Negative edge correlation across a three-terminal forest interface

Scott Schubr
Luminary Oddities

November 2026

Abstract

Start with an arbitrary weighted graph and three distinguished vertices. Add independent sites attached only to those vertices and allow additional edges among the three vertices. In the weighted unrooted forest, any two added edges are nonpositively correlated. For positive activities the correlation is strictly negative exactly when the positive-support graph has a cycle containing both edges. The proof uses a five-state partition at the interface. Stability bounds its signed deviation from a triangle state, and the relevant Rayleigh difference is minimized at a realizable star or triangle state. A nonnegative deviation admits an exact triangle-and-star replacement for all attached laws.

1 Statement

For a finite nonnegatively weighted graph G, write

$$Z_G = \sum_{F \text{ forest of } G} \prod_{e \in F} w_e.$$

The empty forest and isolated vertices are included. In the probability measure proportional to these weights, the covariance of distinct edge indicators is $-w_e w_f \Delta_G(e,f)/Z_G^2$, where $\Delta_G(e,f) = (Z_G)_e (Z_G)_f - Z_G (Z_G)_{ef}$. Derivatives are with respect to the labeled activities. The general weighted Rayleigh question for unrooted forests extends the negative-correlation question of Kahn [3]. Exact results for small graphs appear in [4]. High common activity and connectivity covariances are discussed in [5]. Complete bipartite graphs at common activity are treated in [2].

Theorem 1. *Let H be any finite nonnegatively weighted graph with three distinguished vertices $T = \{1,2,3\}$. Add any number of independent sites, each with edges only to T, and any labeled edges among T. For any two distinct added edges their indicators in the unrooted forest are nonpositively correlated. When both activities are positive they are strictly negatively correlated exactly when their two edges belong to a common cycle of the positive-support graph.*

An exhaustive enumeration of forests on random small instances (4241 pairs of added edges, with backgrounds of up to two interior vertices and up to two sites) confirms both the sign and the exact strictness criterion.

The vertex-cover-three argument in [1, Section 7.1] applies when the background consists only of independent sites and terminal edges. It uses the same optional component marks and two Rayleigh-difference identities as below. Theorem 1 allows all edges in the arbitrary background H, and places no factorization constraint on the edges from the added sites to the terminals. It concerns the indicated added edges.

For parallel labels, combine each parallel class by summing its weights. Conditionally on selecting a class, choose one label with probabilities proportional to its weight, independently across classes. This rescales covariance between different classes by a positive factor for positive labels. A pair in the same class cannot both occur and has strict negative covariance when both weights are positive. It forms a cycle of length two. Loops can be discarded, and zero-weight edges are absent. We therefore give the proof for simple graphs. Incident pairs follow directly from [1, Theorems 2.2 and 3.1].

2 The signed terminal state

Let a background B meet the added graph only along T. Its forests induce the five terminal partitions $1|2|3, 12|3, 13|2, 23|1, 123$. Divide their activities by the total $w_0 > 0$ of the first state and write the normalized weights as $1, h, j, k, C$. Put $E = hj + hk + jk$ and $d = C - E$. After choosing a forest of B, whether an added edge set remains acyclic depends only on the terminal partition: contract the background trees and test the remaining edges. Therefore the normalized partition function for any attached graph is a linear function of the five states. Its Rayleigh difference is $\Delta_G(e, f)/w_0^2$.

Lemma 1. *For any such background,*

$$d^2 + dE \le hjk, \qquad d \le d_{\max} := \frac{\sqrt{E^2 + 4hjk} - E}{2}.$$

If $d > 0$ then $h, j, k > 0$ and $d \le jk/(j + k)$, with the corresponding bounds by symmetry.

Proof. The optional component-mark polynomial of [1, Theorem 2.1] is real stable for every weighted background. Give marks only to the three terminals. Its normalized form is

$$\begin{aligned} P(z) = {} & (1 + z_1)(1 + z_2)(1 + z_3) \\ & + h(1 + z_1 + z_2)(1 + z_3) + j(1 + z_1 + z_3)(1 + z_2) \\ & + k(1 + z_2 + z_3)(1 + z_1) + C(1 + z_1 + z_2 + z_3). \end{aligned}$$

Setting $z_1 = z_2 = 0$ and $z_3 = t$, the Rayleigh difference of marks 1 and 2 is nonnegative for every real t. With $Z = 1 + h + j + k + C$ it equals

$$[Z + (1 + h + k)t][Z + (1 + h + j)t] - Z(1 + t)(1 + j + k + t).$$

Its discriminant is $4Z(d^2 + dE - hjk)$, by expansion. Nonnegativity forces this to be nonpositive, also when the quadratic term vanishes. The upper bound by $d_{\max}$ follows from the roots. If $d > 0$, use $d^2 \le hjk$ and $dh(j + k) \le dE \le hjk$. □

