulis, and Tucker Balch. Deep reinforcement learning for optimal portfolio allocation: A comparative study with mean-variance optimization. Technical report, J.P. Morgan AI Research, 2023.

Peter Spirtes, Clark Glymour, and Richard Scheines. *Causation, Prediction, and Search*. MIT Press, 2nd edition, 2001.

Yoshimasa Uematsu and Takashi Yamagata. Discovering the network Granger causality in large vector autoregressive models. *arXiv preprint arXiv:2303.15158*, 2024.

Harm van Seijen, Mehdi Fatemi, Joshua Romoff, Romain Laroche, Tavian Barnes, and Jeffrey Tsang. Hybrid reward architecture for reinforcement learning. In *Advances in Neural Information Processing Systems*, 2017.

Zhicheng Wang, Biwei Huang, Shikui Tu, Kun Zhang, and Lei Xu. DeepTrader: A deep reinforcement learning approach for risk-return balanced portfolio management with market conditions embedding. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021.

Bing Zhang. Graph attention-based heterogeneous multi-agent deep reinforcement learning for adaptive portfolio optimization. *Scientific Reports*, 16:2674, 2026.

APPENDIX

The appendix is organised so a reader can reconstruct every claim from first principles. A tabulates the systems surveyed in §2; B completes and numerically verifies the theory; C states the action selector; D records where every mathematical object came from and what was changed; E derives the statistical machinery; K records, for each of the ten decisions where more than one method was tried, what each gave and why one was adopted; L catalogues the studies; M lists constants and the execution order; N collects the supporting figures; and O records which cited works were actually read.

A WHERE THE GRAPH SITS IN EACH SYSTEM EXAMINED

B COMPLETING AND VERIFYING THE THEORY

B.1 FULL PROOF OF PROPOSITION 1

The factor form gives $w^\top \Sigma w = \sigma_m^2\, w^\top \beta\beta^\top w + w^\top \Sigma_u w$. *(i)* Since $\beta^\top w \in \mathbb{R}$, $w^\top \beta\beta^\top w = (\beta^\top w)^2$: the factor term collapses to $g_{\mathrm{mkt}}(\beta^\top w)$, a function of *one* scalar however many assets there are. *(ii)* Using $\Sigma_{u,ij} = \Sigma_{u,ji}$ to collect the two copies of each off-diagonal pair, $w^\top \Sigma_u w = \sum_i \Sigma_{u,ii} w_i^2 + 2\sum_{i<j} \Sigma_{u,ij} w_i w_j$; times $-\gamma/2$ the first sum is the quadratic part of φ_i and the second is $\sum_{i<j} \psi_{ij}$. *(iii)* ψ_{ij} vanishes identically whenever $\Sigma_{u,ij} = 0$, so the sum over $i < j$ equals the sum over E^* by definition of E^*; this is definitional, not an approximation. *(iv)* $w^\top \mu = \sum_i w_i \mu_i$ and, under A2, the cost is $\sum_i c|w_i - w_{\mathrm{prev},i}|$; collecting single-index terms gives φ_i. □

B.2 FULL PROOF OF PROPOSITION 4

The statement, restated here in full:

Proposition 4 (False-discovery control bounds the value error)**.** *Let $\widehat{E}$ come from a procedure controlling the false discovery rate at level q; let $\widehat{Q}$ be* (3) *on $\widehat{E}$ with estimated magnitudes $\widehat{\Sigma}_u$ and*

Table 1: placement of the graph in each system. The *value function* column matters because the construction studied here places the graph there: no verified incumbent does, and two of the six have no action-conditioned value function at all.

System	Graph placement	Value function	Edges tested?
DeepTrader (Wang et al., 2021)	spatial attention + GCN in the asset-scoring unit	none (policy gradient, no critic)	no
Zhang (2026)	GAT feature encoder, three functional agents	softmax consensus over simplex weights	no
DeepPocket (Soleymani & Paquet, 2021)	correlation edges into a GCN in actor and critic	monolithic scalar $V(s)$; critic takes no action argument	no
Sood et al. (2023)	none (covariance only in the MVO baseline)	Differential Sharpe (Moody et al., 1998)	n/a
HATS (Kim et al., 2019)	hierarchical graph attention	not an RL method	**yes, in this paper** (§6.6)
FactoredRL (Balaji et al., 2020)	hand-specified factored graph as I/O masks	block-diagonal $Q = \sum_k Q_k$, no pairwise terms	no

Q the same expression on E^* with the true Σ_u, the two sharing φ_i and g_{mkt}. Then the expected fraction *of pairwise coupling terms in $\widehat{Q}$ that are spurious is at most q, and*

$$\big|\widehat{Q}-Q\big| \;\le\; \gamma\Big[\underbrace{\sum_{(i,j)\in\widehat{E}\setminus E^*}\big|\widehat{\Sigma}_{u,ij}w_iw_j\big|}_{\text{(a) variance: spurious edges}} + \underbrace{\sum_{(i,j)\in E^*\setminus\widehat{E}}\big|\Sigma_{u,ij}w_iw_j\big|}_{\text{(b) bias: omitted mass}} + \underbrace{\sum_{(i,j)\in\widehat{E}\cap E^*}\big|(\widehat{\Sigma}_u-\Sigma_u)_{ij}w_iw_j\big|}_{\text{(c) error on found edges}}\Big]. \tag{8}$$

The fraction. Every edge of $\widehat{E}$ contributes exactly one pairwise term to $\widehat{Q}$, and that term is spurious precisely when its edge lies in $\widehat{E}\setminus E^*$. A procedure controlling the false discovery rate at level q satisfies $\mathbb{E}\big[|\widehat{E}\setminus E^*|/\max(|\widehat{E}|,1)\big]\le q$, which is that fraction. Note what this does *not* say: FDR bounds the expectation of a *ratio*, so nothing follows about the *number* $|\widehat{E}\setminus E^*|$ of spurious terms, which grows with $|\widehat{E}|$.

The bound. Write the two pairwise parts, using ψ_{ij} from (5),

$$\widehat{\Psi}(w)=\sum_{(i,j)\in\widehat{E}}-\gamma\widehat{\Sigma}_{u,ij}w_iw_j,\qquad \Psi^*(w)=\sum_{(i,j)\in E^*}-\gamma\Sigma_{u,ij}w_iw_j.$$

Since φ_i and g_{mkt} are common to $\widehat{Q}$ and Q by hypothesis, $\widehat{Q}-Q=\widehat{\Psi}-\Psi^*$. Split both index sets on $\widehat{E}\cap E^*$: $\widehat{E}=(\widehat{E}\setminus E^*)\,\dot\cup\,(\widehat{E}\cap E^*)$ and $E^*=(E^*\setminus\widehat{E})\,\dot\cup\,(\widehat{E}\cap E^*)$. The shared block leaves a difference of magnitudes, so

$$\widehat{\Psi}-\Psi^*=-\gamma\Big[\sum_{(i,j)\in\widehat{E}\setminus E^*}\widehat{\Sigma}_{u,ij}w_iw_j\;-\sum_{(i,j)\in E^*\setminus\widehat{E}}\Sigma_{u,ij}w_iw_j\;+\sum_{(i,j)\in\widehat{E}\cap E^*}(\widehat{\Sigma}_u-\Sigma_u)_{ij}w_iw_j\Big],$$

and taking absolute values term by term gives (8). □

The three groups are the paper's (a) variance, (b) bias and (c) estimation error. Group (c) is not decorative: T9 asserts the three-term bound holds on every replication, and **T9c asserts that the two-term bound, with (c) deleted, is violated on at least one**, so restoring the shorter form fails the verifier rather than passing silently.

B.3 Full proof of Proposition 2

($\Rightarrow$) Suppose $f(w) = \sum_i \varphi_i(w_i) + \sum_{i<j} \psi_{ij}(w_i, w_j)$ and let i, j, k be distinct. Differentiating term by term: a node term φ_ℓ depends on one variable, so at least two of the three derivatives are taken of a constant and it vanishes; a pairwise term $\psi_{\ell m}$ depends on two, so at least one of the three indices is absent from it and it vanishes. Hence $\partial^3 f / \partial w_i \partial w_j \partial w_k = 0$.

($\Leftarrow$) Suppose $\partial^3 f / \partial w_i \partial w_j \partial w_k = 0$ for all distinct i, j, k. Fix $i \neq j$. For any $k \notin \{i, j\}$, $\partial / \partial w_k \big(\partial^2 f / \partial w_i \partial w_j\big) = 0$, so the mixed second partial is constant in every coordinate other than w_i, w_j; write it $h_{ij}(w_i, w_j)$. Define $\psi_{ij}(w_i, w_j) = \int_0^{w_i} \int_0^{w_j} h_{ij}(s, t)\, dt\, ds$, so that $\partial^2 \psi_{ij} / \partial w_i \partial w_j = h_{ij}$ and ψ_{ij} has no other nonzero mixed second partial over distinct indices. Set $g = f - \sum_{i<j} \psi_{ij}$. Then for every $i \neq j$, $\partial^2 g / \partial w_i \partial w_j = h_{ij} - h_{ij} = 0$, so g is separable: $g(w) = \sum_i \varphi_i(w_i)$ with $\varphi_i(w_i) = \int_0^{w_i} \partial g / \partial w_i$ up to a constant, which may be absorbed into any one term. This gives the representation. □

B.4 Full proof of Proposition 3

Orders here are **degrees in position scale**, not orders in N. Writing $w = g\hat{w}$ with $\|\hat{w}\|_1$ fixed, a term is $O(1)$ when it is homogeneous of degree 0 in g and $O(w)$ when degree 1. The two readings are independent, and conflating them is what makes the coupling appear to decay.

Fix asset i. Let $\bar{\sigma}^2 = \max_{i \neq j} |\Sigma_{u,ij}| < \infty$ (bounded residual off-diagonals), $\bar{\beta} = \max_i |\beta_i| < \infty$, and $\mu_{\min} = \min_i |\mu_i| > 0$ (own alpha bounded below); let $\deg_{\widehat{E}}(i) = O(\delta N)$ at density δ and $|w_j| = O(g/N)$.

The coupling term.

$$\Big|\gamma \sum_{j:(i,j)\in\widehat{E}} \Sigma_{u,ij} w_j\Big| \;\le\; \gamma\, \bar{\sigma}^2\, \deg_{\widehat{E}}(i)\, \max_j |w_j| \;=\; \gamma\, \bar{\sigma}^2 \cdot O(\delta N) \cdot O(g/N) \;=\; O\big(\gamma \bar{\sigma}^2 \delta g\big), \quad (9)$$

in which N cancels: the degree grows at exactly the rate the weights shrink. Dividing by $|\mu_i| \ge \mu_{\min}$ gives a ratio $O(\gamma \bar{\sigma}^2 \delta g / \mu_{\min})$, **constant in** N. □

The market term is N-invariant as well, by the same algebra, and this strengthens the argument rather than weakening it: $|\beta^\top w| \le \bar{\beta}\, N \cdot O(g/N) = O(\bar{\beta} g)$, so $\gamma \sigma_m^2 \beta_i (\beta^\top w) = O(\gamma \sigma_m^2 \bar{\beta}^2 g)$, again independent of N. **Two of the three risk terms of** (7) **are therefore N-invariant aggregates**, so inertness cannot arise from asymptotic decay in either of them. It arises from absolute size at realistic g, which is what §5.4 claims and §6.2 measures.

Two respects in which the statement is conservative. It aligns signs that do not align: at the discrete optimum the realised signed coupling is 2.5–4.7× smaller than that aligned sum. And it holds δ fixed, whereas on the real panel δ *falls* as N grows (0.711 at $N = 10$, 0.389 at $N = 20$, 0.240 at $N = 50$, 0.220 at $N = 116$, at gross 1×), so the realised coupling shrinks *faster* than the proposition permits.

The proposition is verified on the real graph rather than assumed: `src/verify_n_invariance.py` measures the coupling's objective gain over own alpha across $N \in \{10, 20, 50, 116\}$ and gross $\in \{1, 5, 20\}$, with the acceptance rule fixed before the run, and §6.2 reports it.

Remark 1 (Lead–lag relations are first order)**.** The graph in this paper validates *contemporaneous* residual correlation, and a reader expecting "stocks that affect each other" may look for lead–lag edges. Proposition 1 says where they would go. Write the return as $r_t = \mu(s_t) + e_t$ with $\mu(s_t) = \mathbb{E}[r_t \mid \mathcal{F}_{t-1}]$; a lagged relation $r_{j,t-1} \to r_{i,t}$ is a term of $\mu_i(s_t)$, and μ_i multiplies w_i inside φ_i. The pairwise terms ψ_{ij} come from the conditional covariance, which is the *innovation* covariance Σ_e and is unchanged by any lag matrix. So a validated lead–lag graph is a *state feature*, first order in position and, by Proposition 2's order argument, not inert; it is never a coupling term. It belongs to the graph-as-feature literature of §2, and this paper's inertness result says nothing about it. Verified as T11 in the verification table below: on a VAR(1) with a cross-lag matrix supported off $\mathrm{supp}(\Sigma_e)$, every ψ_{ij} is identical across lag matrices, φ_i moves on exactly the lagged rows, and a lag-only pair has $\psi_{ij} = 0$ with $\partial \varphi_i / \partial u_{j,t-1} = w_i A_{ij}$.

B.5 NUMERICAL VERIFICATION

An algebraic proof can be correctly derived and incorrectly implemented, so the theory is additionally verified on a synthetic linear-quadratic market where Σ, β and E^* are known *by construction* and ground truth is therefore available (`src/verify_factored_q.py`, seed 20260811). The table is generated at build time from the script's own record of what it asserted, never transcribed: 16 assertions, all passing.

