Supplementary Materials: Exact Permutation and Bootstrap Distribution of Generalized Pairwise Comparisons Statistics

William N. Anderson [1] and Johan Verbeeck [2,*]

S1. Introduction

In a general situation the GPC evaluation will result in a skew matrix U, which has in part the following interpretation

- $u_{ij} < 0$ if subject i has a less favorable outcome than subject j.
- $u_{ij} = 0$ if the subjects cannot be compared, or are actually tied.
- $u_{ij} > 0$ if subject i has a more favorable outcome than subject j.

We are interested in a trial with two arms, treatment and control. We let N be the number of subjects in the trial, with m subjects in the treatment group and n subjects in the control group. Moreover, let the indicator $D_i = 1$ for subjects in the treatment group, and $D_i = 0$ for patients in the control group.

The matrix entry u_{ij} is defined as a *treatment win* if $D_i = 1$, $D_j = 0$, and $u_{ij} > 0$. Similarly the matrix entry u_{ij} is defined as a *control win* if $D_i = 0$, $D_j = 1$, and $u_{ij} > 0$. Accordingly each win results from the comparison of a treatment subject and a control subject.

The *win sum* for subject i is defined by

$$W_i = \sum_{j=1}^{N} u_{ij} \qquad \text{where } u_{ij} > 0, D_i \neq D_j.$$

The win sums for the treatment and control groups are then defined by

$$W_T = \sum_{i=1}^{N} W_i D_i \tag{S1}$$

$$W_C = \sum_{i=1}^{N} W_i (1 - D_i) \tag{S2}$$

Where necessary for clarity, we will let the symbols W_T^{obs} and W_C^{obs} denote the win sums from the original observed data, as opposed to the sums from another win matrix that appears in the course of the algorithm.

We wish to determine the mean and variance of the pair (W_T, W_C) resulting from the permutation distribution of the trial arms, or from bootstrapping from the patients. These means and variances are computed over all permutations of the trial arms, or all bootstrap samples. These values are accordingly the expected means and variances when randomized permutation or bootstrap sampling is performed. The comparisons within trial arms do not correspond to wins, and hence do not enter into the computations of (W_T, W_C). However, the corresponding matrix entries will be needed for some of the variance computations below.

It should be noted that if the entries of U are restricted to $\{-1, 0, 1\}$, the values (W_T, W_C) merely count the wins.

S2. Graphical model

The U matrix can also be viewed as the adjacency matrix of a directed graph $\mathbb{G}$, with N vertices. If $u_{ij} > 0$, the graph has an *edge* from vertex i to vertex j. For visualization, we can draw this edge as an arrow with *head* at vertex j and *tail* at vertex i. The value u_{ij} is called the *weight* of the edge, and is denoted by w_e when referring to an edge e. The graph $\mathbb{G}$ will have E edges; in terms of the original matrix U, E is the number of positive u_{ij}. For notational purposes we number the edges $1, 2, \ldots E$, without prescribing any relationship between the edge numbers and the row and column numbers in the matrix U. If a vertex v is the head or tail of an edge e, then e is said to be *adjacent* to v. We do not use any actual theorems from graph theory, but the graphical model may aid in understanding the algorithms to follow.

The *indegree* $\mathrm{ID}(v)$ of a vertex v, is the number of edges whose head is the vertex v. This is equivalent to the number of positive entries in the vertex or subject associated column of the score matrix U, or the number of pairs that represent a loss for that vertex or subject. Similarly, the *outdegree* $\mathrm{OD}(v)$ is the number of edges whose tail is the vertex v, which is the number of positive entries in the associated row of the score matrix U, or the number of pairs that represent a win for that subject. These quantities do not depend on the trial arm assignments.

In addition to these general graph theory concepts, we need some notations that are specific to our situation.

- In the observed data a vertex v is called a *treatment vertex* if the corresponding $D_v = 1$, and a *control vertex* otherwise.
- In the observed data an edge corresponds to a treatment win if the corresponding $u_{ij} > 0, D_i = 1,$ and $D_j = 0$; such an edge is called a *treatment edge*. Similarly, if $u_{ij} > 0, D_i = 0,$ and $D_j = 1$ the edge is a *control edge*. If $u_{ij} > 0$, but $D_i = D_j$, the edge is a *neutral edge*.
- The following quantities are defined for every sample, including, of course, the observed data.
  - For a vertex v, let F_v denote the number of times that the vertex v appears in the sample. For the permutation distribution, $F_v = 1$ for all samples. For the bootstrap distributions, F_v can be zero or higher.
  - The value T_e is defined as the number of times that edge e is a treatment edge in the sample, and the value C_e is the number of times that edge e is a control edge in the sample. If the treatment edge e has vertices (v, w), then $T_e = F_v F_w$.
  - The vectors $\mathbf{T}$ and $\mathbf{C}$ are the $E \times 1$ column vectors composed of the various T_e and C_e. We also let $\mathbf{W}$ be the column vector of the weights w_e.
    In this context, we rewrite (S1) and (S2) as

$$W_T = \mathbf{W}^t\mathbf{T} = \sum_e w_e T_e = [w_1, w_2, \ldots, w_E]\mathbf{T}$$
$$W_C = \mathbf{W}^t\mathbf{C} = \sum_e w_e C_e = [w_1, w_2, \ldots, w_E]\mathbf{C}$$

S3. General Variance Computations

In the next sections, the expectations and variances of W_T and W_C, and their covariance will be developed for the permutation (Section S4), two-sample bootstrap (Section S5) and one-sample bootstrap (Section S6) in detail. They follow however, the same general pattern.

The expectation of W_T is derived by:

$$\begin{aligned}
\mathbb{E}(W_T) &= \mathbb{E}(\mathbf{W}^t\mathbf{T}) \\
&= [w_1, w_2, \ldots, w_E]\mathbb{E}(\mathbf{T}) \\
&= \sum_e w_e \mathbb{E}(T_e).
\end{aligned}$$

We will show below that $\mathbb{E}(T_e)$ is the same for all edges.

Similarly, the expectation of W_C is derived by:

$$\begin{aligned}\mathbb{E}(W_C) &= \mathbb{E}(\mathbf{W}^t\mathbf{C})\\ &= [w_1, w_2, \ldots, w_E]\mathbb{E}(\mathbf{C})\\ &= \sum_e w_e \mathbb{E}(C_e).\end{aligned}$$

The variance is derived by

$$\begin{aligned}\mathrm{Var}([W_T, W_C]) &= \mathrm{Var}\left(\left[\mathbf{W}^t, \mathbf{W}^t\right]\begin{bmatrix}\mathbf{T}\\ \mathbf{C}\end{bmatrix}\right)\\ &= \left[\mathbf{W}^t, \mathbf{W}^t\right]\mathrm{Var}\left(\begin{bmatrix}\mathbf{T}\\ \mathbf{C}\end{bmatrix}\right)\begin{bmatrix}\mathbf{W}\\ \mathbf{W}\end{bmatrix}\\ &= \left[\mathbf{W}^t, \mathbf{W}^t\right]\left(\mathbb{E}\left(\begin{bmatrix}\mathbf{T}\\ \mathbf{C}\end{bmatrix}\begin{bmatrix}\mathbf{T}\\ \mathbf{C}\end{bmatrix}^t\right) - \left(\mathbb{E}\begin{bmatrix}\mathbf{T}\\ \mathbf{C}\end{bmatrix}\right)\left(\mathbb{E}\begin{bmatrix}\mathbf{T}\\ \mathbf{C}\end{bmatrix}^t\right)\right)\begin{bmatrix}\mathbf{W}\\ \mathbf{W}\end{bmatrix}\\ &= \left[\mathbf{W}^t, \mathbf{W}^t\right]\left(\mathbb{E}\left(\begin{bmatrix}\mathbf{T}\mathbf{T}^t & \mathbf{T}\mathbf{C}^t\\ \mathbf{C}\mathbf{T}^t & \mathbf{C}\mathbf{C}^t\end{bmatrix}\right) - \begin{bmatrix}\mathbb{E}(\mathbf{T})\mathbb{E}(\mathbf{T}^t) & \mathbb{E}(\mathbf{T})\mathbb{E}(\mathbf{C}^t)\\ \mathbb{E}(\mathbf{C})\mathbb{E}(\mathbf{T}^t) & \mathbb{E}(\mathbf{C})\mathbb{E}(\mathbf{C}^t)\end{bmatrix}\right)\begin{bmatrix}\mathbf{W}\\ \mathbf{W}\end{bmatrix} && \text{(S3)}\\ &= \begin{bmatrix}V_{TT} & V_{TC}\\ V_{CT} & V_{CC}\end{bmatrix} && \text{(S4)}\end{aligned}$$