For $h, j, k > 0$, the state $(1, h, j, k, E + d_{\max})$ is realized by a single site with activities

$$\zeta_1 = hj/d_{\max}, \quad \zeta_2 = hk/d_{\max}, \quad \zeta_3 = jk/d_{\max}.$$

Indeed its five forest weights are $1 + \sum \zeta_i, \zeta_1\zeta_2, \zeta_1\zeta_3, \zeta_2\zeta_3, \zeta_1\zeta_2\zeta_3$. The defining quadratic gives $1 + \sum \zeta_i = hjk/d_{\max}^2$, verifying the stated normalization. If one pair weight vanishes, then $d_{\max} = 0$ and the triangle with weights h, j, k realizes the state, with zero edges omitted.

3 Disjoint pairs

For disjoint targets move all edges into the background except the targets and the other edges at the sites touched by them. This leaves two pair types.

3.1 A terminal edge and a site edge

Take $e = 12$, $f = 3x$. Let x_1, x_2 be the other weights at x, possibly zero. The five-state deletion-contraction identity of [1, Section 7.1, case (a)] gives the normalized Rayleigh difference

$$(jx_2 - kx_1)^2 + jx_2(1 + j + x_2) + kx_1(1 + k + x_1) \\ - d(x_1 + x_2)(1 + x_1 + x_2).$$

If $d \le 0$ the displayed terms are nonnegative. For $d > 0$ use $d \le c = jk/(j + k) \le \min(j, k)$ and

$$c(x_1 + x_2) \le jx_2 + kx_1, \qquad c(x_1 + x_2)^2 \le kx_1^2 + jx_2^2.$$

The second inequality is weighted Cauchy–Schwarz. The remaining terms include $j^2x_2 + k^2x_1$. Hence the difference is positive whenever $jx_2 + kx_1 > 0$. A common cycle exists exactly when one of these products is positive: if the cycle leaves x along $x1$, it uses a background path from 2 to 3 avoiding 1, and conversely such a path and the three added edges form a cycle. This path exists exactly when $k > 0$. The case of leaving along $x2$ uses $j > 0$.

3.2 Edges at distinct sites

Take $e = 1x$, $f = 2y$, $x \ne y$. The other weights are a_2, a_3 at x and b_1, b_3 at y, with subscripts indicating the terminal. Set

$$A = 1 + a_2 + a_3, \quad B' = 1 + b_1 + b_3, \quad \alpha = a_2a_3/A, \quad \beta = b_1b_3/B'.$$

If e is absent, deleting x gives a factor A and an effective edge of weight α on 23. If e is contracted, the remaining edges become edges from 1 to 2 and 3; the same applies at y. Adding terminal edges increases the corresponding pair weights and leaves d unchanged. With $\phi(H, J, K) = 1 + H + J + K + HJ + HK + JK$, this gives

$$\phi_1 = \phi(h + a_2, j + a_3 + \beta, k),$$
$$\phi_2 = \phi(h + b_1, j, k + b_3 + \alpha),$$
$$\phi_3 = \phi(h + a_2 + b_1, j + a_3, k + b_3),$$
$$\phi_4 = \phi(h, j + \beta, k + \alpha).$$

The normalized Rayleigh difference is

$$AB'[(\phi_1 + d)(\phi_2 + d) - (\phi_3 + d)(\phi_4 + d)].$$

These are the state identities of [1, Section 7.1, case (b)]. Expansion shows this expression is affine in d, with slope $-AB'M$, where

$$M = a_2(b_3 - \beta) + b_1(a_3 - \alpha) + a_3b_3 + \alpha\beta \ge 0.$$

Thus it suffices to evaluate the maximal state of Lemma 1, even if the actual d is negative. The maximal state is realized on the three terminals with one additional site or triangle edges. Adding x, y gives a simple graph on at most six vertices. Erickson's all-weights finite theorem [4, Theorem 5.3.1 and Corollary 5.3.2] gives the desired nonnegativity after removing zero weights.

3.3 Strictness for distinct sites

The finite theorem of [4] gives more than the sign. Its equation (5.2.1) expresses each Rayleigh difference on a simple graph of at most seven vertices as sums of activity monomials times squares. Inside the summand indexed by a subset S of a common cycle minus $\{e, f\}$, the sign of a term depends only on S and the unique common cycle in that term. If a common cycle exists, take one of maximal length and let S be all its edges except e, f. No different common cycle can contain $S \cup \{e, f\}$. All signs in this summand therefore agree, and the choice of the forest S occurs. The summand is strictly positive when all its cycle weights are positive. This proves strictness in the finite comparison graph when it retains a common cycle.