#	Assertion	Statistic
T1	decomposition identity reproduces $r(w)$ exactly	max abs error 1.39×10^{-17} over 500 random w
T2	pairwise terms are supported exactly on supp_off(Σ_u)	17 edges, density 0.187; every off-support term is 0
T3	supp_off(Σ) is complete while supp_off(Σ_u) is sparse	91/91 vs 17/91
T4	market term depends on w only through $\beta^\top w$ (rank one)	expansion error 1.39×10^{-17}; rank 1
T5	aggregate-augmented exact DP equals brute-force argmax	8/8 trials, n=11, 3^{11}=177,147 joint actions
T6	dropping the rank-one market term changes the argmax	7/8 trials
T7	market-neutral scoping moves the coupling into a constraint, does not remove it	7/8 trials
T8	BH at level q controls the expected fraction of spurious coupling terms	mean FDP 0.0877 $\pm$ 0.0041 vs q = 0.1, 200 replications
T8b	a cross-sectional projection manufactures edges and breaks FDR control	mean FDP 0.7938 $\pm$ 0.0026 vs 0.0877 for the time-series regression, target q = 0.1, 200 replications
T9c	the TWO-term bound (estimation error dropped) IS violated, as intended **(deliberately failing)**	violated on 57/200 replications
T9	three-term value-error bound is never violated	200/200 replications
T9b	spurious-edge magnitude scales as $1/\sqrt{T}$	ratio 4.64 vs expected 4.00
T10	BH on raw correlations is near-complete; on market residuals it recovers supp(Σ_u)	raw density 1.000; residual 0.057 vs true 0.061 (synthetic only)
E7	BY controls the expected spurious fraction under negative dependence among statistics	BY mean FDP 0.0129 $\pm$ 0.0024 vs q=0.1 (level q/H_m=0.0170); BH at q gives 0.0814; 26.9% of statistic pairs below -0.20
E8	Prop. 3 converse: ψ_{ij} recovered by the proof's double integral leaves a separable remainder; a genuine three-way term defeats the construction	remainder mixed partials 0.0×10^{0} (pairwise f) vs 2.220 (with w0*w1*w2); third-order partials 0.0×10^{0} vs 1.000
T11	lead-lag relations enter the conditional mean (first order, phi_i) and never the coupling (ψ_{ij} depends on the innovation covariance alone)	identity holds with lagged mu; max—d psi— = 0.0×10^{0} across lag matrices; phi moves on exactly the 3 lagged rows; a lag-only pair has psi = 0 and d phi_i/d u_j,t-1 = w_i A_ij (err 1.6×10^{-12})

T9c deserves comment: it is a *deliberately failing* assertion. An earlier draft of Proposition 4 omitted the estimation-error term, group (c) of (8); T9c asserts that the two-term bound *is* violated on some replication, so that if the shorter version is ever restored the verifier fails.

Proposition 2's criterion, evaluated. The third-order mixed partials are taken by a $2\times2\times2$ central stencil in three distinct coordinates, on the real covariance at a held book, over 200 random triples of held names (`third_order.json`). The mean–variance quadratic's third partial is *identically* zero in exact arithmetic, so what a finite difference returns for it is roundoff and scales as h^{-3}; risk-parity dispersion's is a real derivative and does not move with h. Quoting one step size would therefore quote an arbitrary number, so the step is swept and the separation reported at each:

step h	mean–variance	risk parity	separation
1.00×10^{-4}	1.36×10^{-8}	65.4	9.7 orders
1.00×10^{-3}	1.36×10^{-11}	65.1	12.7 orders
1.00×10^{-2}	1.36×10^{-14}	46.1	15.5 orders

The mean–variance column falls by exactly three decades per decade of h, the signature of an exact zero rather than a small number, while risk parity's is stable to two figures over the two smaller steps and drifts only at $h = 10^{-2}$, where the stencil's $O(h^2)$ truncation error bites. An earlier draft of this paper quoted 5.6×10^{-8} and 3.6×10^{-1} for the same comparison, values that lived only in a source comment, with no artifact behind them and no step size stated. They are replaced here by the measurement.

C THE ACTION SELECTOR

Algorithm 1 Multi-restart coordinate ascent (HEURISTIC; no optimality guarantee)

Require: node table $\varphi \in \mathbb{R}^{N\times|L|}$; pairwise payoffs $\psi \in \mathbb{R}^{|\widehat{E}|\times|L|\times|L|}$; loadings β; restarts R

$a^* \leftarrow \varnothing$; $v^* \leftarrow -\infty$
for $r = 1$ to R **do**
  $a \leftarrow$ all-flat if $r = 1$, else uniform random over L^N
  repeat
    for $i = 1$ to N **do**
      $a_i \leftarrow \arg\max_{\ell\in L}$ of the **true** objective (3) with a_{-i} fixed {scored under the full objective, market term included}
    end for
  until no coordinate changed
  if $Q(s, a) > v^*$ **then**
    $a^* \leftarrow a$; $v^* \leftarrow Q(s, a)$
  end if
end for
return a^*

Scoring each candidate under the *full* objective rather than the factored surrogate matters: the market term g_{mkt} is dense, so a coordinate move that improves the node and edge terms can worsen the total. The same term bounds what "exact" can mean here: on a graph of small induced width the node-plus-pairwise part admits exact variable elimination, but g_{mkt} couples every asset through the scalar $\beta^\top w$, so exactness is per value of a dual multiplier on that scalar, with candidates scored under the true objective; it is a primal heuristic with a duality gap, not a global optimum. Benchmarks are in §6.1 and Figure 7.

Exactness where it can be checked (E17). The FDR-controlled materiality graph at $\varepsilon = 0.30$ (Appendix L, E15) has induced width at most 8 in every window, so the node-plus-pairwise objective can be maximised *exactly* by bucket elimination, giving the first exact joint optimum at $N = 116$ in this work. Over 52 windows $\times$ 6 states, Algorithm 1 returned the exact maximiser, value and actions alike, in **100%** of pairs; and under the exact optimiser the graph changes at most **2** of 116 actions, the same count the heuristic gives. So on every graph where exactness is feasible, the "actions changed" figures are not a selector artifact. The heuristic's inexactness, where it exists, lives in the dense market term: solving the Lagrangian exactly for each of 41 dual multipliers and scoring under the true objective is *worse* than coordinate ascent in 64% of pairs. The dense Benjamini–Yekutieli graph (width ≥ 26, E13) remains beyond exact reach, and there the claim rests on Proposition 3 and the continuous relaxation (D2).

D PROVENANCE OF EVERY MATHEMATICAL OBJECT

What a finite universe costs, and which term pays. A1 assumes approximate factor structure, which is identified only as N grows, so a reader may ask what $N = 116$ costs. Proposition 1 costs nothing: it is an identity for *any* split of Σ into a rank-one part plus a remainder, at any N, sparse or not. Two things are asymptotic. First, the *identification* of the split: Chamberlain & Rothschild (1983) define an approximate K-factor structure as a sequence with $\sup_N \lambda_1(R_N) < \infty$, and their uniqueness result pins $BB^\top$ down exactly at each N *given* that hypothesis, which no single N can verify. Their own statement of the residual cost is that the loadings carry an N subscript, so the premium assigned to an asset can move as more assets enter; D1 (one factor against five nested) and E23 (factors refit out of window) vary exactly that and change no conclusion. Second, the *estimation* of Σ_u from finite samples, whose effect on the value is already bounded: a misallocated split leaves the identity exact and moves mass into Σ_u, which is group (c) of Proposition 4, the estimation term over edges correctly found with magnitudes estimated. Worth noting that their explicit finite-N bound, which holds at every N rather than in the limit, bounds the *pricing error* $\sum_i(\mu_i - \tau_0 - \sum_j \tau_j\beta_{ij})^2$ by $\delta^2\lambda_{K+1}$: it falls on **alpha**, the term §6.3 shows the graph does not inform. Two further points are consistent with A1 rather than against it: a residual with nonzero off-diagonals is their own canonical example of the *approximate* case, a banded R_N with uniformly

bounded eigenvalues, and the sparsity this paper validates is what keeps $\lambda_1(\Sigma_u)$ bounded. Whether it *is* bounded on this panel is not measured here and is not claimed.

This appendix answers, for each quantity the paper uses: what it is, where it came from, and what, if anything, was changed. Objects marked **new** are introduced here.

Table 2: provenance. "Adapted" means the object is used in a setting or with a modification its source does not cover; those modifications are the paper's methodological content.

Object	Definition	Source	What was changed
Log return	$r_t = \ln(P_t/P_{t-1})$	standard	none
Mean–variance objective	$w^\top \mu - \frac{\gamma}{2} w^\top \Sigma w$	Markowitz (1952)	added the ℓ_1 cost term
Proportional cost	$c\|w - w_{\text{prev}}\|_1$	standard execution model	none
Approximate factor model	$\Sigma = \sigma_m^2 \beta\beta^\top + \Sigma_u$	Chamberlain & Rothschild (1983); Fan et al. (2013)	**used to *define* the edge set**, not merely to estimate Σ
Approximate sparsity	$m_p = \max_i \sum_j \|\sigma_{u,ij}\|^q$ bounded	POET, Fan et al. (2013)	adopted verbatim as A1
Covariance thresholding	n/a	Bickel & Levina (2008)	background for A1
Fisher z	$z = \operatorname{arctanh}(\hat{\rho})\sqrt{T-3}$	Fisher's transform (App. E.1)	applied to *residual* correlations
Benjamini–Hochberg	reject $p_{(k)} \le kq/m$	Benjamini & Hochberg (1995)	applied over all $N(N-1)/2$ pairs; retained as a sensitivity arm
Benjamini–Yekutieli	$q/\sum_i 1/i$, arbitrary dependence	Benjamini & Yekutieli (2001), Thm 1.3	**adopted as primary** after measuring the dependence
Factored Q	$\sum_i \varphi_i + \sum_{(i,j)\in E} \psi_{ij}$	Guestrin et al. (2001); Böhmer et al. (2020)	E **derived from** $\operatorname{supp}(\Sigma_u)$ **rather than chosen**
Parameter sharing across edges	n/a	Böhmer et al. (2020)	adopted verbatim
Variable elimination	cost $O(N\|L\|^{\omega+1})$	Guestrin et al. (2001)	measured ω and rejected the method
Max-plus	message passing on the CG	Böhmer et al. (2020)	benchmarked against brute force and rejected
Treewidth lower bound	degeneracy (max k-core), and minor-min-width (contraction)	graph theory, standard	infeasibility certificates; minor-min-width raises the k5 bound from a median of 10 to 21 and decides §6.1
Treewidth upper bound	min-degree, min-fill, and 2,000 random-tiebreak orderings	graph theory, standard	a feasibility certificate: it exhibits an elimination order
Multi-restart coordinate ascent	Algorithm 1	**new**	replaces max-plus; no optimality guarantee, and named accordingly
Risk contribution	$RC_i = w_i(\Sigma w)_i/(w^\top \Sigma w)$	Maillard et al. (2010)	used as an alternative objective
Diversification ratio	n/a	Choueifaty & Coignard (2008)	tested and *discarded* after its control failed
Differential Sharpe	online Sharpe derivative	Moody et al. (1998)	shown to carry no coupling (§5)
Parameter-randomization control	untrained-network baseline	Adebayo et al. (2018)	**transferred from saliency maps to edge reliability**

Algorithm 2 Validated edge inference. Steps 1–4 run on one estimation window; step 5 keeps only what recurs across the four most recent windows.

Require: returns R up to the end of window t; level q

1: $f_t \leftarrow \frac{1}{N}\sum_i R_{ti}$; for each i, regress $R_{\cdot i}$ on $[\mathbf{1}, f]$ by OLS and set $u_{\cdot i} \leftarrow$ residual {**time series**, never cross-sectional (§4.1)}

2: for each pair $i < j$: $\hat{\rho}_{ij} \leftarrow \mathrm{corr}(u_{\cdot i}, u_{\cdot j})$, $z_{ij} \leftarrow \mathrm{arctanh}(\hat{\rho}_{ij})\sqrt{T-3}$, $p_{ij} \leftarrow 2(1-\Phi(|z_{ij}|))$

3: sort $p_{(1)} \le \cdots \le p_{(m)}$, $m = N(N-1)/2$; $k^* \leftarrow \max\{k : p_{(k)} \le kq/m\}$

4: $\widehat{E}_t \leftarrow$ pairs of the k^* smallest p {$q = q_{\mathrm{BY}}$ for the primary graph (§4.1)}

5: repeat 1–4 on windows $t-1, t-2, t-3$ and **return** $\{(i,j) : (i,j) \in \widehat{E}_{t'}$ for at least 3 of $t' = t, \ldots, t-3\}$ {persistence, 3 of 4}

Object	Definition	Source	What was changed
Moving-block bootstrap	resample contiguous blocks	standard (App. E.4)	block length set to span non-overlapping windows
Chance-corrected Jaccard	$\big(J - k^2/n\big)/(2k - k^2/n)$	**new here** (App. E.5)	derived for top-k selection from n edges
GAT attention	$\alpha_{ij} = \mathrm{softmax}_j$ $(\mathrm{LeakyReLU}(a^\top[Wh_i \| Wh_j]))$	graph-attention networks, standard formulation	read edge strength off it; symmetrized over directions
Third-order decomposability test	$\partial^3 f/\partial w_i \partial w_j \partial w_k = 0$	elementary (interaction order ≤ 2); proved for completeness	Proposition 2
Cross-impact cost	$\frac{\lambda}{2}\Delta w^\top \Lambda \Delta w$	linear market-impact models	**assumed** $\Lambda \propto \Sigma_u$; stated as a modelling assumption

E STATISTICAL MACHINERY, DERIVED

E.1 FISHER'S VARIANCE-STABILISING TRANSFORM

A sample correlation $\hat{\rho}$ is awkward to test directly: it is bounded in $[-1, 1]$ and its sampling distribution is skewed whenever the true $\rho \neq 0$, with a variance that depends on ρ. The transform $z = \mathrm{arctanh}(\hat{\rho})$ removes that dependence. Apply the delta method: for T observations $\mathrm{Var}(\hat{\rho}) \approx (1-\rho^2)^2/T$, and $g(\rho) = \mathrm{arctanh}(\rho)$ has $g'(\rho) = 1/(1-\rho^2)$, so

$$\mathrm{Var}(z) \;\approx\; \big(g'(\rho)\big)^2 \,\mathrm{Var}(\hat{\rho}) \;=\; \frac{1}{(1-\rho^2)^2} \cdot \frac{(1-\rho^2)^2}{T} \;=\; \frac{1}{T}. \tag{10}$$

The variance is *free of* ρ, which is the entire purpose of the transform. The refinement from T to $T-3$ is a higher-order bias correction; at $T = 504$ the difference is 0.3%.

Remark 2 (The assumption a reader should attack first)**.** This derivation assumes T *independent* draws from a bivariate normal. Daily equity residuals are neither: they are heavy-tailed (measured excess kurtosis 5.72 at the median, 82.12 at the maximum) and volatility-clustered. That is exactly why the realized false-discovery proportion is calibrated under kurtosis-matched and GARCH errors instead of trusting the nominal level (Figure 3).

E.2 WHY MULTIPLICITY CONTROL IS UNAVOIDABLE HERE

With $m = 6,670$ pairs tested at $\alpha = 0.05$ and no true edges, the expected number of false discoveries is $\alpha m \approx 334$, and $\mathbb{P}(\text{at least one}) = 1 - 0.95^m$ is indistinguishable from 1. Controlling the family-wise error rate by Bonferroni requires testing at $\alpha/m = 7.50 \times 10^{-6}$, so stringent that almost no real edge survives: one trades a graph full of noise for an empty graph. False-discovery control is the right criterion because a perfectly clean graph is not required; what is required is to know what *fraction* of the edges acted upon is real, which is exactly what Proposition 4 converts into a bound on value error.