Computing the variance of the GPC statistics requires two counting steps. In the first step, we compute the expected values $\mathbb{E}(T_e)$, $\mathbb{E}(C_e)$ and the expected value of an ordered pair of edges (e, f), not necessarily distinct, $\mathbb{E}(T_eT_f)$, $\mathbb{E}(T_eC_f)$, $\mathbb{E}(C_eT_f)$, and $\mathbb{E}(C_eC_f)$. This computation involves elementary calculations involving binomial coefficients. The calculations differ between edge pairs, depending on the trial arm assignments and the geometric relationship of the edges. We note that because the variance matrix is symmetric, computing the individual terms $\mathbb{E}(C_eT_f)$ is redundant, but we will present those calculations anyway. In the second step, we compute the number of times that each of these geometric configurations of edges is present in the dataset.

S4. The permutation distribution

In a permutation test, subjects are randomly re-sampled to the treatment groups without replacement. If all possible permutations of m treatment assignments and n control assignments are considered, the W_T and W_C in each of these permutation samples will lead to their permutation distribution. The expectations, variances, and covariance of this permutation distribution of W_T and W_C can be calculated explicitly. An edge will always join the same subjects, but whether or not this edge contributes to the treatment wins or control wins depends on the treatment assignment in the permutation sample.

For expected values, we use the symbol $\mathbb{E}_P$, as a specification of the general expection symbol $\mathbb{E}$.

S4.1. Expectations

For a single edge $e = (i, j)$ to contribute to the treatment wins, the treatment assignments need to be $D_i = 1$ and $D_j = 0$. Since, the total number of trial arm assignments is $\binom{N}{m}$, for a single edge $e = (i, j)$, the number of assignments with $D_i = 1$ and $D_j = 0$ is $\binom{N-2}{m-1}$. Hence for all edges e,

$$\mathbb{E}_P(T_e) = \binom{N-2}{m-1} \Big/ \binom{N}{m} = \frac{mn}{N(N-1)}, \tag{S5}$$

and

$$\mathbb{E}_P(W_T) = \mathbb{E}_P(\sum_e w_e T_e) = \frac{mn}{N(N-1)} \sum_e w_e. \tag{S6}$$

Since having $D_i = 1$ and $D_j = 0$ is equally likely to having $D_i = 0$ and $D_j = 1$, it follows that $\mathbb{E}_P(C_e) = \mathbb{E}_P(T_e)$, and $\mathbb{E}_P(W_C) = \mathbb{E}_P(W_T)$.

S4.2. Variances and covariance

For the variance computation using (S4), we will additionally calculate the terms $\mathbb{E}_P(T_e T_f)$, $\mathbb{E}_P(T_e C_f)$, $\mathbb{E}_P(C_e T_f)$, and $\mathbb{E}_P(C_e C_f)$ for all ordered pairs of edges e, f. There are various cases depending on the geometric relationship between the edges in the pair. For pairs of edges that meet at a vertex there are separate formulas for pairs of identical edges (case 1), for pairs that have a common head (case 2), a common tail (case 3), or one head and one tail at the vertex (cases 4 and 5). For all pairs of edges that do not meet at a vertex (case 6), there is one formula.

In principle some of the above cases could be combined, by considering a mixture of ordered and unordered pairs. It seems clearer to use ordered pairs throughout.

Computations at a single vertex .

Case 1: An ordered pair of identical edges e with head v:

- In the permutation case, T_e is 0 or 1, we have using (S5)

$$\mathbb{E}_P(T_e^2) = \mathbb{E}_P(T_e) = \frac{mn}{N(N-1)}.$$

- Similarly,

$$\mathbb{E}_P(C_e^2) = \mathbb{E}_P(C_e) = \frac{mn}{N(N-1)}.$$

- Since an edge cannot be simultaneously a win for treatment and for control, we have $T_e C_e = 0$ for all edges e and all permutations, and thus

$$\mathbb{E}_P(T_e C_e) = 0.$$

Case 2: An ordered pair (e, f) of distinct edges, each with head v:

- We have an ordered pair of edges $e = (v_i, v_j)$ and $f = (v_k, v_j)$ where $\{v_i, v_j, v_k\}$ are distinct; that is vertex v_j is the head for both edges. Then $T_e T_f = 1$ iff $D_i = D_k = 1$ and $D_j = 0$. There remain $N-3$ patients to be assigned trial arms, with $m-2$ treatment patients and $n-1$ control patients. The number of trial arm permutations satisfying these conditions is $\binom{N-3}{m-2}$. Accordingly.

$$\mathbb{E}_P(T_e T_f) = \binom{N-3}{m-2} \Big/ \binom{N}{m} = \frac{mn(m-1)}{N(N-1)(N-2)}.$$

- Similarly,

$$\mathbb{E}_P(C_e C_f) = \binom{N-3}{n-2} \Big/ \binom{N}{m} = \frac{mn(n-1)}{N(N-1)(N-2)}.$$

- T_e and C_f cannot both be 1, because the former requires $D_j = 0$, and the latter requires $D_j = 1$. Hence $T_e C_f = 0$ always. Accordingly

$$\mathbb{E}_P(T_e C_f) = \mathbb{E}_P(C_e T_f) = 0.$$

Case 3: An ordered pair (e, f) of distinct edges, each with tail v:

- We have an ordered pair of edges $e = (v_i, v_j)$ and $f = (v_i, v_l)$ where $\{v_i, v_j, v_l\}$ are distinct; that is vertex v_i is the tail for both edges. Then $T_e T_f = 1$ iff $D_i = 1$ and $D_j = D_l = 0$. The number of trial arm permutations satisfying these conditions is $\binom{N-3}{m-1}$. Accordingly

$$\mathbb{E}_P(T_e T_f) = \binom{N-3}{m-1} \Big/ \binom{N}{m} = \frac{mn(n-1)}{N(N-1)(N-2)}.$$

  Similarly,

$$\mathbb{E}_P(C_e C_f) = \binom{N-3}{n-1} \Big/ \binom{N}{m} = \frac{mn(m-1)}{N(N-1)(N-2)}.$$

- T_e and C_f cannot both be 1, because the former requires $D_i = 1$, and the latter requires $D_i = 0$. Hence $T_e C_f = 0$ always. Accordingly

$$\mathbb{E}_P(T_e C_f) = \mathbb{E}_P(C_e T_f) = 0.$$

Case 4: An ordered pair (e, f) of distinct edges, where vertex v is the tail of edge e and the head of edge f:

- We have a pair of edges $e = (v_i, v_j)$ and $f = (v_k, v_i)$ where $\{v_i, v_j, v_l\}$ are distinct; that is vertex v_i is the tail of edge e and the head of edge f. Then T_e and T_f can never both be 1, because the former requires $D_i = 1$ and the latter requires $D_i = 0$. Accordingly

$$\mathbb{E}_P(T_e T_f) = \mathbb{E}_P(C_e C_f) = 0.$$

- $T_e C_f = 1$ iff $D_i = 1$ and $D_j = D_k = 0$. Accordingly

$$\mathbb{E}_P(T_e C_f) = \binom{N-3}{m-1} \Big/ \binom{N}{m} = \frac{mn(n-1)}{N(N-1)(N-2)}.$$

- $C_e T_f = 1$ iff $D_i = 0$ and $D_j = D_k = 1$. Accordingly

$$\mathbb{E}_P(C_e T_f) = \binom{N-3}{n-1} \Big/ \binom{N}{m} = \frac{mn(m-1)}{N(N-1)(N-2)}.$$

Case 5: An ordered pair (e, f) of distinct edges, where vertex v is the head of edge e and the tail of edge f:

- We have a pair of edges $e = (v_i, v_j)$ and $f = (v_j, v_k)$ where $\{v_i, v_j, v_k\}$ are distinct; that is vertex v_j is the head of edge e and the tail of edge f. Then T_e and T_f can never both be 1, nor can C_e and C_f. Accordingly

$$\mathbb{E}_P(T_e T_f) = \mathbb{E}_P(C_e C_f) = 0.$$

- $T_e C_f = 1$ iff $D_i = D_k = 1$ and $D_j = 0$. Accordingly

$$\mathbb{E}_P(T_e C_f) = \binom{N-3}{m-2} \Big/ \binom{N}{m} = \frac{mn(m-1)}{N(N-1)(N-2)}.$$

  Similarly,

- $$\mathbb{E}_P(C_e T_f) = \binom{N-3}{n-2} \Big/ \binom{N}{m} = \frac{mn(n-1)}{N(N-1)(N-2)}.$$

Computations for the remaining edge pairs.

Case 6: An ordered pair (e, f) of non-intersecting edges:

- Consider a pair of edges $e = (v_i, v_j)$ and $f = (v_k, v_l)$ where $\{v_i, v_j, v_k, v_l\}$ are distinct; that is the edges do not meet. Then $T_e T_f = 1$ iff $D_i = D_k = 1$ and $D_j = D_l = 0$. The number of trial arm permutations satisfying these conditions is $\binom{N-4}{m-2}$. The same counting argument applies to $C_e C_f$, $T_e C_f$, and $C_e T_f$. Accordingly

$$\begin{aligned}
\mathbb{E}_P(T_e T_f) = \mathbb{E}_P(T_e C_f) = \mathbb{E}_P(C_e T_f) = \mathbb{E}_P(C_e C_f) \\
&= \binom{N-4}{m-2} \Big/ \binom{N}{m} \\
&= \frac{mn(m-1)(n-1)}{N(N-1)(N-2)(N-3)}
\end{aligned} \tag{S7}$$

The algorithm is now complete. We evaluate the variance (S4) by examining each ordered pair of edges (e, f), evaluating $\mathbb{E}_P(T_e T_f)$, $\mathbb{E}_P(T_e C_f)$, $\mathbb{E}_P(C_e T_f)$, and $\mathbb{E}_P(C_e C_f)$. This gives an $2E \times 2E$ matrix, where E is the number of edges in the graph $\mathbb{G}$. Pre- and post-multiply by the W terms. The variance matrix for the pair (W_T, W_C) is obtained by summing each of the four $E \times E$ blocks of this larger matrix.

Unfortunately, the algorithm as stated is $O(N^4)$, because the number of edges is $O(N^2)$, and we have explicitly computed all the entries of an $2E \times 2E$ matrix. For a practical algorithm we need to reduce the entire computation to $O(N^2)$.

S4.3. Practical algorithm in the permutation case

The key to a practical computation is to look at cases 1-5 at each vertex separately. Then sum those values to include in the four entries of (S4). Finally, compute the number of pairs in case 6, and include those terms also.

Define the following values of the U-matrix:

$$I_v = \sum_i w_e \quad \text{over edges } e \text{ such that the head of } e \text{ is vertex } v \text{ or the column sums of } w_e \text{ for vertex } v$$

$$I_v^s = \sum_i w_e^2 \quad \text{over edges } e \text{ such that the head of } e \text{ is vertex } v \text{ or the column sums of } w_e^2 \text{ for vertex } v$$

$$O_v = \sum_j w_e \quad \text{over edges } e \text{ such that the tail of } e \text{ is vertex } v \text{ or the row sums of } w_e \text{ for vertex } v.$$

$$O_v^s = \sum_j w_e^2 \quad \text{over edges } e \text{ such that the tail of } e \text{ is vertex } v \text{ or the row sums of } w_e^2 \text{ for vertex } v.$$

We note that if the entries of U are restricted to $\{-1, 0, 1\}$, then $I_v = I_v^s = \mathrm{ID}(v)$, and $O_v = O_v^s = \mathrm{OD}(v)$.

Case 1: At each vertex v, the contribution to the square terms in the variance is

$$\begin{aligned}
\frac{mn}{N(N-1)} I_v^s &\quad \text{for } V_{TT} \\
\frac{mn}{N(N-1)} O_v^s &\quad \text{for } V_{CC} \\
0 &\quad \text{for } V_{TC} \text{ and } V_{CT}
\end{aligned}$$

Case 2: For an ordered pair (e, f) of distinct edges each with head v, the contribution to the square terms in the variance V_{TT} is $w_e w_f \frac{mn(m-1)}{N(N-1)(N-2)}$. The total combination for all the ordered pairs of edges is then

$$\begin{aligned} \frac{mn(m-1)}{N(N-1)(N-2)}(I_v^2 - I_v^s) &\quad \text{for } V_{TT} \\ \frac{mn(n-1)}{N(N-1)(N-2)}(I_v^2 - I_v^s) &\quad \text{for } V_{CC} \\ 0 &\quad \text{for } V_{TC} \text{ and } V_{CT} \end{aligned}$$

The reason for the subtraction in $(I_v^s - I_v)$ is that we are looking at pairs of *distinct* edges; the ordered pairs of identical edges were already counted in case 1.

Case 3: For an ordered pair (e, f) of distinct edges each with tail v, the contribution to the square terms in the variance V_{TT} is $w_e w_f \frac{mn(n-1)}{N(N-1)(N-2)}$. The total combination for all the ordered pairs of edges is then

$$\begin{aligned} \frac{mn(n-1)}{N(N-1)(N-2)}(O_v^2 - O_v^s) &\quad \text{for } V_{TT} \\ \frac{mn(m-1)}{N(N-1)(N-2)}(O_v^2 - O_v^s) &\quad \text{for } V_{CC} \\ 0 &\quad \text{for } V_{TC} \text{ and } V_{CT} \end{aligned}$$

Case 4: At each vertex v, the contribution to the square terms in the variance is

$$\begin{aligned} \frac{mn(n-1)}{N(N-1)(N-2)} I_v O_v &\quad \text{for } V_{TC} \\ \frac{mn(m-1)}{N(N-1)(N-2)} O_v I_v &\quad \text{for } V_{CT} \\ 0 &\quad \text{for } V_{TT} \text{ and } V_{CC} \end{aligned}$$

Case 5: At each vertex v, the contribution to the square terms in the variance is

$$\begin{aligned} \frac{mn(m-1)}{N(N-1)(N-2)} O_v I_v &\quad \text{for } V_{TC} \\ \frac{mn(n-1)}{N(N-1)(N-2)} I_v O_v &\quad \text{for } V_{CT} \\ 0 &\quad \text{for } V_{TT} \text{ and } V_{CC} \end{aligned}$$