If $1x$ and $2y$ belong to a common cycle in the original graph, deleting x, y from the cycle leaves two disjoint paths through terminals in the background, allowing a path of length zero. At most one path is nontrivial because there are only three terminals. If a nontrivial path connects a terminal pair, it avoids the third, and its terminal pair weight is positive. Replace it by the corresponding path through the maximal-state star, or by the positive triangle edge when $d_{\max} = 0$. This preserves the common cycle in the finite comparison graph. The normalized Rayleigh difference is then strictly positive there, and monotonicity in d gives strict positivity in the original graph. If no cycle contains both edges, they belong to different cycle-matroid components and the forest law factors across those components. Together with the terminal–site and incident cases, this completes Theorem 1.

4 Exact replacement and the signed terminal state

If an arbitrary background has $d \geq 0$, its full terminal state is reproduced by a triangle plus one site. For $d = 0$ use the triangle alone. Otherwise all h, j, k are positive. Take the maximal-state star for pair weights (th, tj, tk), $0 < t \leq 1$. Its quadratic root d_t is continuous with $d_t \to 0$ as $t \downarrow 0$ and $d_1 = d_{\max}$. Choose t with $d_t = d$. Add triangle weights $(1-t)h, (1-t)j, (1-t)k$, which add to the pair weights and preserve d_t. The five-state composition therefore gives exact equality for every normalized attached partition function and joint law of attached edge indicators. The Rayleigh theorem used the upper constraint on d, so it also covers possible negative defects.

Nonnegative defect is preserved under terminal extension. Join a new terminal $1'$ to 1 by an edge of weight $w > 0$ and regard $1', 2, 3$ as the terminals. In the new normalization $D = 1 + h + j + w$, the pair weights are $wh/D, wj/D, [(1+w)k + C]/D$, and the connected weight is wC/D. Indeed if the new edge is present it contributes w times the old state, and if it is absent only the states $1'|2|3$ and $23|1'$ occur. Consequently the new defect is

$$d' = \frac{w[(1+w)d + hj]}{D^2}.$$

This extends exact replacement to arbitrary terminals in any weighted cactus. A subgraph attached through two vertices can be replaced by an edge with weight equal to its connected-to-separated forest weight ratio. A subgraph attached through one vertex factors out. Apply these rules on the terminal subtree of the block-cut tree, regarding nonarticulation terminals as leaves at their blocks. If its junction is a vertex, its arms reduce to a star or a path. If its junction is a cycle block, replace the arcs between the attachment vertices by triangle edges, then apply terminal extensions along the arms. A bridge block has at most two attachments. When the terminals lie in different components at most one pair weight is positive and $d = 0$. Thus the replacement holds on all these backgrounds.

The sign for a general three-terminal state can be written as a covariance of effective conductances. For a forest of B, let R be the interior vertices in components that meet T, and $S = V(B) \setminus (T \cup R)$. For fixed R the partition sum factors into an unrooted forest on S of weight sum $F(S)$ and a grove on $T \cup R$, where every grove component meets T. Let D_R be the determinant of the interior block of the induced Laplacian on $T \cup R$. Terms with $D_R = 0$ have no such grove. By the rooted-forest determinant formula [6], the marked grove polynomial on T equals D_R times the marked determinant for the Schur complement onto T. For three terminals that complement is a triangle Laplacian with conductances $c_{12}(R), c_{13}(R), c_{23}(R)$. Reading the homogeneous mark coefficients identifies the five conditional grove weights as D_R times $1, c_{12}, c_{13}, c_{23}, \sum_{e<f} c_e c_f$. In degree two the three pair-partition mark polynomials are linearly independent. Consequently, under probability proportional to $F(S)D_R$, the signed state coordinate is

$$d = \sum_{e<f} \mathrm{Cov}(c_e(R), c_f(R)).$$

This expresses the signed state coordinate, which governs exact triangle-and-star replacement, as a covariance of effective conductances. The theorem on selected added edges holds for every sign of d.

Acknowledgement

OpenAI Codex and Anthropic Claude assisted with proof development, independent checking, numerical verification and manuscript preparation. The author is responsible for the work.

References

[1] S. Schubr. Stable stars, rank-one replacement and negative edge correlation in random forests. Archive edition 1 PDF, 2026.

[2] S. Schubr. Negative edge correlation in complete bipartite forests at every fugacity. Archive edition 1 PDF, 2026.

[3] G. R. Grimmett and S. N. Winkler. Negative association in uniform forests and connected graphs. *Random Structures and Algorithms*, 24:444–460, 2004.

[4] A. Erickson. Negative correlation properties for matroids. Master's thesis, University of Waterloo, 2008. University repository.

[5] R. A. Çiçeksiz and M. Ravichandran. Negative correlations for forests and the $q < 1$ random cluster model. arXiv:2610.00549.

[6] S. Chaiken. A combinatorial proof of the all minors matrix tree theorem. *SIAM Journal on Algebraic and Discrete Methods*, 3:319–329, 1982.