E.3 Benjamini–Hochberg, Benjamini–Yekutieli, and their guarantees

Given p-values $p_1, \ldots, p_m$, sort them, find the largest k with $p_{(k)} \le kq/m$, and reject the corresponding k hypotheses. If the p-values are independent or positive regression dependent, the procedure satisfies

$$\mathrm{FDR} \;=\; \mathbb{E}\!\left[\frac{V}{R \vee 1}\right] \;\le\; \frac{m_0}{m}\, q \;\le\; q, \tag{11}$$

with V false discoveries among R rejections and m_0 true nulls (Benjamini & Hochberg, 1995; Benjamini & Yekutieli, 2001). The bound is conservative by m_0/m: if half the hypotheses are real, the realized rate is about $q/2$. Replacing q by $q/\sum_{i=1}^{m} i^{-1}$ gives the level used in §4.1, which Benjamini & Yekutieli (2001) prove controls (11) under *arbitrary* joint dependence, the version adopted here, because §4.1 measures the positive-dependency condition to be implausible here.

Remark 3 (Benjamini–Hochberg *is* a magnitude threshold)**.** Because every pair here is tested on the same T, and $z = \operatorname{arctanh}(\hat{\rho})\sqrt{T-3}$ is strictly monotone in $|\hat{\rho}|$, the p-value ordering *is* the magnitude ordering. Two consequences follow that are easy to get wrong. First, the objection "the procedure ignores economic magnitude" is false: it lands on an adaptive magnitude threshold, measured here at $|\rho| = 0.0925$. Second, a top-K-by-magnitude truncation of a discovery set *is itself* a discovery set, namely $\widehat{E}(q^*)$ for the level q^* yielding K discoveries, so such a truncation retains a valid guarantee at a stricter level. This was verified directly: at $q = 10^{-6}$ the discovery set and the top-300 by $|\hat{\rho}|$ have symmetric difference **zero**. An earlier version of this work asserted the opposite, and was wrong.

Remark 4 (FDR level is not a sparsity dial at this sample size)**.** At $T = 504$ a residual correlation of $|\rho| = 0.3$ carries a p-value near 10^{-12}, so edges that strong survive any level one would write down. In 16 of 52 windows, *no* level down to 10^{-10} reduces the induced width below the feasibility cap. Tuning q cannot buy tractability.

E.4 The moving-block bootstrap

The 52 estimation windows are 504 days long and step by 63 days, so consecutive windows share 87.5% of their observations. Resampling windows independently would treat heavily redundant windows as independent facts and give intervals far too narrow. Contiguous *blocks* are therefore resampled, each of $\lceil 504/63 \rceil = 8$ windows, so each block spans non-overlapping underlying data.

Remark 5 (The effective sample size)**.** 52 overlapping windows contain roughly $52/8 \approx 6$ independent ones. Wherever this paper reports a quantity "over 52 windows", the reader should mentally substitute "six independent observations".

E.5 Chance-corrected top-k agreement

For two independently drawn k-subsets of n items the expected intersection is k^2/n, so the expected Jaccard index is

$$\mathbb{E}[J_{\mathrm{chance}}] \;=\; \frac{k^2/n}{2k - k^2/n}. \tag{12}$$

With $n = 1,467$ and $k = 146$ this gives 0.0524, and the reported statistic is $J_{\mathrm{adj}} = J_{\mathrm{obs}} - \mathbb{E}[J_{\mathrm{chance}}]$ throughout.

Remark 6 (Why the chance level is not the right baseline)**.** A chance level shows that a score exceeds random selection. It does *not* show that *training* produced the score. For that one needs a control sharing the architecture and differing only in the thing being credited, which here is an untrained network across random initialisations. In §6.5 that control reverses the conclusion the chance level alone would support.

E.6 Power as a procedure

For a paired comparison with effect size δ, standard deviation σ, significance α and target power $1 - \beta$, the required sample size is approximately

$$n \;\approx\; \left(\frac{(z_{1-\alpha/2} + z_{1-\beta})\,\sigma}{\delta}\right)^2. \tag{13}$$

Power is treated as a *procedure* rather than a post-hoc calculation: run a pilot, measure σ, evaluate (13), and if n exceeds a pre-set cap declare the hypothesis underpowered *in advance* and report it as uninformative rather than as a null. In §6.4 this returned 43,761 from 8 pilot seeds, against a cap of 100; the confirmatory run was not performed (Figure 16). Because σ is itself estimated, the seed count inherits the χ^2 interval on σ, and it was re-estimated on 32 seeds: σ falls from 7.471 to 4.957 and the count to 19,261 $[12,380,34,044]$. The estimand's denominator is a dispersion that spans $281\times$ across seeds and falls below 10^{-4} in 2 of 32 cells, which is why the eight-seed figure was so movable; the scale-free $\log G_1 - \log G_4$ gives 4,356 $[3,790,4,501]$. No estimand and no seed count brings it within an order of magnitude of the cap, so the verdict does not depend on which is chosen; all are reported.

Remark 7 (The implementation used here is conservative by $2\times$)**.** The code evaluating (13) uses the two-independent-samples form, which carries a leading factor of 2, on a *paired* statistic. The paired form has no such factor, so the seed counts reported here are twice what the paired formula gives. No conclusion changes, since both figures dwarf the cap, but the discrepancy is real and is recorded here rather than quietly correcting the number.

F THE RELIABILITY AUDIT IN FULL

If a model learns which assets are related, independent runs should learn the *same* relations. Over 20 seeds, the top 146 of 1,467 edges by learned strength were taken and chance-corrected cross-seed agreement was computed, together with a *parameter-randomization control* (Adebayo et al., 2018): the same architecture, *untrained*. Against the analytic chance level alone (0.0524) the trained edges look weakly reliable; the control removes that reading, since untrained networks agree **more** (0.1895) than trained ones (0.1106), leaving a training contribution of **-0.0789**. On the primary Benjamini–Yekutieli graph a budget curve to **5,000 steps**, ten times the budget that raised the undertraining hypothesis, has a negative trained-minus-untrained difference at every checkpoint, with seed-level intervals excluding zero at four of six and the 5,000-step interval at **-0.060 [-0.107, -0.007]** over 24 seeds, a count fixed in advance from the dispersion at eight; the trend is $+0.001$ per $1{,}000$ steps, indistinguishable from flat. Training does not recover agreement; the contribution replicates at -0.0730. **A chance baseline shows that a score is above random; it does not show that the training procedure produced it.**

The same audit on a multi-head GAT gave an implausibly high cross-seed Jaccard of 0.849, because attention is a softmax over each node's neighbourhood: a node with two neighbours assigns each about 0.33 and one with fifty about 0.02, and against a $1/(d_i{+}1) + 1/(d_j{+}1)$ proxy the Spearman correlation is 0.983. **Raw GAT attention weights are largely a node-degree readout, so their cross-seed consistency is not evidence of stable learned relational structure**, relevant wherever edge importances are read off a softmax over a *variable-size* neighbourhood, which is the mechanism: a node with two neighbours assigns each about 0.33 and one with fifty about 0.02. Architectures that soft-max over a *fixed* number of sampled neighbours cannot produce it, and this was *measured* rather than argued. HATS samples exactly 20 slots per relation whatever the true degree; on its own trained checkpoints the attention–degree correlation is 0.31 and 0.25, against 0.99 for the GAT built here, and its *untrained* control gives 0.01 and 0.01, so what association exists there is learned, whereas in the GAT built here it is architectural and training moves it by 0.001. **The readout finding is a property of variable-size neighbourhood softmaxes, not of attention in general.** Residualising on the degree proxy leaves a difference of +0.0097 at 20 seeds, $[-0.0711, +0.0955]$, spanning zero; at 48 (E6) it is -0.0311, still spanning zero; and at 128 self-contained seeds (E27) it resolves at **-0.0394** $[-0.0543, -0.0239]$, which is 50% of the effect on the module built here, with 100% of 30 bootstrap streams excluding zero and both split halves negative. **The degree readout masks this rather than merely inflating it**: the raw contrast is $+0.0069$, the opposite sign. A study checked on the untrained *level* instead of the contrast would have voided here in one direction and at E6's block in the other. Across three blocks (the two above and a third whose own verdict is discarded as void, although its measurements remain valid measurements), the level spans 0.029 while the contrast spans 0.013, all three negative.

G THE DIAGNOSTIC ON A PUBLISHED SYSTEM: PROTOCOL AND PREDICTIONS

Scope. Six published systems were in scope (Table 1); only HATS (Kim et al., 2019) releases runnable code and data, and it is a supervised movement classifier rather than an RL agent, so the diagnostic is applied to its long/short book (their `filter_topk`: the 15 highest-$P(\text{up})$ names long, the 15 highest-$P(\text{down})$ short; their released code uses 30/30, and both are reported). "Positions changed" is therefore one layer below actions, and this is stated wherever the number appears.

Reproduction gate, fixed before training. The run counts only if the unmodified system's test macro-F1 lands within 0.02 of the figure in their Table 2 for that phase (0.3314, 0.3347, 0.3100 for phases 1–3), evaluated under their own protocol (best-eval checkpoint, that epoch's neighbour sample). Phases 1 and 2 are **void** (0.2857 and 0.3026); phases 3 and 4 pass (0.3098 and 0.3289). Eval-F1 trajectories peaked at epochs 208 and 133 of 300 and were flat thereafter, so the shortfall is not undertraining.

Environment. Their code needs TensorFlow 1.15, which no current GPU stack supports. It was run on TF 2.20 through a 16-line compatibility shim (all renames), admitted only after the same checkpoint on the same inputs under TF 1.15 and under the shim produced a worst class-probability difference of 1.49×10^{-7} and identical books on all 99 test days at both book sizes.

Why the bound does not generalise to the GAT built here (E29). The bound derived above needs the relational pathway to enter additively before a linear readout. That GAT computes `node_utilities` $=$ MLP(GAT(s)) with a ReLU inside, so no exact separation exists and only the Lipschitz form survives: with ReLU 1-Lipschitz, the graph moves a utility by at most $L\|\Delta h\|$ for $L = \|W_2\|_2\|W_1\|_2$. Six agents trained exactly as D8/E6 train them give L between 1.02 and 1.41, so looseness in L is *not* the problem, yet the bound certifies **nothing** on all six. The reason is the ratio it must beat: the allowance $2L\|\Delta h\|$ runs 12–42$\times$ the median decision margin, because in this architecture the graph *is* the representation rather than a small residual: HATS's median $\|R\|/\|S\|$ is 10^{-4}, whereas here it is of order one. Soundness held (no flip inside the safe set at any scale), but both pre-registered usefulness predictions failed: Spearman 0.83 against a 0.90 threshold, and $s(1) > 0.10$ on 0 of six models. The pre-registration committed to narrowing the claim if that happened, and §7 does. A directional bound along the realised Δh, rather than a worst-case one over all directions, might recover something; that has not been tested and is not claimed.

Their published metrics against chance. On the reproduced split, HATS's published accuracy (0.3859) and macro-F1 (0.3100) sit at or below the two baselines the task admits: the majority class (0.3876) and a label permutation (0.3333). Their labels are terciles, so both are computable from the label distribution alone and neither is a strawman. This is reported because a reader comparing the reproduced macro-F1 with the published one should know what range the task's metrics occupy; it is not a claim that the system is uninformative, which its own ablations, not those run here, would have to settle.

The graph they actually use. Their configuration excludes 14 of 85 relations, among them the four densest, carrying 89.4% of the edge mass; the effective union density is 0.1480 over 71 relations, not the 0.9306 of the paper's description. The paper does not mention the exclusion.

Arms. *graph* (theirs); *zeroed* (the relational pathway's output set to zero, which is the architecture-level removal and the figure quoted); *none* (self-loops through their sampler, which is a different graph rather than no graph and overstates the effect sixfold against *zeroed*); *shuffled*; twenty *degree-matched random* graphs, exact by construction because their model takes the per-node neighbour count as a separate input. The comparison is the paired per-day difference between graph-vs-none and graph-vs-random with a block-bootstrap interval and the 0.05 materiality used throughout.

Two defects in the released code (`verify_index_offset.py`). Neighbour lists are 0-based company indices; the lookup table is $\text{concat}([0,\ \text{LSTM}])$ with company c at row $c{+}1$, so index j returns company $j{-}1$ and index 0 the zero pad. The padding mask `att_mask_mat` is computed

and never applied. Both are asserted against their source and data, and the offset is demonstrated on a five-company table whose values are the company indices.

D7b: the paired wiring experiment (pre-registered 2026-09-06). Each phase trained twice, identically, differing only in a $+1$ on the neighbour indices. Predictions: P1, the correction moves macro-F1 by less than 0.0069 (their margin over their strongest baseline): **holds**, mean -0.0020, but between separately trained models whose code seeds nothing, so it carries a seed term; P2, the correction changes fewer book positions than a degree-matched random graph: **not established**; measured seed-free (the other wiring's graph fed to the same weights), the paired interval spans zero in eight of eight models, so the two are indistinguishable and both under 0.001 of positions; P3, neither wiring's graph beats its random control: **holds** in eight of eight.

The bound (pre-registered 2026-09-07). With $z = W^\top(R + S) + b$ and leaky-ReLU monotone, a class flip for company i requires $\max_{c \neq c'} \|w_c - w_{c'}\| \, \|R_i\|$ to exceed its top-two margin without the pathway. Q1 (median $\|R\|/\|S\| < 0.10$), Q2 ($> 90\%$ of company-days out of reach) and Q3 (zeroing the pathway moves fewer positions than a random graph, as a paired interval) were fixed before the probe ran. All three hold on the reproduced model; Q1 and Q2 hold on all eight models (98.2–99.7%); Q3 holds with the interval excluding zero in seven of eight and is indistinguishable in one. The bound was first run on synthetic ground truth where a deliberately large pathway flipped 27.7% of decisions: zero flips fell inside the set the bound calls unreachable there, and zero in every real model. Two-thirds of the pathway's output (0.66) is a common shift across companies, which moves every logit and no ranking.

Profitability (E10, pre-registered before the run). Their Table 3 states the contribution in money, not classification, so the same book was priced with the stored per-day returns. The smallest meaningful effect is theirs: on phase 3 they report Sharpe 1.2503 against the strongest baseline's 1.0642 (margin 0.186) and daily return 0.0703% against 0.0454% (margin 0.0249 pp). Zeroing the relational pathway moves Sharpe by +0.031 and daily P&L by +0.0009 pp [+0.0001, +0.0020], an order of magnitude inside their own margin, on every arm. On phase 4 the same measurement reproduces their profitability as well: the book returns +0.1705% a day against their published 0.1779%, inside their own 0.0350 pp margin, and zeroing the pathway moves it by [-0.0055, +0.0000] pp. On phase 3 the reproduced book's Sharpe is -0.37 against their published $+1.25$, and an earlier draft called that an unexplained discrepancy. It is not: a Sharpe ratio on 99 days carries an annualised standard error near 1.6, and the block-bootstrap intervals are $[-3.32, +2.59]$ on phase 3 and $[+0.44, +7.33]$ on phase 4, each containing the published value (itself a five-run average). Two point estimates had been set side by side without an interval, which is the error of §6.4 committed in this project's own ledger. A sweep of every profitability convention constructible from their code and text (E18) adds two facts about the released repository: it computes no Sharpe ratio, and its profitability metric is a half-universe long/short over a top-30 filter with neutral predictions excluded, not the 15/15 book of their §5.1.