Case 6: For the non-intersecting edges, we need to sum the term (S7) for all pairs of edges, except those already considered in cases 1-5. Accordingly we define

$$F_P = \sum_v \left[I_v^s + (I_v^2 - I_v^s) + (O_v^2 - O_v^s) + 2 I_v O_v \right].$$

Then the variance term is

$$\frac{mn(m-1)(n-1)}{N(N-1)(N-2)(N-3)} \left(\left(\sum_e w_e \right)^2 - F_P \right) \quad \text{for } V_{TT}, V_{TC}, V_{CT}, \text{and } V_{CC}$$

S4.4. Complexity of the final permutation algorithm

For a specific vertex v, in order to compute the various vertex dependent terms in the variance formulas, $(I_v, I_v^s, O_v, \text{and } O_v^s)$ each other vertex must be examined once. Thus the complexity of the computation at each vertex is $O(N)$, and computing at all vertices is thus $O(N^2)$. The total number of edges and the wins W_T and W_C are computed from these numbers in an additional $O(N)$ steps, and the final computations are $O(1)$. Accordingly the time complexity of the permutation algorithm is $O(N^2)$. No algorithm can have time complexity less than $O(N^2)$, because we must examine each entry of the original $N \times N$ matrix at least once.

S5. The two-sample bootstrap distribution

In a two-sample bootstrap test, subjects are randomly re-sampled with replacement within their treatment group. There are accordingly m^m bootstrap samples from the treatment population, and n^n bootstrap samples from the control population, for a total of $m^m n^n$ bootstrap samples. If all possible bootstrap samples, are considered, the the treatment win sum W_T and control win sum W_C will have the complete bootstrap distribution. The means, variances, and covariance of this bootstrap distribution of W_T and W_C will be calculated explicitly. These values will also be the expected means, variances, and covariances from a randomized bootstrap sample.

Neutral edges in the observed data can never correspond to wins in a bootstrap sample, and such edges play no role in the computations below.

We will use a number of easily derived identities involving multinomial coefficients. These are summarized in section S7.

S5.1. Expectations

The expected values of the vectors **T** and **C** are computed using elementary sums of multinomial coefficients.

We note that the number of ways a treatment vertex can appear k_1 times in the sample from the treatment vertices is $\sum_{k_2+\ldots k_m=m-k_1} \binom{m}{k_1,k_2,\ldots,k_m}$. Accordingly the expected number of times that a treatment vertex appears in the complete set of bootstrap samples is given by

$$\begin{aligned}
\sum_1^m k_1 \sum_{k_2+\ldots k_m=m-k_1} & \binom{m}{k_1, k_2, \ldots, k_m} \\
&= \sum_{k_1+k_2+\ldots k_m=m} k_1 \binom{m}{k_1, k_2, \ldots, k_m} \\
&= m^m
\end{aligned}$$

where we have used identity (S14) from the appendix.

Since there are a total of m^m bootstrap samples from the treatment vertices, it follows from the above that the expected number of times that a specific treatment vertex appears in a bootstrap sample is 1. This particular point can easily be derived without using the multinomial coefficients, but sums of those coefficients appear to be necessary for some of the later computations.

Similarly, the number of ways a control vertex ℓ_1 appears in the sample from the control vertices is n^n.

Consider a treatment edge e. Then the head is a control vertex, and the tail is a treatment vertex. If the frequency of the treatment vertex is k_1, and frequency of the control vertex is ℓ_1 , then the edge e appears $k_1\ell_1$ times in the bootstrap sample. Accordingly, for the expected value of T_e, we have

$$\begin{aligned}\mathbb{E}_{B2}(T_e) &= \frac{1}{m^m n^n}\sum_{k_1+\ldots k_m=m} k_1\binom{m}{k_1,k_2,\ldots,k_m}\sum_{l_1+\ldots l_l=l} l_1\binom{n}{l_1,l_2,\ldots,l_m} \\ &= \frac{1}{m^m n^n} m^m n^n = 1. \end{aligned} \tag{S8}$$

For this edge we will have $\mathbb{E}_{B2}(C_e) = 0$, since this edge cannot correspond to a control win in any bootstrap sample. Similarly, if an edge f is a control edge, then $\mathbb{E}_{B2}(T_f) = 0$ and $\mathbb{E}_{B2}(C_f) = 1$. Consequently $\mathbb{E}_{B2}(W_T)$ is the observed treatment win sum, and $\mathbb{E}_{B2}(W_C)$ is the observed control win sum.

S5.2. Variances

For the variance, again the expectations of $\mathbb{E}_{B2}(T_eT_f)$, $\mathbb{E}_{B2}(C_eC_f)$, $\mathbb{E}_{B2}(T_eC_f)$ and $\mathbb{E}_{B2}(C_eT_f)$ need to be calculated, and these depend on the geometric configuration of the edges. We considered various cases of *ordered* pairs. Cases 3-10 fall into pairs, with the odd case for a treatment vertex and the even case for a control vertex. (The computations for the different vertices are slightly different.) The following cases are considered: an ordered pair of identical treatment edges (case 1), an ordered pair of identical control edges (case 2), an ordered pair of distinct treatment edges adjacent to a vertex (case 3 respectively case 4), an ordered pair of distinct control edges adjacent to a vertex (case 5 respectively case 6), an ordered pair of a treatment edge and a control edge adjacent to a vertex (case 7 respectively case 8), and an ordered pair of a control edge and a treatment edge adjacent to a vertex (case 9 respectively case 10). Finally, for non-intersecting edges, 4 cases are considered: an ordered pair of distinct treatment edges (case 11), an ordered pair of distinct control edges (case 12), an ordered pair of a treatment and a control edge (case 13), and an ordered pair of a control edge and a treatment edge (case 14).

In principle some of the above cases could be combined, by considering a mixture of ordered and unordered pairs of edges. It seems clearer to use ordered pairs throughout.

Computations at a single vertex.

Case 1: An ordered pair of identical treatment edges, adjacent to the vertex v:

- Let the edge be $e = (v_1, v_2)$, with corresponding vertex frequencies k_1, k_2. The number of possibilities for $T_e > 0$ is k_1k_2. Then

$$\begin{aligned}\mathbb{E}_{B2}(T_e^2) &= \frac{1}{m^m n^n}\sum_{k_1+\ldots k_m=m} k_1^2\binom{m}{k_1,k_2,\ldots,k_m} \\ &\quad\times\sum_{l_1+\ldots l_l=l} l_1^2\binom{n}{l_1,l_2,\ldots,l_m} \\ &= \frac{1}{m^m n^n}(2m-1)m^{m-1}(2n-1)n^{n-1} \\ &= \frac{(2m-1)(2n-1)}{mn}, \end{aligned} \tag{S9}$$

where we have used formula (S15) from section S7. For this same edge, $C_e = 0$ for all bootstrap samples, and hence

$$\mathbb{E}_{B2}(C_e^2) = \mathbb{E}_{B2}(T_eC_e) = \mathbb{E}_{B2}(C_eT_e) = 0.$$

Case 2: An ordered pair of identical control edges, adjacent to the vertex v:

- The computations are similar to case 1.

$$\mathbb{E}_{B2}(T_e^2) = \mathbb{E}_{B2}(T_eC_e) = \mathbb{E}_{B2}(C_eT_e) = 0.$$

$$\mathbb{E}_{B2}(C_e^2) = \frac{(2m-1)(2n-1)}{mn}.$$

Case 3: An ordered pair (e, f) of distinct treatment edges, each adjacent to the treatment vertex v:

- Let the edges be $e = (v_1, v_2)$ and $f = (v_1, v_3)$, with corresponding vertex frequencies k_1, k_2, k_3. Then the common vertex v_1 is a treatment vertex. The number of possibilities for $T_e > 0$ is k_1k_2, and for $T_f > 0$ is k_1k_3. Then the number of samples that produce both $T_e > 0$ and $T_f > 0$ is $k_1^2k_2k_3$, and

$$\begin{aligned}\mathbb{E}_{B2}(T_eT_f) &= \frac{1}{m^mn^n} \sum_{k_1+\dots k_m=m} k_1^2 \binom{m}{k_1, k_2, \dots, k_m} \\ &\quad\times \sum_{l_1+\dots l_n=n} l_2l_3 \binom{n}{l_1, l_2, \dots, l_n} \\ &= \frac{1}{m^mn^n}(2m-1)m^{m-1}(n-1)n^{n-1} \\ &= \frac{(2m-1)(n-1)}{mn},\end{aligned}$$

where we have used formulas (S15) and (S16) from section S7. Because both edges correspond to treatment wins, we have

$$\mathbb{E}_{B2}(C_eC_f) = \mathbb{E}_{B2}(T_eC_f) = \mathbb{E}_{B2}(C_eT_f) = 0.$$

Case 4: An ordered pair (e, f) of distinct treatment edges, each adjacent to the control vertex v:

- Let the edges be $e = (v_1, v_2)$ and $f = (v_3, v_2)$, with corresponding vertex weights are k_1, k_2, k_3. Then the common vertex v_2 is a control vertex. The number of possibilities for $T_e > 0$ is k_1k_2, and for $T_f > 0$ is k_3k_2. Then the number of samples that produce both $T_e > 0$ and $T_f > 0$ is $k_1k_2^2k_3$, and

$$\begin{aligned}\mathbb{E}_{B2}(T_eT_f) &= \frac{1}{m^mn^n} \sum_{k_1+\dots k_m=m} k_1k_3 \binom{N}{k_1, k_2, \dots, k_N} \\ &\quad\times \sum_{l_1+\dots l_n=n} l_2^2 \binom{n}{l_1, l_2, \dots, l_n} \\ &= \frac{1}{m^mn^n}(m-1)m^{m-1}(2n-1)n^{n-1} \\ &= \frac{(m-1)(2n-1)}{mn},\end{aligned}$$

where we have used formulas (S15) and (S16) from section S7.
Because both edges correspond to treatment wins, we have

$$\mathbb{E}_{B2}(C_eC_f) = \mathbb{E}_{B2}(T_eC_f) = \mathbb{E}_{B2}(C_eT_f) = 0.$$

Case 5: An ordered pair (e, f) of distinct control edges, each adjacent to the treatment vertex v:

- The expected value computations are similar to those of case 3.

$$\mathbb{E}_{B2}(C_e C_f) = \frac{(2m-1)(n-1)}{mn}.$$
$$\mathbb{E}_{B2}(T_e T_f) = \mathbb{E}_{B2}(T_e C_f) = \mathbb{E}_{B2}(C_e T_f) = 0.$$

Case 6: An ordered pair (e, f) of distinct control edges, each adjacent to the control vertex v:

- The expected value computations are similar to those of case 4.

$$\mathbb{E}_{B2}(C_e C_f) = \frac{(m-1)(2n-1)}{mn}.$$
$$\mathbb{E}_{B2}(T_e T_f) = \mathbb{E}_{B2}(T_e C_f) = \mathbb{E}_{B2}(C_e T_f) = 0.$$

Case 7: An ordered pair (e, f) of distinct edges, where e is a treatment edge and f is a control edge, each adjacent to the treatment vertex v:

- The expected value computations are similar to those of case 3.

$$\mathbb{E}_{B2}(T_e C_f) = \frac{(2m-1)(n-1)}{mn}.$$
$$\mathbb{E}_{B2}(T_e T_f) = \mathbb{E}_{B2}(C_e T_f) = \mathbb{E}_{B2}(C_e C_f) = 0.$$

Case 8: An ordered pair (e, f) of distinct edges, where e is a treatment edge and f is a control edge, each adjacent to the control vertex v:

- The expected value computations are similar to those of case 4.

$$\mathbb{E}_{B2}(T_e C_f) = \frac{(m-1)(2n-1)}{mn}.$$
$$\mathbb{E}_{B2}(T_e T_f) = \mathbb{E}_{B2}(C_e T_f) = \mathbb{E}_{B2}(C_e C_f) = 0.$$

Case 9: An ordered pair (e, f) of distinct edges, where e is a control edge and f is a treatment edge, each adjacent to the treatment vertex v:

- The expected value computations are similar to those of case 3.

$$\mathbb{E}_{B2}(C_e T_f) = \frac{(2m-1)(n-1)}{mn}.$$
$$\mathbb{E}_{B2}(T_e T_f) = \mathbb{E}_{B2}(T_e C_f) = \mathbb{E}_{B2}(C_e C_f) = 0.$$

Case 10: An ordered pair (e, f) of distinct edges, where e is a control edge and f is a treatment edge, each adjacent to the control vertex v:

- The expected value computations are similar to those of case 4.

$$\mathbb{E}_{B2}(C_e T_f) = \frac{(m-1)(2n-1)}{mn}.$$
$$\mathbb{E}_{B2}(T_e T_f) = \mathbb{E}_{B2}(T_e C_f) = \mathbb{E}_{B2}(C_e C_f) = 0.$$

- This case is not actually needed, because the variance matrix is symmetric.

Computations for the remaining edge pairs.

Case 11: An ordered pair (e, f) of non-intersecting edges, both treatment edges:

- Let the edges be $e = (v_1, v_2)$ and $f = (v_3, v_4)$, with corresponding vertex frequencies k_1, k_2, k_3, k_4. The number of possibilities for $T_e > 0$ is $k_1 k_2$, and for $T_f > 0$

is k_3k_4. Then the number of samples that produce both $T_e > 0$ and $T_f > 0$ is $k_1k_2k_3k_4$, and

$$\begin{aligned}\mathbb{E}_{B2}(T_eT_f) &= \frac{1}{m^m n^n} \sum_{k_1+\ldots k_m=m} k_2k_4 \binom{N}{k_1, k_2, \ldots, k_N} \\ &\quad\times \sum_{l_1+\ldots l_n=n} l_1l_3 \binom{n}{l_1, l_2, \ldots, l_n} \\ &= \frac{1}{m^m n^n}(m-1)m^{m-1}(n-1)n^{n-1} \\ &= \frac{(m-1)(n-1)}{mn}.\end{aligned}$$

Because both edges correspond to treatment wins, we have

$$\mathbb{E}_{B2}(C_eC_f) = \mathbb{E}_{B2}(T_eC_f) = \mathbb{E}_{B2}(C_eT_f) = 0.$$

Case 12: An ordered pair (e, f) of non-intersecting edges, both control edges. Following similar reasoning as in case 11 we have:

$$\begin{aligned}\mathbb{E}_{B2}(C_eC_f) &= \frac{(m-1)(n-1)}{mn}. \\ \mathbb{E}_{B2}(T_eT_f) &= \mathbb{E}_{B2}(T_eC_f) = \mathbb{E}_{B2}(C_eT_f) = 0.\end{aligned}$$