Prior reports (E11). Before either defect was described, the repository's issue tracker was read in full: all 9 issues open and closed, 0 pull requests, 0 forks with commits ahead, checked on 2026-09-08. **Neither the index offset nor the unused mask appears anywhere.** Two other issues bear on the claims made here, and they are cited rather than having their observations restated as new. Issue #3 (open, unanswered since 2020) asks why "*in the data there are 85 relations, 14 of which are marked as not used. However in the paper, totally 72 relations are presented*", so the exclusion of §G was noticed by a user six years ago; what is new here is its size, that the excluded relations carry 89.4% of the edge mass. Issue #4 (open, unanswered since 2020) reports an independent reproduction scoring "*macro f1 : 0.2762 . . . much lower than the figures reported in the paper*", the band in which the two void phases here land. This is evidence that the reproduction difficulty belongs to the released code rather than to the environment or compatibility shim used here.

A third defect: relation attention normalised over the wrong axis (E16). `aggregate_relation_reps` applies `softmax(dense(x), 1)` to an $[R, N, 1]$ tensor, so axis 1 is the *company* axis: each relation receives a distribution over the 423 companies, rather than each company receiving one over the 71 relations. Measured on their own trained checkpoints, the 423-sized axis sums to 1.0000 for every (day, relation), while the per-company

sum across relations is 0.1678, exactly $71/423$, with a range of $[0.08, 8.12]$, so some companies end with roughly $8\times$ the total relational weight of others. On a two-epoch checkpoint that range is $[0.168, 0.168]$: training makes the imbalance severe. Reported as a property of the released code at commit `854b160`, with its consequence bounded by the pathway result above.

The two phases that do not reproduce (E21). Phases 1 and 2 were trained a *second* time, same command, nothing changed; their trainer seeds nothing, so this is an independent draw. All 8 (phase, wiring, draw) cells miss the gate again: the two draws differ by 0.0070 on average while both sit 0.0341 below the published figures, so **the shortfall is 4.9× the draw-to-draw variability** and is a property of the released code rather than of one unlucky run. It holds under both wirings, so it is not the indexing defect either. Issue #4 on their tracker reports an independent reproduction at macro-F1 0.2762, the band in which the runs here land. The paper states what does not reproduce and stops there: nothing here identifies a cause.

What is not claimed. Any explanation of why phases 1 and 2 fall short; or that this generalises beyond one system.

H When the graph's information arrives, and what it is made of

A distinction worth stating first. "Asset correlation" in the Basel capital framework denotes the correlation between an obligor's asset return and a common systematic factor (Ho et al., 2022), an asset-to-*market* quantity rather than an asset-to-asset one. It is therefore the component that the residualization *removes*, not the one that the graph occupies, and the two should not be read as the same object. Measured that way on the present panel it is large: mean 0.6701 across the 116 names, consistent with $\sqrt{0.4056}$ from the market regression. Mega-caps are overwhelmingly systematic, which is why Σ_u is a small residue and the graph on it is sparse.

Everything below is fitted strictly on $t < 1,600$ and scored on the 2,172 days that follow. The market is removed with in-window coefficients *applied* out of window, never refitted, so no out-of-sample quantity sees its own betas. Artifact: `results/tables/graph_timing.json`.

Table 3: out-of-sample R^2 of the market residual $u_{i,t}$ on the mean residual of a peer set, same day and lagged one day. The lag-1 column is the point: no structure, at any granularity, carries predictive content. Random controls are degree-matched (pairwise) and size-matched (clusters).

peer set	peers	same-day R^2	lag-1 R^2
market factor (free to every agent)	n/a	0.4056	n/a
validated graph, all neighbours	12.3	0.0907	-0.0010
validated graph, top-10 by $\lvert\rho\rvert$	10	0.1134	n/a
validated graph, top-4 by $\lvert\rho\rvert$	4	0.1636	n/a
raw top-10 by $\lvert\rho\rvert$, no validation	10	0.0957	n/a
raw top-23 by $\lvert\rho\rvert$, no validation	23	0.0183	n/a
validated, top-23	23	0.0970	n/a
sector labels	15	0.1009	-0.0010
Average-linkage clusters, $K = 20$	10	0.1579	-0.0009
Average-linkage clusters, $K = 40$	4	0.2001	-0.0005
degree-matched random graph	12.3	0.0030	-0.0013
size-matched random clusters	15	0.0038	-0.0009

Concentration, and what validation is for. Peer *count* dominates: four peers is worth roughly double twenty-three, for clusters and pairwise alike. Against raw correlation with no validation at all, validation wins decisively at high k (0.0970 vs. 0.0183 at $k = 23$) and has nearly converged by $k = 4$ (0.1636 vs. 0.1548). **Multiple-testing control earns its keep by deciding where the real edges stop, not by ranking the strongest ones**. This is also an implementation gap in the

agent studied here: the theory's ψ_{ij} is a function of $\Sigma_{u,ij}$, but the ψ network is given only the two endpoints' features and never the edge strength.

Lead–lag by size. Splitting the panel into residual-volatility terciles (a size proxy; the cache carries no volume) and regressing each follower tercile on the lagged leader tercile gives six ordered pairs, every one negative, the largest -0.00084. The present panel spans 0.129–0.237 annualised residual volatility, which is a narrow range of large caps, so this rules out the hypothesis *within* the universe held here rather than in general.

Most of the graph is sector structure, and the fix is not adopted. Adding a leave-one-out own-sector factor to the residualization and re-running the identical BY plus persistence pipeline takes the graph from 715 edges to 154 (78% removed; 11.2% of the original survives). The survivors are *more* same-sector, not less (0.7403, lift 6.11×): removing the sector *mean* leaves pairs that co-move beyond both market and sector. **This is reported, and the market-only graph is kept as primary**, for a reason that is not stylistic. Sector labels here are not point-in-time, and no look-ahead-free residualization reproduces the same edge set: in-window principal components give a *different* graph at every depth tried (Jaccard 0.1159, 0.1967 and 0.2163 at $k = 2, 6, 12$ against the sector-residualized one). Adopting the sector residualization would move a documented look-ahead from a baseline arm into the primary pipeline.

I AN OBJECTIVE WITH NO RETURN TERM: FULL RESULTS

What this arm is, and the baseline it does not yet carry. No agent and no learning appear anywhere in this study: the allocator is the global minimum-variance portfolio $\Sigma^{-1}\mathbf{1}/\mathbf{1}^\top\Sigma^{-1}\mathbf{1}$, arms differ only in $F_{\mathrm{arm}}(\Sigma_u)$, and what is compared is covariance models. The contrast that carries the claim is against a **density-matched random support**, which is a statement about structure, not an estimator race. Read as an estimator race it is incomplete in one respect that should be stated: the practitioner default for minimum-variance covariance is **linear shrinkage** (Ledoit & Wolf, 2004), which is not among the arms here, and Sood et al. (2023) decline the sample covariance for exactly the reason our own diagnostics show, namely that its minimum-variance solution is ill-conditioned and highly levered (arm B2 carries median condition number 2,261 and median gross 6.19× against B3k4's 804 and 3.29×). The validated support also beats an equal-weighted book, a diagonal residual and a clustering covariance, but a shrinkage arm is the comparison a reader from that literature will want and it is left for future work.

Why masking is not an option. The first design set Σ_u's off-diagonals to zero off the support, and it fails on conditioning. At a 250-day window, across three states, the masked validated arm has condition numbers 2.4×10^9–4.5×10^9, minimum eigenvalues down to -1.2×10^{-4}, gross exposure up to $3,156\%$ and a single position at 133% of net asset value, against 2.3×10^3 for the unmasked sample covariance.

The cause is masking at this sample size, not the validated edges, and an earlier draft of this appendix got that wrong. A *single* random mask of the same density is equally broken (condition number up to 4.4×10^9, minimum eigenvalue -3.8×10^{-4}); what made the voided run's random arm appear sound is that it *averaged 20 draws*, which is itself shrinkage the validated arm never received (4.3×10^2 for that average). So the voided comparison was confounded twice over (one mask against a mean of twenty), and the disposition stands for a better reason than first stated. The fixed-support estimator below needs no repair at all: 9.2×10^2 on the same supports and states. That design is void and nothing is drawn from it.

Fixed-support maximum likelihood. The estimator instead maximises $\log\det K - \mathrm{tr}(SK)$ subject to $K_{ij} = 0$ for $(i,j) \notin \widehat{E}$, which is the covariance-selection problem of Dempster (1972), solved by block coordinate descent for known zeros: sweeping column j, the regression is restricted to j's neighbours and the *unpenalised* system is solved. Equivalently it is the zero-penalty limit of an ℓ_1-penalised Gaussian graphical model in which the penalty *pattern* is fixed rather than the penalty tuned, so there is no cross-validated free parameter. The estimate is positive definite by construction and satisfies $\widehat{\Sigma}_{ij} = S_{ij}$ exactly on $\widehat{E}$ and on the diagonal, which is asserted rather than assumed.

Three known answers check the solver before use: diagonal support returns $\mathrm{diag}(S)$ exactly, complete support returns S to 5×10^{-19}, and the off-support precision is zero to 4×10^{-10} *relative* to $\|K\|$, since an absolute threshold is meaningless when $\Sigma_u \sim 10^{-4}$ puts $K \sim 5 \times 10^4$.

Table 4: E34 contrasts on the confirmatory window ($t \geq 1,862$, 90 windows, evaluated on by no other experiment in this paper). Primary endpoint is out-of-sample Gaussian log-likelihood per held-out day; volatility is secondary and reported for comparability. Positive log-likelihood and negative volatility both favour the first arm.

	market residual		market + sector residual	
contrast	$\Delta \log L$	vol %	$\Delta \log L$	vol %
validated top-4 vs. random	+12.86	-10.42	+3.47	-4.99
validated top-10 vs. random	+9.89	-5.54	+3.40	-3.46
validated all vs. random	+4.57	+0.25	+2.98	-4.49
concentration (top-4 vs. all)	+6.43	-4.01	+0.64	+0.27
vs. best clustering arm	+4.53	-6.04	+0.75	-4.08
vs. full sample covariance	+58.27	-18.66	+41.37	-11.72
clustering vs. random clusters	+10.24	-13.87	+2.15	-2.36
validated vs. raw clustering	+0.02	-0.11	-0.44	-0.16

The adversarial pass, and the number claimed. Because Appendix H shows most of the graph is sector structure, the entire design was re-run on market + sector residuals with the decision rule written down first: retain at least 50% of the effect on both endpoints, or the claim is re-scoped. Retention was 27.0% on log-likelihood and 47.9% on volatility, so the verdict is **SECTOR-DRIVEN** and the headline is the *deflated* figure. The axis is nested and asserted so: with the sector factor switched off the pipeline reproduces the market-residual run to 6×10^{-16} on returns and 1×10^{-13} on log-likelihood, with identical edge counts. Validity gates held throughout: every $\widehat{\Sigma}$ positive definite with no repair (worst condition number 1.0×10^4), the maximum-likelihood property asserted to 5.0×10^{-18}, and every null support distinct from the claim support it controls.

Both pre-registered windows. The covariance window was an axis fixed in advance, $W \in \{250, 750\}$, at T/N of roughly 2.2 and 6.5. The headline contrast holds at both, on the confirmatory window: $W = 250$ gives -10.42% $[-17.04, -5.48]$ and $W = 750$ gives -12.75% $[-21.41, -5.76]$, with log-likelihood differences of +12.86 and +13.32. All four intervals exclude zero, so the result is **not** an artefact of the window where the sample covariance is weakest. The windows do differ in the way the axis was designed to expose: the reference contrast (full Σ against diagonal) is -0.84% at $W = 250$ and -20.65% at $W = 750$, so the unfiltered sample covariance is nearly useless at $T/N \approx 2.2$ and strongly useful at 6.5.

Is the interval calibrated? A 21-day variance-ratio comparison can over-reject badly under fat tails and volatility clustering, so the size of the test actually used (a block bootstrap over holding periods) was measured against that null: a common GARCH(1,1) volatility path with $t(5)$ innovations and two arms of identical unconditional variance by construction, 400 replications. Rejection rates are 0.0000–0.0100 against a nominal 0.05, i.e. the test is **conservative, not liberal**. Two consequences follow, and both are stated: the contrasts reported above cleared a bar *stricter* than nominal, and a *non-significant* contrast here is correspondingly weak evidence of no effect.

Scope. This is an *allocation* result, not a rescue of the factored-Q agent. It identifies a different objective, one in which the covariance enters at $O(1)$ because there is no return term to crowd it out, and shows that validated edges used as *sparse precision support* pay there. Two honest limits: the all-edge arm is weak, so the gain requires the top-k restriction rather than the edge set wholesale; and $\widehat{E}$ is validated on *marginal* correlations while it is used here as *precision* support, two objects whose top-k sets overlap at Jaccard 0.0648 on the present data. That it works despite the mismatch is a finding, not a reassurance, and a support validated directly on partial correlations is untested.

Table 5: every load-bearing result under both false-discovery corrections. Benjamini–Yekutieli is the primary pipeline, because its guarantee holds under arbitrary dependence and Benjamini–Hochberg's does not, the positive-dependency condition having been measured implausible (Assumption 1). **No conclusion differs.**

Result	**BH** q=0.10	**BY** q=0.0107	**Verdict**
median edges retained	100%	59%	n/a
residual graph density	0.4062	0.2552	sparser, same direction
paired difference vs raw	0.5774	0.7115	CI excludes 0 in both
implied $\|\rho\|$ cutoff	0.0925	0.1352	larger edges under BY
treewidth ub ≤ 10	0%	0%	exact VE infeasible in both
persistence winner reverses	yes	yes	criterion-dependent in both
actions changed, median	0	0	inert in both
windows with any change	13/52	12/52	below threshold in both
worst window	1.3%	1.7%	below the 5% threshold in both
reliability, training contribution	-0.0789	-0.0730	untrained wins in both
pilot, names differing	60.9%	65.6%	K8 silent in both
pilot, seeds for 80% power	13,626	43,761	underpowered in both

J BOTH FALSE-DISCOVERY CORRECTIONS, SIDE BY SIDE

Table 5 sets every load-bearing result against both corrections. Benjamini–Yekutieli is the primary pipeline throughout, because its guarantee holds under the dependence actually measured in §4.1; Benjamini–Hochberg is retained as a sensitivity arm. **No conclusion differs between them**, and the only systematic difference is sparsity.

K METHODS TRIED, AND WHAT EACH ONE GAVE

Nine decisions in this paper were made by trying more than one method and measuring. Each table below lists what was tried, what it gave, and which was adopted *and why*. Several of the rejected options are the natural first choice, and two of the adopted ones replaced a method after it had already produced a number that had to be withdrawn.

1. Joint action selection. The coordination-graph literature's method against a generic local search, on 30 real 12-asset subgraphs small enough to enumerate all 3^{12} joint actions.

Method	**Exact matches**	**Worst gap**	**Outcome**
Max-plus (Böhmer et al., 2020)	57%	7.95×	rejected: does not converge at $N = 116$
Max-plus + coordinate polish	73%	4.64×	rejected: the polish did the work
Multi-restart coordinate ascent	**80%**	**2.35**×	**adopted**, as a labelled heuristic
Exact variable elimination	100%	0	infeasible on the BY graph (width > 10); feasible on the ε=0.30 FDR graph, where coordinate ascent matches it in 100% of 312 states (E17)

Why. Coordinate ascent wins on every axis measured (more exact matches, a smaller worst-case gap, and ∼90× faster), and an ablation showed random initialisation beat max-plus outright, so the message passing was contributing nothing. Neither carries an optimality guarantee and the paper says so wherever a selector output appears.

2. False-discovery correction. Benjamini–Hochberg's guarantee needs positive dependency, which was measured to be implausible here.

Method	Edges kept	Guarantee	Outcome
Benjamini–Hochberg, $q = 0.10$	100% (reference)	needs PRDS, not established	sensitivity arm, quoted at realized 0.12
Benjamini–Yekutieli, q/H_m	**59%**	**arbitrary dependence**	**adopted as primary**

Why. Every load-bearing result was re-run on the stricter graph and none changed (App. J). A paper criticising a literature for not controlling false discovery should not decline the correction that removes an assumption it cannot support; 41% of the edges is what that costs.

3. Building a sparse graph. Four routes to thinning the complete correlation graph, scored by whether the result is sparse, guaranteed, and tractable for exact elimination.

Route	Median $\lvert E\rvert$	FDR	Feasible windows	Outcome
Raw correlation + BH	n/a	yes	n/a	rejected: density 0.98, near-complete
Market-residual + BY	1458	yes	0/52	**adopted** for every decision measurement
Five-factor residual	862	yes	0/52	sparser, still infeasible (E13)
Magnitude cut $\lvert\rho\rvert \geq 0.30$	253	**no**	29/52	diagnostic only: no guarantee
Materiality null $\lvert\rho\rvert \leq 0.30$, BY	**95**	**yes**	**52/52**	**adopted** where tractability is the question

Why. The last row is the only construction that is simultaneously sparse, false-discovery controlled and exactly solvable, and it is *sparser* than the unguaranteed magnitude cut at the same threshold, so the guarantee costs nothing here and buys tractability. The residual-BY graph remains primary for every *decision* measurement because that is the graph the theory is about; the materiality graph answers the tractability question specifically.

4. Treewidth bounds. Two lower bounds and three upper bounds, and the choice decided a claim.

Direction	Method	Median, one-factor graph	Outcome
Lower	Degeneracy (max k-core)	31	superseded
Lower	**Minor-min-width**	**44**	**adopted**
Upper	Min-degree, **min-fill**	**84**	**adopted**
Upper	2,000 random orderings	85	kept, but improved on min-fill in only 19% of windows

Why it mattered. Degeneracy is valid but loose. On the five-factor graph it reported a median width of 10 (minimum 5), which had been read as *exact elimination is no longer provably infeasible in about half the windows*. Minor-min-width raises it to 21 (minimum 11) and all 52 windows are infeasible again. **A conclusion that rested on the weaker of two available bounds was wrong, and the reversal is reported rather than the original.**

5. Estimating a graph-restricted precision matrix. This choice reversed a finding previously reported in this project.

Method	Ridge needed	Verdict it produced	Outcome
Zeroing covariance entries	n/a	baseline beat every graph	rejected: marginal, not conditional independence
Symmetrised neighbourhood + ridge	105/204 fits	"two neutral criteria disagree"	**withdrawn**
Graph-constrained Gaussian MLE	**0/204**	validated graph wins on both	**adopted**

Why. The ridge was needed in half the fits and the disagreement it produced was an artifact of it. Under the maximum-likelihood estimate for the graph's exact zero pattern, which needs no ridge and was self-tested on synthetic ground truth, the validated graph beats both the naive and the random graph on *both* criteria, intervals excluding zero.

6. Ablating a published system's graph. Three ways to remove a graph from HATS, and they differ by a factor of six.

Arm	Book positions changed / day	Outcome
`none`: empty neighbour lists	n/a	rejected: divides by the original counts, or by 10^{-10}
`none`: self-loops via their sampler	0.61 of 30	rejected: a different graph, not no graph
`zeroed`: pathway output set to 0	**0.10 of 30**	**adopted**: the architecture-level removal

Why. Only the third answers "what does this model do without its graph". The self-loop arm overstates the graph's effect sixfold because it still feeds a representation through the attention and aggregation. The first was caught by a validity check when it changed exactly zero positions.

7. Running a 2019 codebase. Two routes, one admitted only after measurement.

Route	Result	Outcome
NVIDIA's TF-1.15 fork, Python 3.8	does not build on current free GPU tooling	rejected
TF 2.20 compatibility shim	worst probability difference 1.49×10^{-7}, every book identical	**adopted**

Why. "It is only renames" is an argument; identical long/short books from the same checkpoint on the same inputs is evidence. The shim was gated on that comparison and would have been abandoned had a single book differed.

8. Comparing two arms. The methodological choice this project got wrong most often.

Method	What happened	Outcome
Two point estimates side by side	produced four verdicts that had to be withdrawn	rejected
Two confidence intervals	the difference-of-significance fallacy	rejected
Paired difference with an interval	verdicts flipped when applied	**adopted everywhere**

Why. Each of the four cases was caught by re-scoring an existing result, not by a number looking wrong. Three of them changed a verdict.

8b. The resampling unit, once the comparison is across seeds. Cross-seed agreement is an average over *seed pairs*, and the obvious resampling unit is therefore the wrong one.

Method	Unit	At 24 seeds, 5,000 steps	Outcome
Bootstrap over seed *pairs*	pair	pairs share seeds; intervals too narrow	rejected *a priori*
Per-seed paired test	seed	-0.056 $[-0.081, -0.032]$	agrees, but anti-conservative
Leave-one-seed-out jackknife	seed	-0.060 $[-0.113, -0.006]$	agrees
Seed-level bootstrap	seed	$\mathbf{-0.060}$ $[\mathbf{-0.107}, \mathbf{-0.007}]$	**adopted**

Why. Seed pairs are not independent, because each seed appears in $n-1$ of them, so resampling pairs understates dispersion. The seed-level bootstrap is the widest of the three valid options and was adopted for that reason; the other two are reported because a verdict that survives only under its most favourable test is not a verdict. All three agree here. They did *not* agree at eight seeds, where the bootstrap upper bound wandered across zero under Monte Carlo re-draw alone (8% of 40 streams excluded zero) while the per-seed test already read as significant, which is why the eight-seed result was refused in advance and the seed count raised to 24, at which 100% of streams exclude zero.

9. Objectives, and where the coupling lives. Whether a graph can matter at all is a property of the objective, not of the graph.

Objective	Decomposes?	Actions changed vs control	Outcome
Mean–variance with costs	**yes** (exact)	0 of 116	**the paper's setting**: coupling exact, effect inert
Mean–variance + quadratic cross-impact	**yes**	0 of 116 under a valid Λ	inert; the 18 first reported was an indefinite matrix (E22)
Risk-parity dispersion	no	18 vs 4 random	secondary: graph is a prior, not the structure
Diversification ratio	no	65.5 vs **67.5** random	discarded: control moved more
Differential Sharpe	**no coupling at all**	exactly 0	not well-posed: no structure to occupy

Why. Proposition 2 says which objectives admit the construction, and the measurements agree: the one that decomposes exactly is the one where the coupling is provably inert, and the published reward family (Differential Sharpe) has no coupling structure for a graph to occupy at all.

L THE PRE-REGISTERED STUDIES

The four void continuous-relaxation attempts. D2 took five runs. The first used a Σ that was not positive semi-definite, so its "optimum" sat on a direction of negative risk. The second rescaled the solution to a target gross exposure, which destroys the optimality the comparison depends on. The third and fourth pinned gross exposure as a constraint, which in mean–variance *forces* the implied risk aversion to 158 against the 3.0 used everywhere else, which makes it a different problem rather than a different solver. Every one produced numbers that looked reasonable; every one was caught by the same pre-registered rule, that a maximiser must score at least as well as its comparator under its own objective. The fifth run, box-constrained at the project's own γ, is the one §6.2 reports.

The table carries one row per pre-registration directory. A few study codes name sub-analyses pre-registered inside another study's document (E10 and E11 within D7) or follow-up analyses carrying conclusions only (E21); those have no row of their own. Each study has its own pre-registration fixing predictions, decision rules and the outcome that would count against it before the run. The notes below record what is not visible in the table.

E3 (discrete regime re-measurement): three void runs. D6 was re-measured with the discrete $\arg\max$, and the first three attempts were void, each caught by a validity check written before the run, none by a number looking wrong. The first applied a transaction cost the continuous problem it re-measures does not have; the second used one selector seed where the claim it tests uses six; the third revealed that D6's "slack" volatility target is in fact *binding* for the with-graph arm, so the two arms were not solving the same problem. The instrument that finally licensed the study was a harness check reproducing the earlier claim in 52/52 windows; once it passed, the rows were *re-scored* from the saved run rather than recomputed. Counting these and E20, a whole study voided on its exchangeability check, five studies contributed eleven void runs between them.

Gate F (false-discovery calibration). The FDR-discovery literature drawn on here (Uematsu & Yamagata, 2024) assumes i.i.d. sub-Gaussian errors, and the Fisher-z statistic used here carries a normality assumption of its own; real equity residuals satisfy neither, so the gap was measured and

Table 6: the pre-registered studies and their verdicts. Protocols, the eleven void runs, the two withdrawn verdicts and the corrected one are in Appendix L.

	Question	**Outcome**
D1	Why only one factor?	Density 0.406→0.145; still 0 actions changed
D2	Artifact of the action grid?	No: 1.32% of the book. **4 of 5 runs void**
D3	Hold on the policy's own states?	Yes; the i.i.d. protocol was the *generous* one
E1	Hold on a *trained* agent's states?	Yes; median 0, max 2.6% of names over 2,016 agent-days
D4	Are the edges economically large?	No: strongest 253 edges alone change nothing
D5	Quadratic cross-impact costs	Effect appeared and **was an artifact**: its Λ were indefinite; **superseded by E22**
E22	D5 with valid cross-impact matrices	Zero actions changed on every support at the calibrated λ; max 6 at $4\times$
D6	Binding volatility budget	Effect appears; rounded metric, **superseded by E3**
E3	D6 with the discrete argmax	Graph changes 1→16.5 of 116 names as the budget tightens; validated beats random by 1–3 names, **below materiality**
D7	Diagnostic on a published system (HATS)	Two of four phases reproduce the published macro-F1 (gap 0.0002); on it the graph does not beat a degree-matched random graph; §6.6
D7b	Wiring, and could *any* graph matter?	Their released code reads neighbours off by one; correcting it is indistinguishable from a random rewiring; 98.2–99.7% of decisions provably out of any graph's reach
D8	Reliability for a real GAT?	Verdict **withdrawn**: attention is a degree readout
D9	A criterion-neutral edge test	Ridge-regularised fit; verdict **withdrawn** by E2
E2	D9 with the graph's exact MLE	Validated edges win on **both** neutral criteria, intervals excluding zero
D10	Is PRDS plausible? BY's cost?	PRDS implausible; BY keeps 59.0% of edges
D11	FDR guarantee under a width cap	Yes; it always had one, at $q^* \approx 1.1 \times 10^{-7}$
D12–13	Noise floor; higher frequency	Partly answered in §6.4; intraday **not run**
D14–17	Everything again under BY	All eighteen predictions held (App. J)
E5	Does a bigger budget recover edge agreement?	No, at ten times the budget: -0.060 $[-0.107, -0.007]$ over 24 seeds, trend flat
E6	Same audit on a real GAT, powered	Degree-corrected -0.0394 $[-0.0543, -0.0239]$ at 128 seeds (E27): C25 generalises; the raw contrast $+0.0069$ hides it
E12	An FDR guarantee for *material* edges	Shifted-null BY on $\lvert\rho\rvert \le \varepsilon$; actions changed stay under materiality at every ε
E13	Exact elimination after sparsification?	Factor removal buys none (52/52 certified *infeasible* under a tighter bound); magnitude thresholding does
E23	D1 with factors fitted out of window	Denser graph (density $+0.12$ $[+0.07, +0.20]$ at k=5), still 0 actions changed
E17	Is "0 of 116" a selector artifact?	No: on the exactly solvable graph the exact optimum changes ≤ 2 actions and the selector is exact in 100% of states
E16	Is HATS's attention a degree readout too?	No: Spearman 0.31/0.25 against 0.99 for the GAT built here; its softmax is over a fixed 20 slots
E20	E6 replicated at 128 seeds	**Void**: exchangeability check failed, since E6's untrained block is a 1-in-40 high draw (percentile 100 of 20 blocks), E20's is typical; the reported contrast agrees across both
E19b	The noise floor at 32 seeds	19,261 $[12, 380, 34, 044]$ seeds; the 8-seed figure moved $75\times$ under leave-one-out, this one $2.0\times$
E26	Where does that noise come from?	74% training seed, 26% evaluation window; removing evaluation noise entirely leaves 28,579
E27	C25 on graph attention, self-contained	**Yes**: -0.0394 $[-0.0543, -0.0239]$ at 128 seeds; the raw contrast $+0.0069$ masks it
E24	E3 at the true discrete optimum	Validated beats random by +2.8 names, below the 5% line and inside the selector's own noise floor
E15	Can one graph be both guaranteed and tractable?	Yes: at $\varepsilon = 0.30$ all 52/52 windows are certified solvable, width exactly 5, none empty
E18	Every profitability convention in their code	Their repository computes no Sharpe ratio, and its profitability metric is a half-universe long/short over a top-30 filter with neutral predictions excluded, not the 15/15 book of their §5.1
E28	Is the reachability bound a diagnostic or only a cap?	A diagnostic: it tracks the realised effect at Spearman 0.98 over a $2{,}048\times$ sweep, and the system needs a $2048\times$ larger pathway to reach half its decisions
E29	Does that bound generalise past a linear readout?	**No**: sound but vacuous on 6 of 6 models, the allowance running 12–$42\times$ the decision margin. The number rules a graph *out*, never in
E30	How many seeds must be averaged?	Sixteen: disjoint-half agreement 0.508 trained against 0.438 untrained, the sign reversing at $K = 8$. The deficit is *variance*, not absence
E31	The conditional-mean channel, which is $O(1)$	Channel confirmed (effect ratio 0.67 from $1\times$ to $20\times$ gross, so it does not scale like the coupling) but the graph is **empty**: 2 edges of 13,340 ordered pairs
E32	Does the graph in the *representation* earn its level?	**No**: a GAT moves 11–39% of actions, but a count-matched random graph moves as many (-0.0181 $[-0.0247, -0.0109]$)
E33	Feudal budget partition by graph community	**Void**: the per-group caps did not bind on the groups that mattered, so all three arms returned exactly zero. Three validity checks passed and none asked whether the constraint constrains
E34	An objective with no return term	Validated top-k support beats a density-matched random support by -10.42% volatility, deflated to -4.99% by a pre-registered adversarial pass; §6.7