Case 13: An ordered pair (e, f) of non-intersecting edges, where e is a treatment edge and f is a control edge. Following similar reasoning as in case 11 we have:

$$\begin{aligned}\mathbb{E}_{B2}(T_eC_f) &= \frac{(m-1)(n-1)}{mn} \\ \mathbb{E}_{B2}(T_eT_f) &= \mathbb{E}_{B2}(C_eT_f) = \mathbb{E}_{B2}(C_eC_f) = 0.\end{aligned}$$

Case 14: An ordered pair (e, f) of non-intersecting edges, where e is a control edge and f is a treatment edge. Following similar reasoning as in case 11 we have:

$$\begin{aligned}\mathbb{E}_{B2}(C_eT_f) &= \frac{(m-1)(n-1)}{mn} \\ \mathbb{E}_{B2}(T_eT_f) &= \mathbb{E}_{B2}(T_eC_f) = \mathbb{E}_{B2}(C_eC_f) = 0.\end{aligned}$$

The algorithm is now complete. We evaluate the variance (S4) by examining each ordered pair of edges (e, f), evaluating $\mathbb{E}_{B2}(T_eT_f)$, $\mathbb{E}_{B2}(T_eC_f)$, $\mathbb{E}_{B2}(C_eT_f)$, and $\mathbb{E}_{B2}(C_eC_f)$. This gives an $2E \times 2E$ matrix, where E is the number of edges in the graph $\mathbb{G}$. Pre- and post-multiply by the W terms. The variance matrix for the pair (W_T, W_C) is obtained by summing each of the four $E \times E$ blocks of this larger matrix.

Unfortunately, the algorithm as stated is $O(N^4)$, because the number of edges is $O(N^2)$, and we have explicitly computed all the entries of an $2E \times 2E$ matrix. For a practical algorithm we need to reduce the entire computation to $O(N^2)$.

S5.3. Practical algorithm in the two-sample bootstrap case

The key to a practical computation is to look at cases 1-10 at each vertex separately. Then sum those values to include in the four entries of (S4). Finally, compute the numbers of pairs in each of cases 11 - 14, and include those terms also.

We also need a bit of notation that is specific to the bootstrap computations.

Let the indicator $D_v = 1$ when $D_i = 1$ and $D_j = 0$, and $D_v = 0$ when $D_i = 0$ and $D_j = 1$. An edge corresponding to a treatment win ($u_{ij} > 0$ for $D_v = 1$ and $u_{ij} < 0$ for $D_v = 0$) or a control win ($u_{ij} < 0$ for $D_v = 1$ and $u_{ij} > 0$ for $D_v = 0$) in the observed data will be called a *treatment edge* or a *control edge*.

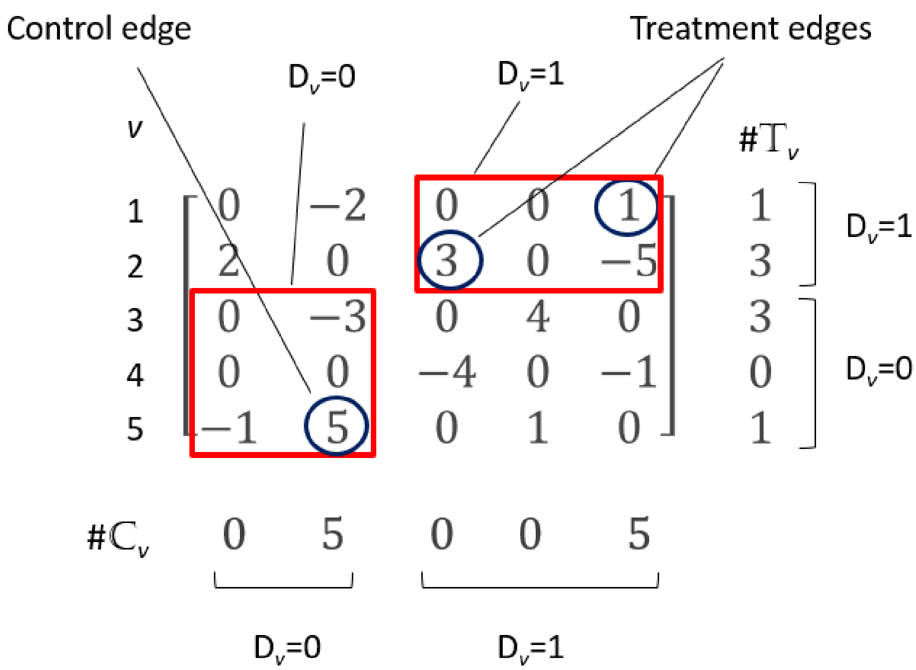

Figure S1. Relevant terminology for the bootstrap distribution of a score matrix U. The edges in the red rectangles are the edges contributing to the wins, $D_v = 1$ is representing the treatment subjects and $D_v = 0$ the control subjects and $\#\mathbb{T}_v$ respectively $\#\mathbb{C}_v$ denoting the number of treatment or control edges adjacent to the vertex v.

$$
\begin{aligned}
\#\mathbb{T}_v &= \sum w_e \quad \text{over treatment edges adjacent to vertex } v \text{ or} \\
&\quad \text{the row sums of } w_e > 0\ (w_e < 0) \text{ for vertex } v \text{ with } D_{v=1}\ (D_{v=0}). \\
\#\mathbb{C}_v &= \sum w_e \quad \text{over control edges adjacent to vertex } v \text{ or} \\
&\quad \text{the column sums of } w_e > 0\ (w_e < 0) \text{ for vertex } v \text{ with } D_{v=0}\ (D_{v=1}). \\
\#\mathbb{T}_v^s &= \sum w_e^2 \quad \text{over treatment edges adjacent to vertex } v \text{ or} \\
&\quad \text{the row sums of } w_e^2 \text{ for } w_e > 0\ (w_e < 0) \text{ for vertex } v \text{ with } D_{v=1}\ (D_{v=0}). \\
\#\mathbb{C}_v^s &= \sum w_e^2 \quad \text{over control edges adjacent to vertex } v \text{ or} \\
&\quad \text{the column sums of } w_e^2 \text{ for } w_e > 0\ (w_e < 0) \text{ for vertex } v \text{ with } D_{v=0}\ (D_{v=1}).
\end{aligned}
$$

or due to symmetry of the U-matrix (Figure S1):

$$
\begin{aligned}
\#\mathbb{T}_v &= \sum w_e \quad \text{the row } (D_{v=1}) \text{ and column } (D_{v=0}) \text{ sums of } w_e > 0 \text{ for vertex } v. \\
\#\mathbb{C}_v &= \sum w_e \quad \text{the row } (D_{v=1}) \text{ and column } (D_{v=0}) \text{ sums of } w_e < 0 \text{ for vertex } v. \\
\#\mathbb{T}_v^s &= \sum w_e^2 \quad \text{the row } (D_{v=1}) \text{ and column } (D_{v=0}) \text{ sums of } w_e^2 \text{ for } w_e > 0 \text{ vertex } v. \\
\#\mathbb{C}_v^s &= \sum w_e^2 \quad \text{the row } (D_{v=1}) \text{ and column } (D_{v=0}) \text{ sums of } w_e^2 \text{ for } w_e < 0 \text{ vertex } v.
\end{aligned}
$$

The distinction between $\#\mathbb{T}_v^s$ and $\#\mathbb{T}_v^2$ should be made clear. In the former the appropriate values w_e are squared, and then summed; in the latter the values w_e are summed, and then the sum is squared.