the nominal level tested against it. (The pipeline uses Fisher-z on partial correlations with BY/BH, not the debiased-lasso procedure of that paper, whose lagged component was not built.) Measured residual excess kurtosis is 5.72 at the median, roughly $11\times$ the heaviest errors simulated in the source method, whose total kurtosis is 3.53 (their §5.1, verbatim: *"skewness 1.86 and kurtosis 3.53"*), and matching a Student-t gives $\nu = 4 + 6/\text{excess} = 5.05$. Over 200 replications the worst realized false-discovery proportion is 0.1159, inside the pre-registered 0.15 threshold, so the Benjamini–Hochberg arm is usable at its *realized* FDR of ≈ 0.12, which is the figure quoted, never the nominal 0.10. The primary Benjamini–Yekutieli pipeline needs no such calibration, because its guarantee depends on neither the tails nor the dependence structure. The *first* run of this gate returned FDP 0.1834 and appeared to fail; it was not a failure but an invalid instrument, since statistical power was 0.007 and the proportion was computed over a denominator of roughly one discovery. The simulator was repaired rather than the threshold, and a minimum-power precondition was added (≥ 0.20; observed 0.256).

D2 (continuous relaxation): four void runs. Described above. **No intermediate number looked implausible** in any of them; all four were caught by one rule fixed before the run. Pre-registration is usually defended against selective reporting; here it guarded against being confidently wrong about the construction under study, which is arguably the more valuable function. Two rival explanations for the null were tested inside the surviving run and both fail. At zero transaction cost the no-graph policy moves 76 of 116 names off its previous position (it is fully active), and the graph still changes nothing; at 41 action levels only 4 of 116 names sit at a grid boundary, and the graph still changes nothing. The median largest single-name move is 0.335 of one grid step. The continuous maximiser is *local*: the graph-restricted Σ is not positive semi-definite in any window (thresholding a covariance's entries does not preserve it), so the relaxed problem is not concave and coordinate ascent from six restarts carries no global guarantee, the same limitation as the discrete selector, which is the like-for-like comparison intended. What a global optimum could add is bounded by the order argument of Proposition 3, not by the search.

D1 and E23 (factor residualisation). D1 fitted its market loading and principal components on the window it then tested. E23 refits every factor on the *preceding* window: the residual graph is denser (density $+0.121$ $[+0.072, +0.198]$ at $k = 5$, paired over 44 windows), and the coupling still changes 0 actions. The in-window fit was conservative for the claim: it made the graph sparser, not the coupling larger.

E1 (trained-agent occupancy). The greedy check below cannot answer whether a *trained* agent visits states where the coupling matters, and §6.4 shows training can amplify a mechanism. Eight independently trained mean–variance factored-Q agents were rolled out of sample for 252 days each and the with/without-graph $\arg\max$ counted at every state they visited: median 0, mean 0.196 [0.151, 0.244], maximum 3 of 116 on any of 2,016 agent-days. All 16 runs passed two pre-registered validity checks (no degenerate policy; the i.i.d. comparator reproducing §6.2 on the same days). The adversarial variant, with agents trained *with* the graph so that the coupling itself shaped their occupancy, gives 0.207, a paired $+0.011$ $[-0.054, +0.077]$. One pre-registered prediction **failed**: D3's tenfold concentration of flips on high-turnover days is *absent* here (0.180 against 0.211), because a trained agent's turnover is 0.385 against the greedy policy's 0.0039; with no quiet baseline there is nothing for flips to concentrate against, so that mechanism statement belongs to a policy at a fixed point.

D3 (on-policy states). Mean turnover is 0.0039 on-policy against 0.2092 i.i.d. Among the residual changes, the rate on above-median-turnover days is 0.145 against 0.015 on the rest, a tenfold concentration that sharpens the mechanism statement: the coupling can only alter a decision that was already close to changing.

D4 (economic magnitude). The implied cutoff is $|\rho| = 0.0925$ [0.0870, 0.0980], with 50.5% of selected edges below 0.15 and 90.2% below 0.30. A fixed-threshold graph carries no false-discovery guarantee and is used here as a diagnostic only; the principled version, a shifted null $H_0 : |\rho| \leq \varepsilon$ under Benjamini–Yekutieli, so the discovery set carries a valid guarantee *for the claim* $|\rho| > \varepsilon$ (E12), equals the BY set at $\varepsilon = 0$, is a subset of the fixed-threshold graph at every ε, and changes at most 4.3% of names.

D8 (graph attention): a withdrawn verdict. All four pre-registered predictions "held" on a raw difference of -0.0007 with an implausibly high Jaccard across twenty independent initialisations. The degree analysis explained why, and the verdict was withdrawn and replaced by the degree-corrected analysis reported in §6.5. The Spearman correlation of raw attention with the degree proxy is 0.983 trained and 0.984 untrained.

D9 and E2 (criterion-neutral edge tests). Persistence alone is **criterion-dependent**: scored by residual significance the validated graph persists at 0.593 against naive correlation's 0.466; scored by raw correlation the ordering reverses (0.677 against 0.420). Each method wins under the criterion its own pipeline optimises. Two criteria that *neither* pipeline optimises were therefore built: out-of-sample Gaussian log-likelihood under the graph-restricted precision matrix, and held-out portfolio-variance forecast error on random books. D9's first estimate of that precision matrix needed a ridge in 51.5% of fits and reported the two criteria *disagreeing*; its pre-registration required a proper constrained maximum-likelihood fit before the number could stand. E2 supplies it (the graphical-model MLE for the graph's exact zero pattern, self-tested on synthetic ground truth; 204/204 fits converged with no ridge), and **the disagreement was the ridge**. Under the MLE the validated graph beats naive correlation on log-likelihood by +1.88 [+1.54, +2.16] and random edges by +4.74 [+3.43, +5.98], and on variance-forecast error (lower is better) by -0.019 [-0.023, -0.015] against naive and -0.011 [-0.017, -0.005] against random, with every interval excluding zero. On criteria neither pipeline targets, validated edges are the better model both of the joint distribution and of the quantity a portfolio depends on; the D9 reversal is withdrawn. This says nothing about whether those edges change *decisions*, which is §6.2's question and is answered there in the negative.

The first run of D9 was also void. It zeroed entries of the *covariance* (marginal independence) rather than of the *precision* matrix (conditional independence), which is the graph-restricted object, and in that version arm G3 came out identical to G4 because Benjamini–Hochberg's ordering *is* the magnitude ordering (Remark 3). Both defects were found by the pre-registered validity check.

M HYPERPARAMETERS, CONSTANTS AND EXECUTION ORDER

What would close the limitations that remain. Four limitations in §7 cannot be closed with the data on hand, and each has a definite closer. *Survivorship* needs a point-in-time constituent list, which converts the panel from present-day large caps to the universe an agent would actually have faced; every study here reruns unchanged on such a panel. *Intraday* needs bar data beyond the 730 days that the data source provides; the theory is frequency-free, and the empirical question is whether residual coupling is larger at higher frequency, where microstructure induces it. *A real cross-impact matrix* needs execution data; E22 shows that on every support tried a *valid* matrix is inert, so what execution data could add is a support that none of those tested here resembles. *The performance hypothesis* needs 19,261 $[12,380,34,044]$ seeds at the dispersion measured over 32 of them, and **no evaluation design removes that**: decomposing the variance (E26) puts 74% of it in the *training seed* and only 26% in the evaluation window, so averaging each agent over infinitely many windows would still leave 28,579, which is 285$\times$ the cap. Two seeds produce different agents, and nothing about how one evaluates them changes that. This is a fact about the comparison, not a gap in how it was executed.

Why no GPU. One forward pass of the 6,796-parameter network takes 0.34 ms against 75 ms for one joint-action selection, leaving tensor algebra 0.46% of the work, so by Amdahl's law an *infinitely* fast accelerator would buy 1.005$\times$ end to end. The rest is sequential coordinate ascent with data-dependent branching. (The published-system study of §6.6 is the exception: HATS is a TensorFlow model and was trained on two T4 GPUs.)

The six arms. They differ *only* in which pairs carry a ψ_{ij} term. **G0**: none. **G1**: random pairs, *count*-matched to G4 (same number of edges, placed at random, so the mean degree matches and the degree *sequence* does not; the HATS controls of Appendix G are degree-matched exactly and say so), the control that rules out "more edges helped". **G2**: sector membership, which carries a **disclosed weak look-ahead** because its sector labels (Yahoo Finance's taxonomy, not GICS) are current rather than point-in-time. **G3**: naive rolling correlation, unvalidated, edge-count matched to

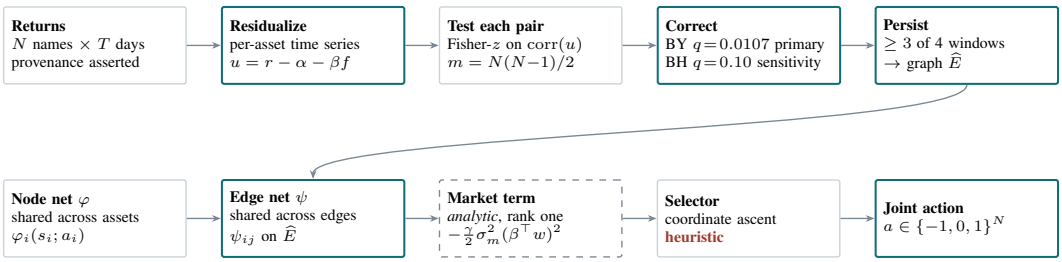


Figure 2: the method end to end. Top row: edge inference, one graph per estimation window (Algorithm 2). Bottom row: the factored action-value function (3) and joint-action selection. Only the edge network is restricted to $\widehat{E}$; the market term is analytic because Proposition 1 shows it is the reward's own structure.

G4. **G4**: the validated residual graph of Algorithm 2. **G5**: learned attention payoffs on a dense-capped topology. Every arm asserts its provenance before a gradient step is taken.

Constants. $\gamma = 3.0$; transaction cost $c = 0.001$ (10 bps per unit turnover); discount 0.99; action levels $L = \{-1, 0, +1\}/N$; gross target 1.0; estimation window 504 days; step 63 days; persistence 3-of-4; $q = 0.10$ (sensitivity arm) and $q_{\mathrm{BY}} = 0.01066$ (primary); selector restarts 8 for evaluation and 2 during training; block-bootstrap block 8.

Every number in this paper is loaded at build time from a saved artifact; the manuscript source is generated, and a missing artifact fails the build rather than producing a plausible value. No figure contains a hardcoded literal. Provenance is asserted at data load and the pipeline raises rather than falling back to simulated data. Execution order: `data_pipeline` → `measure_c11_c15` → `gate_fdr` → `verify_factored_q` → `benchmark_c18` → `benchmark_selector_v2` → `verify_n_invariance` → `gate_option3` → `gate_option3b` → `h3_persistence` → `h2_reliability` → `h2_budget` → `run_h4_pilot`, followed by the study scripts.

N SUPPORTING FIGURES

The figures below support the experiments of §6. They are placed here rather than in the main text because each restates in graphical form a result the text already states numerically.

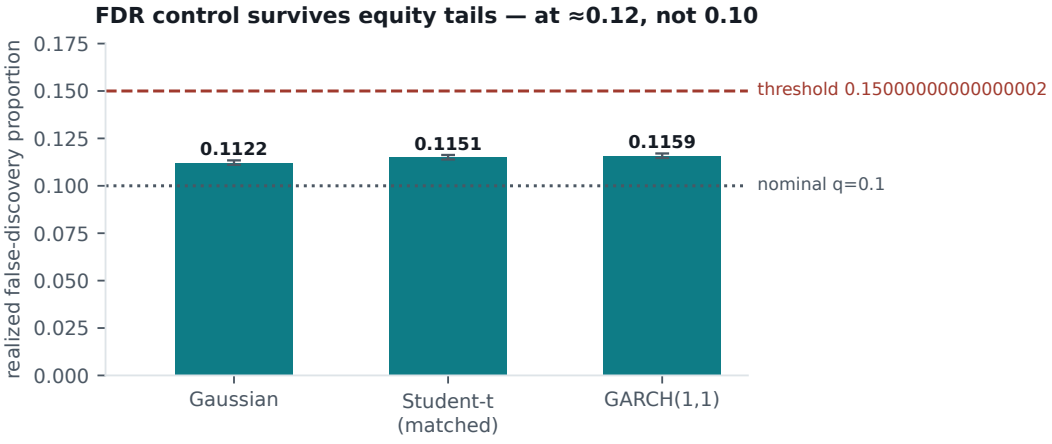


Figure 3: realized false-discovery proportion under Gaussian, kurtosis-matched Student-t and GARCH(1,1) errors, at two true edge densities (§L, gate F). The overshoot is monotone in tail weight, which is what makes it a genuine tail effect rather than simulation noise.

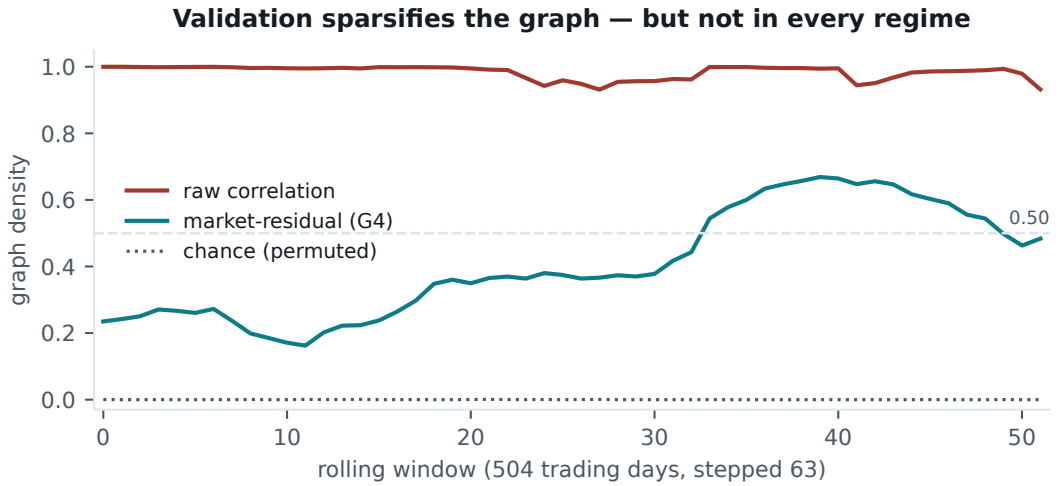


Figure 4: edge density per estimation window for the raw correlation graph, the validated residual graph and a permutation null (§6.1). The raw graph is all but complete, which is the first consequence of Proposition 1, measured.

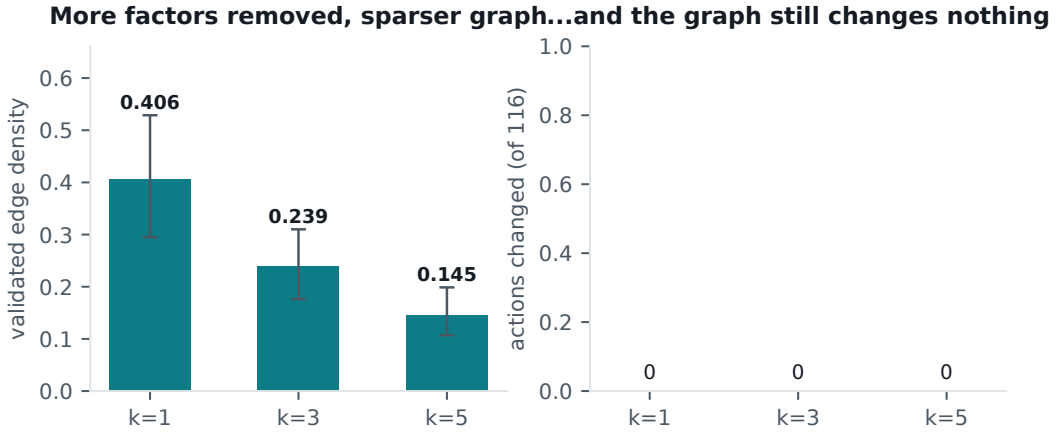


Figure 5: removing more factors gives a sparser graph and changes nothing. Density falls by 2.8× from one to five factors and edge count by 65%, while the actions changed by the graph remain zero at every k.

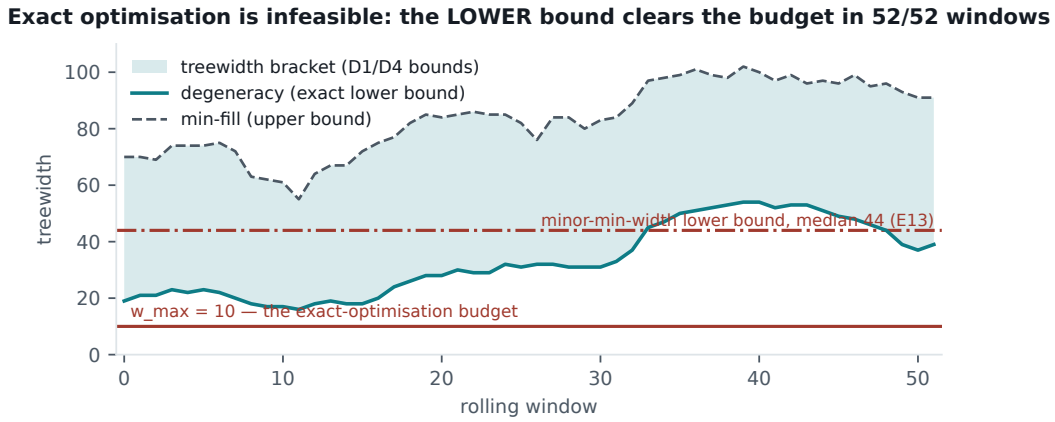


Figure 6: induced width per window, bracketed between a lower bound (degeneracy, tightened to minor-min-width in E13) and min-fill/min-degree upper bounds (§6.1). Exact variable elimination needs $|L|^{\omega+1}$ table entries, so on the residual graphs the lower bound alone rules it out; on the magnitude-thresholded graph the upper bound falls far enough to admit it.

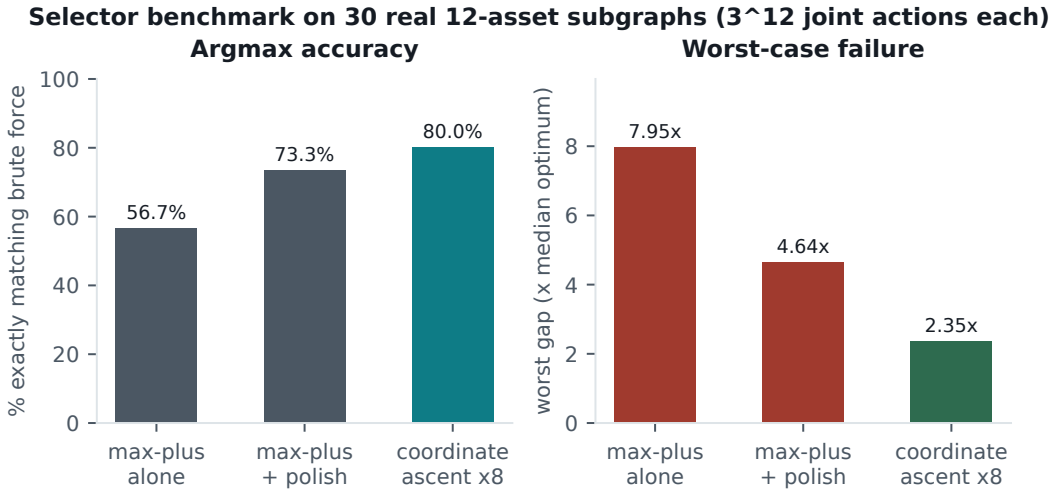


Figure 7: action selectors against the exhaustive optimum on real 12-asset subgraphs (§6.1). Max-plus is compared to multi-restart coordinate ascent; the median gap hid a strictly suboptimal tail (§6.1).

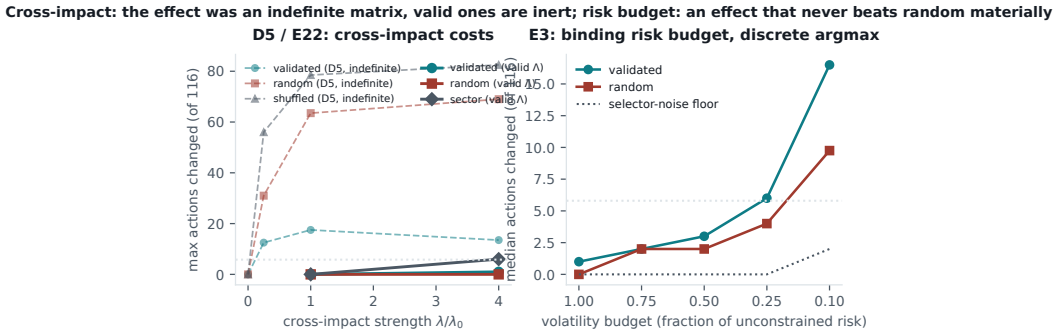


Figure 8: the two regimes in which the order argument of §5.4 predicts the coupling should matter (D5, repaired as E22; D6 re-measured discretely as E3). Under cross-impact the dashed curves are D5's arms as built, whose Λ are indefinite and whose counts are therefore artifacts; the solid curves are the same supports repaired to valid cost matrices, on which every arm changes zero actions at the calibrated λ. Under the risk budget the validated graph moves slightly *more* than random; at the true discrete optimum (E24b) that margin resolves at $+2.8$ $[+0.9, +4.7]$ names, below the 5% line (dotted) and inside the selector's own noise floor (dotted curve) at the median.

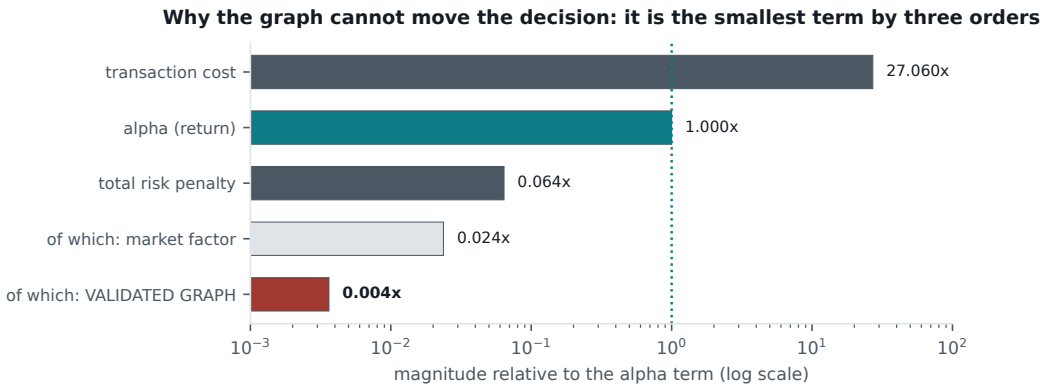


Figure 9: magnitudes of the terms of (7) on random books at gross $1\times$ (§6.2). The validated-graph contribution is three orders of magnitude below the alpha term whose sign it would have to overturn.

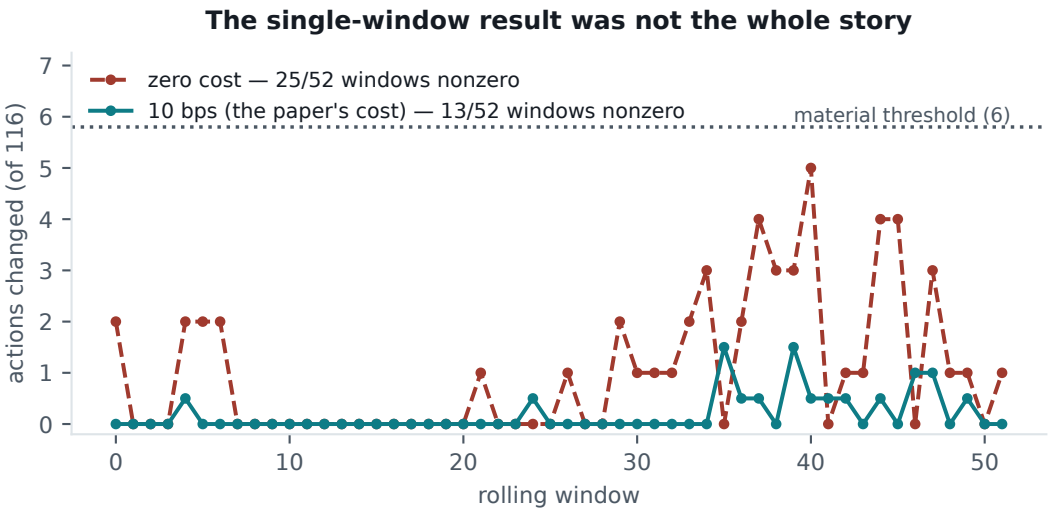


Figure 10: the central measurement, across all 52 estimation windows at two transaction-cost levels (§6.2). The median window changes no actions; a minority (the densest graphs) change one or two, and every window stays below the pre-registered materiality threshold. The first version of this claim was measured at a single window.

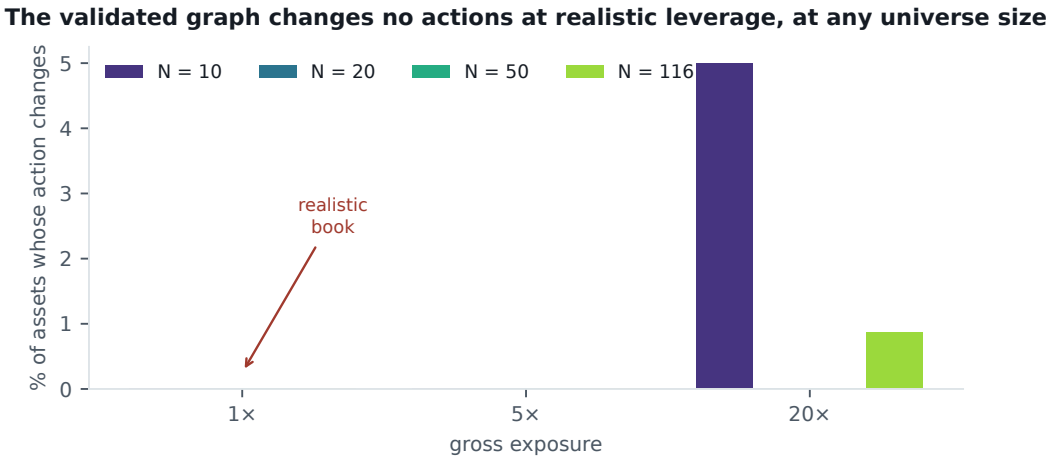


Figure 11: actions changed by the coupling across the sweep over universe size, gross exposure, risk aversion, transaction cost and action granularity (§6.2). An effect appears only at $20\times$ leverage, where the $O(w)$ scaling says it should.

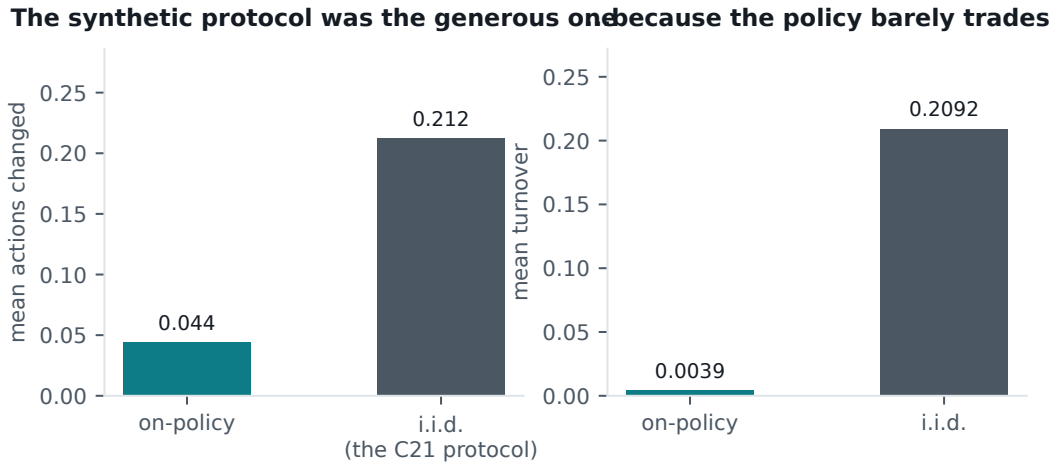


Figure 12: on-policy versus i.i.d. state distributions (D3). Once the greedy policy settles it barely trades, and those are precisely the states in which a small coupling cannot tip anything, so the synthetic protocol was the generous one.