Case 1: For each treatment edge e adjacent to v, the contribution to the square terms in the variance V_{TT} is $w_e^2 \frac{(2m-1)(2n-1)}{mn}$. Summing over all the treatment edges adjacent to v, we have the contribution to the square terms

$$\begin{aligned} \frac{(2m-1)(2n-1)}{mn} \#\mathbb{T}_v^s &\quad \text{for } V_{TT} \\ 0 &\quad \text{for } V_{TC}, V_{CT}, \text{and } V_{CC} \end{aligned}$$

Case 2: For each control edge e adjacent to v, the contribution to the square terms in the variance V_{CC} is $w_e^2 \frac{(2m-1)(2n-1)}{mn}$. Summing over all the control edges adjacent to v, we have the contribution to the square terms

$$\begin{aligned} \frac{(2m-1)(2n-1)}{mn} \#\mathbb{C}_v^s &\quad \text{for } V_{CC} \\ 0 &\quad \text{for } V_{TT}, V_{TC}, \text{and } V_{CT} \end{aligned}$$

Case 3: For an ordered pair (e, f) of distinct treatment edges adjacent to the treatment vertex v, the contribution to the square terms in the variance V_{TT} is $w_e w_f \frac{(2m-1)(n-1)}{mn}$. Summing over all the the pairs, we have the contribution to the square terms

$$\begin{aligned} \frac{(2m-1)(n-1)}{mn} (\#\mathbb{T}_v^2 - \#\mathbb{T}_v^s) &\quad \text{for } V_{TT} \\ 0 &\quad \text{for } V_{TC}, V_{CT}, \text{and } V_{CC} \end{aligned}$$

The reason for the subtraction in $(\#\mathbb{T}_v^2 - \#\mathbb{T}_v^s)$ is that we are looking at pairs of *distinct* edges; the ordered pairs of identical edges were already counted in case 1.

Case 4: For an ordered pair (e, f) of distinct treatment edges adjacent to the control vertex v, the contribution to the square terms in the variance V_{TT} is $w_e w_f \frac{(m-1)(2n-1)}{mn}$. Summing over all the the pairs, we have the contribution to the square terms

$$\begin{aligned} \frac{(m-1)(2n-1)}{mn} (\#\mathbb{T}_v^2 - \#\mathbb{T}_v^s) &\quad \text{for } V_{TT} \\ 0 &\quad \text{for } V_{TC}, V_{CT}, \text{and } V_{CC} \end{aligned}$$

Case 5: For an ordered pair (e, f) of distinct control edges adjacent to the treatment vertex v, the contribution to the square terms in the variance V_{CC} is $w_e w_f \frac{(2m-1)(n-1)}{mn}$. Summing over all the the pairs, we have the contribution to the square terms

$$\begin{aligned} \frac{(2m-1)(n-1)}{mn} (\#\mathbb{C}_v^2 - \#\mathbb{C}_v^s) &\quad \text{for } V_{CC} \\ 0 &\quad \text{for } V_{TT}, V_{TC}, \text{and } V_{CT} \end{aligned}$$

Case 6: For an ordered pair (e, f) of distinct control edges adjacent to the control vertex v, the contribution to the square terms in the variance V_{CC} is $w_e w_f \frac{(m-1)(2n-1)}{mn}$. Summing over all the the pairs, we have the contribution to the square terms

$$\begin{aligned} \frac{(m-1)(2n-1)}{mn} (\#\mathbb{C}_v^2 - \#\mathbb{C}_v^s) &\quad \text{for } V_{CC} \\ 0 &\quad \text{for } V_{TT}, V_{TC}, \text{and } V_{CT} \end{aligned}$$

Case 7: For an ordered pair (e, f) of distinct edges, where e is a treatment edge and f is a control edge, adjacent to the treatment vertex v, the contribution to the square terms in the variance V_{TC} is $w_e w_f \dfrac{(2m-1)(n-1)}{mn}$. Summing over all the the pairs, we have the contribution to the square terms

$$\begin{aligned} \frac{(2m-1)(n-1)}{mn}\#\mathbb{T}_v\#\mathbb{C}_v &\quad \text{for } V_{TC} \\ 0 &\quad \text{for } V_{TT}, V_{CT}, \text{and } V_{CC} \end{aligned}$$

Case 8: For an ordered pair (e, f) of distinct edges, where e is a treatment edge and f is a control edge, adjacent to the control vertex v, the contribution to the square terms in the variance V_{TC} is $w_e w_f \dfrac{(m-1)(2n-1)}{mn}$. Summing over all the the pairs, we have the contribution to the square terms

$$\begin{aligned} \frac{(m-1)(2n-1)}{mn}\#\mathbb{T}_v\#\mathbb{C}_v &\quad \text{for } V_{TC} \\ 0 &\quad \text{for } V_{TT}, V_{CT}, \text{and } V_{CC} \end{aligned}$$

Case 9: For an ordered pair (e, f) of distinct edges, where e is a control edge and f is a treatment edge, adjacent to the treatment vertex v, the contribution to the square terms in the variance V_{CT} is $w_e w_f \dfrac{(2m-1)(n-1)}{mn}$. Summing over all the the pairs, we have the contribution to the square terms

$$\begin{aligned} \frac{(2m-1)(n-1)}{mn}\#\mathbb{T}_v\#\mathbb{C}_v &\quad \text{for } V_{CT} \\ 0 &\quad \text{for } V_{TT}, V_{TC}, \text{and } V_{CC} \end{aligned}$$

Case 10: For an ordered pair (e, f) of distinct edges, where e is a control edge and f is a treatment edge, adjacent to the control vertex v, the contribution to the square terms in the variance V_{CT} is $w_e w_f \dfrac{(m-1)(2n-1)}{mn}$. Summing over all the the pairs, we have the contribution to the square terms

$$\begin{aligned} \frac{(m-1)(2n-1)}{mn}\#\mathbb{T}_v\#\mathbb{C}_v &\quad \text{for } V_{CT} \\ 0 &\quad \text{for } V_{TT}, V_{TC}, \text{and } V_{CC} \end{aligned}$$

Case 11: For an ordered pair (e, f) of non-intersecting treatment edges, the contribution to the square terms in the variance V_{TT} is $w_e w_f \dfrac{(m-1)(n-1)}{mn}$. The sum of the $w_e w_f$ terms for all ordered pairs of treatment edges would be W_T^2. We have already included the terms $w_e w_f$ for single edges (Case 1) and for pairs of intersecting edges (Cases 3 and 4). Accordingly the remaining variance computation is

$$\begin{aligned} \left[W_T^2 - \sum_{D_v=1}\#\mathbb{T}_v^s - \sum_v\left(\#\mathbb{T}_v^2 - \#\mathbb{T}_v^s\right)\right]\frac{(m-1)(n-1)}{mn} &\quad \text{for } V_{TT} \\ 0 &\quad \text{for } V_{TC}, V_{CT}, \text{and } V_{CC} \end{aligned}$$

Case 12: For an ordered pair (e, f) of non-intersecting control edges, the contribution to the square terms in the variance V_{CC} is $w_e w_f \dfrac{(m-1)(n-1)}{mn}$. The sum of the $w_e w_f$ terms for all ordered pairs of treatment edges would be W_C^2. We have already included the terms $w_e w_f$ for single edges (Case 2) and for pairs of intersecting edges (Cases 5 and 6). Accordingly the remaining variance computation is

$$\begin{aligned}\left[W_C^2 - \sum_{D_v=1} \#\mathbb{C}_v^s - \sum_v \left(\#\mathbb{C}_v^2 - \#\mathbb{C}_v^s\right)\right] \frac{(m-1)(n-1)}{mn} \quad &\text{for } V_{CC} \\ 0 \quad &\text{for } V_{TT}, V_{TC}, \text{and } V_{CT}\end{aligned}$$