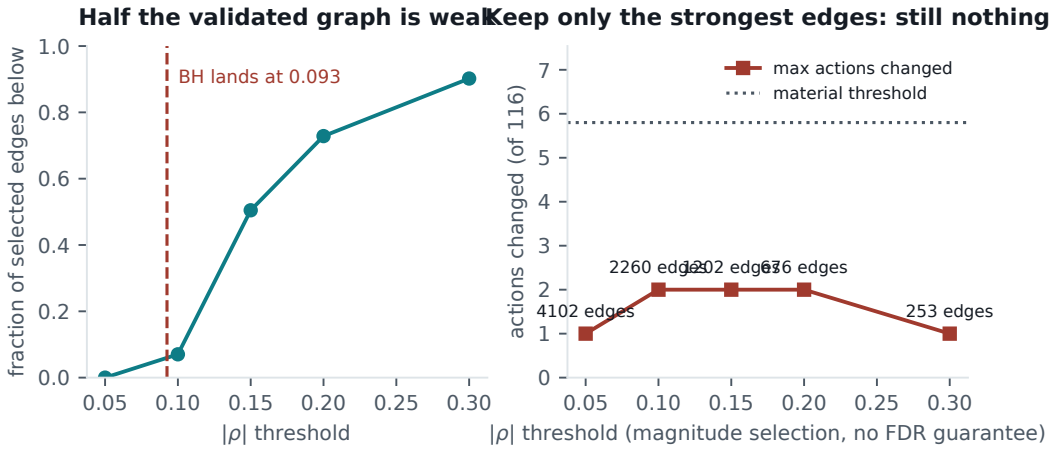


Figure 13: rebuilding the graph by correlation magnitude alone, at $|\rho| \geq \varepsilon$ (D4). Edge smallness is not why the coupling is inert: varying the one quantity the mechanism does not depend on leaves the result where it was.

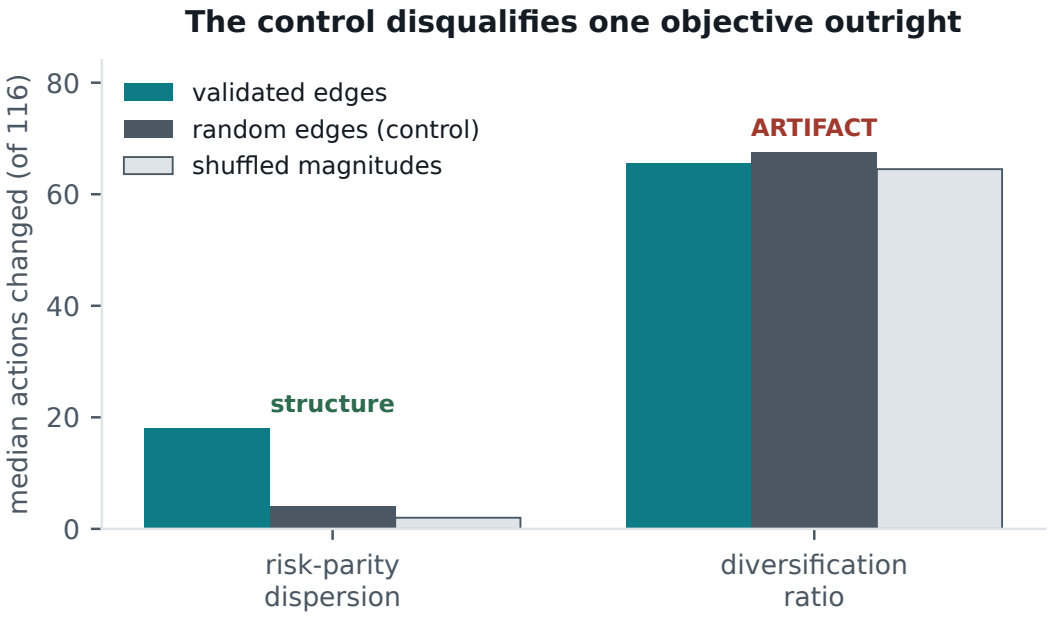


Figure 14: swapping the objective while holding the graph fixed, with both controls (§6.4). The diversification-ratio column is the control doing its job: without it that row would have read as the paper's strongest result.

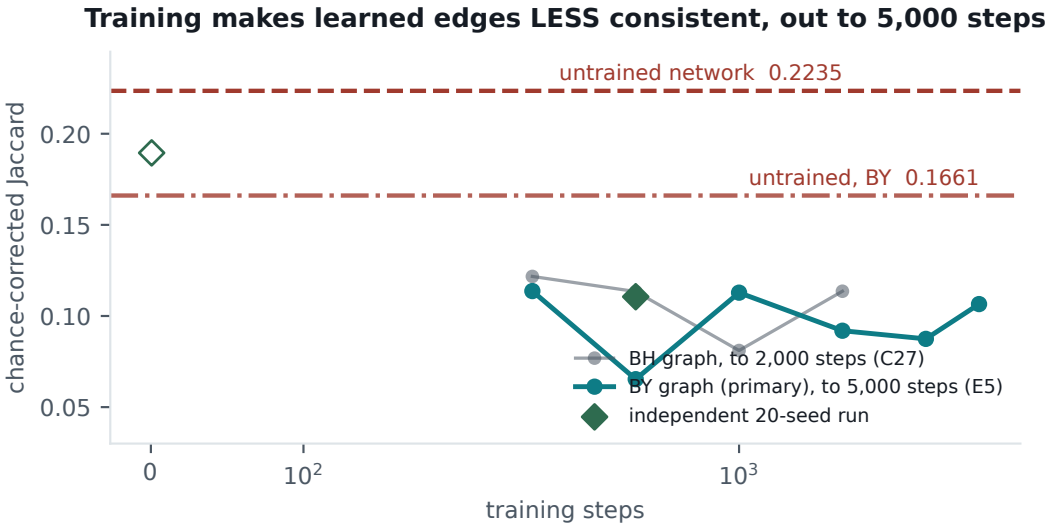


Figure 15: cross-seed agreement of learned edges against training budget (§6.5). Below the untrained baseline at every checkpoint out to 5,000 steps on the primary graph, so the reliability finding is not undertraining.

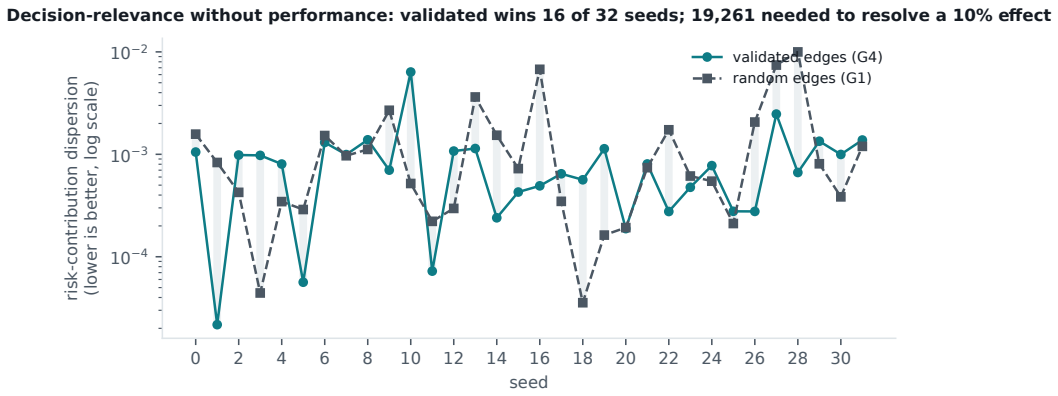


Figure 16: out-of-sample risk-contribution dispersion per seed, validated (G4) against count-matched random (G1) edges, over the 32 seeds the seed-count estimate is drawn from (E19b); each grey bar is one seed's paired gap. The arms interleave (validated wins some seeds and loses others), and the spread *across* seeds dwarfs the gap *within* any one of them, which is why resolving a 10% effect needs 19,261 $[12,380,34,044]$ seeds against a cap of 100. The estimand is a ratio one seed moves $75\times$ at eight seeds and $2.0\times$ at thirty-two; a scale-free one needs 4,356, and E26 puts 74% of the variance in the training seed, where no evaluation design reaches it.

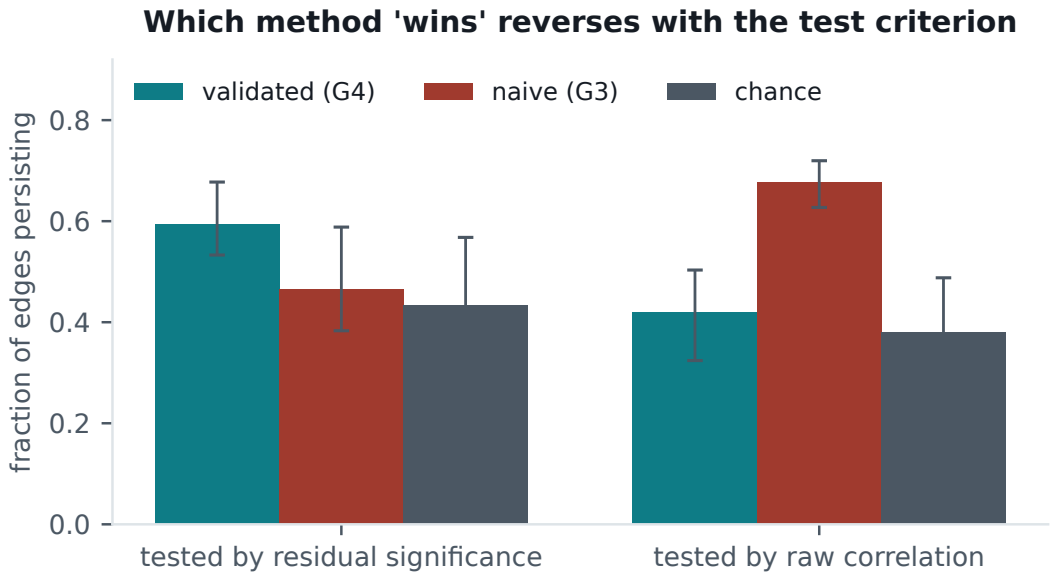


Figure 17: persistence is criterion-dependent (C24): tested by residual significance the validated graph persists more, tested by raw correlation the naive graph does; each wins under the criterion its own pipeline optimises. On two criteria neither pipeline optimises (E2), the validated graph wins both; see Appendix L.

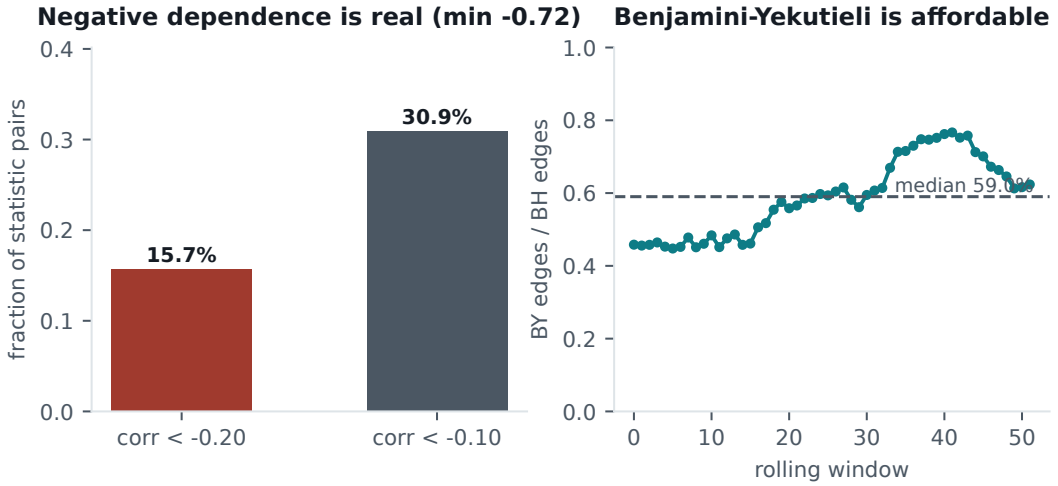


Figure 18: what the Benjamini–Yekutieli correction costs: edges retained per window relative to Benjamini–Hochberg (§4.1). The graph stays non-empty in all 52 windows.

O CITATION VERIFICATION STATUS

This appendix records how each cited work was verified, because two incidents of fabricated citation content in a related project made a search summary an unacceptable source. *read*: the PDF was read locally. *read-partial*: read, but part of the file was not machine-extractable. *secondary*: attributed via a source that was read; the primary was not opened. *unread*: cited as a pointer only; **no claim in this paper rests on it**. *unread, bibliography verified*: as *unread*, and the author, title, venue, volume, pages and DOI were additionally checked against the publisher's own record rather than a search result.

Reference	Status	Load-bearing for
Wang et al. (2021)	read	the motivating gap
Zhang (2026)	read	related-work placement
Soleymani & Paquet (2021)	read	related-work placement
Sood et al. (2023)	read	reward-mismatch critique
Balaji et al. (2020)	read	collision check on factored RL
Kim et al. (2019)	read	related-work placement
Uematsu & Yamagata (2024)	read	FDR framing, kurtosis gate
Curme et al. (2015)	read	validated-network design
Bennett et al. (2022)	read	ruled out as a validation method
Guestrin et al. (2001)	read	exact-elimination cost
Böhmer et al. (2020)	read	max-plus, parameter sharing
Qu et al. (2020)	read	locality claim, scope-checked
van Seijen et al. (2017)	read	reward-decomposition collision check
Fan et al. (2013)	read	Assumption A1
Benjamini & Yekutieli (2001)	read	Assumption A3, the BY level
Maillard et al. (2010)	read-partial	risk-parity objective
Benjamini & Hochberg (1995)	secondary	the BH procedure
Moody et al. (1998)	secondary	Differential Sharpe
Chamberlain & Rothschild (1983)	primary (NBER WP 996 version)	approximate factor models
Ledoit & Wolf (2004)	secondary	shrinkage baseline
Markowitz (1952)	secondary	reward (1)
Adebayo et al. (2018)	verified in a related project	the randomization control
Spirtes et al. (2001)	secondary	the PC algorithm
Bickel & Levina (2008)	secondary	thresholding background
Choueifaty & Coignard (2008)	secondary	diversification ratio
Dempster (1972)	**unread**, bib. verified	nothing; names the estimator
Ho et al. (2022)	read (2018 manuscript)	the Basel-correlation point