Case 13: For an ordered pair (e,f) of non-intersecting edges, where e is a treatment edge and f is a control edge, the contribution to the square terms in the variance V_{TC} is $w_e w_f \dfrac{(m-1)(n-1)}{mn}$. We have already included the terms $w_e w_f$ for pairs of intersecting edges (Cases 7 and 8). Accordingly the remaining variance computation is

$$\begin{aligned}\left[W_T W_C - \sum_v \#\mathbb{T}_v \#\mathbb{C}_v\right] \frac{(m-1)(n-1)}{mn} \quad &\text{for } V_{TC} \\ 0 \quad &\text{for } V_{TT}, V_{CT}, \text{and } V_{CT}\end{aligned}$$

Case 14: For an ordered pair (e,f) of non-intersecting edges, where e is a control edge and f is a treatment edge, the contribution to the square terms in the variance V_{TC} is $w_e w_f \dfrac{(m-1)(n-1)}{mn}$. We have already included the terms $w_e w_f$ for pairs of intersecting edges (Cases 9 and 10). Accordingly the remaining variance computation is

$$\begin{aligned}\left[W_T W_C - \sum_v \#\mathbb{T}_v \#\mathbb{C}_v\right] \frac{(m-1)(n-1)}{mn} \quad &\text{for } V_{CT} \\ 0 \quad &\text{for } V_{TT}, V_{TC}, \text{and } V_{CT}\end{aligned}$$

The variances for W_T and W_C are thus:

$$\begin{aligned}\mathrm{Var}_B(W_T) = &\frac{(2m-1)(2n-1)}{mn} \sum_{D_v=1} \#\mathbb{T}_v^s + \frac{(2m-1)(n-1)}{mn} \sum_{D_v=1} (\#\mathbb{T}_v^2 - \#\mathbb{T}_v^s) + \\ &\frac{(m-1)(2n-1)}{mn} \sum_{D_v=0} (\#\mathbb{T}_v^2 - \#\mathbb{T}_v^s) + \frac{(m-1)(n-1)}{mn} \Big(W_T^2 - \sum_{D_v=1} \#\mathbb{T}_v^s - \\ &\sum_v (\#\mathbb{T}_v^2 - \#\mathbb{T}_v^s) \Big) - W_T^2. \\ \mathrm{Var}_B(W_C) = &\frac{(2m-1)(2n-1)}{mn} \sum_{D_v=1} \#\mathbb{C}_v^s + \frac{(2m-1)(n-1)}{mn} \sum_{D_v=1} (\#\mathbb{C}_v^2 - \#\mathbb{C}_v^s) + \\ &\frac{(m-1)(2n-1)}{mn} \sum_{D_v=0} (\#\mathbb{C}_v^2 - \#\mathbb{C}_v^s) + \frac{(m-1)(n-1)}{mn} \Big(W_C^2 - \sum_{D_v=1} \#\mathbb{C}_v^s - \\ &\sum_v (\#\mathbb{C}_v^2 - \#\mathbb{C}_v^s) \Big) - W_C^2.\end{aligned}$$

and the covariance:

$$\begin{aligned}\mathrm{Cov}_B(W_T W_C) = &\frac{(2m-1)(n-1)}{mn} \sum_{D_v=1} \#\mathbb{T}_v \#\mathbb{C}_v + \frac{(m-1)(2n-1)}{mn} \sum_{D_v=0} \#\mathbb{T}_v \#\mathbb{C}_v \\ &+ \frac{(m-1)(n-1)}{mn} \Big(W_T W_C - \sum_v \#\mathbb{T}_v \#\mathbb{C}_v \Big) - W_T W_C.\end{aligned}$$

S5.4. Complexity of the final bootstrap algorithm

For a specific vertex v, in order to compute the various vertex dependent terms in the variance formulas, ($\#\mathbb{T}_v, \#\mathbb{C}_v, \#\mathbb{T}_v^s$, and $\#\mathbb{C}_v^s$) each other vertex must be examined once. Thus the complexity of the computation at each vertex is $O(N)$, and computing at all vertices is thus $O(N^2)$. The total number of edges and the wins W_T and W_C are computed from these numbers in an additional $O(N)$ steps, and the final computations are $O(1)$. Accordingly the time complexity of the bootstrap algorithm is $O(N^2)$. No algorithm can have time complexity less than $O(N^2)$, because we must examine each entry of the original $N \times N$ matrix at least once.

We also note that the edges within trial arms (the so-called *neutral* edges), never entered into any of the computations. Accordingly these edges do not influence the variance computation. This is in contrast to the permutation situation, where the neutral edges do impact the variance.

S6. The one-sample bootstrap distribution

In a one-sample bootstrap test, subjects are randomly re-sampled with replacement from the entire population. Accordingly, the trial arm counts in a bootstrap sample will typically differ from those in the observed data. Although sampling from the trial arms separately seems to be more common; this method might be used on occasion. With a population of size N patients, there are N^N possible bootstrap samples, and our underlying assumption is that all are equally likely. If all possible bootstrap samples, are considered, the the treatment win sum W_T and control win sum W_C will have the complete bootstrap distribution. The means, variances, and covariance of this bootstrap distribution of W_T and W_C will be calculated explicitly. These values will also be the expected means, variances, and covariances from a randomized bootstrap sample.

Neutral edges in the observed data can never correspond to wins in a bootstrap sample, and such edges play no role in the computations below.

We will use a number of easily derived identities involving multinomial coefficients. These are summarized in the appendix.

We also need a bit of notation that is specific to the bootstrap computations.

For expected values, we use the symbol $\mathbb{E}_{B1}$, as a specification of the general expection symbol $\mathbb{E}$.

$$\begin{aligned}
\#\mathbb{T}_v &= \sum w_e \quad \text{over treatment edges adjacent to vertex } v. \\
\#\mathbb{C}_v &= \sum w_e \quad \text{over control edges adjacent to vertex } v. \\
\#\mathbb{T}_v^s &= \sum w_e^2 \quad \text{over treatment edges adjacent to vertex } v. \\
\#\mathbb{C}_v^s &= \sum w_e^2 \quad \text{over control edges adjacent to vertex } v.
\end{aligned}$$

The distinction between $\#\mathbb{T}_v^s$ and $\#\mathbb{T}_v^2$ should be made clear. In the former the appropriate values w_e are squared, and then summed; in the latter the values w_e are summed, and then the sum is squared.

When sampling from the entire population, a vertex frequency can have any value in the interval $[0, N]$. The number of ways that a vertex frequency can have the value k_1 is given by $\sum_{k_2+\cdots+K_N=N-k_1} \binom{N}{k_1,k_2,\ldots,k_N}$, and the number of ways a pair of vertices can have frequencies (k_1, k_2) is given by $\sum_{k_3+\cdots+K_N=N-k_1-k_2} \binom{N}{k_1,k_2,\ldots,k_N}$.

S6.1. Expectations

Consider a treatment edge e. For the expected value of T_e, we use equation (S14) from the appendix.

$$\mathbb{E}_{B1}(T_e) = \frac{1}{N^N} \sum_{k_1+\ldots k_N=N} k_1 k_2 \binom{N}{k_1, k_2, \ldots, k_N}$$
$$= \frac{1}{N^N}(N-1)N^{N-1} = \frac{N-1}{N} \tag{S10}$$

For this edge we will have $\mathbb{E}_{B1}(C_e) = 0$, since this edge cannot correspond to a control win in any bootstrap sample. Similarly, for a control edge f, $\mathbb{E}_{B1}(T_f) = 0$ and $\mathbb{E}_{B1}(C_f) = (N-1)/N$.

S6.2. Variances

For the variance, again the expectations of $\mathbb{E}_{B1}(T_e T_f)$, $\mathbb{E}_{B1}(C_e C_f)$, $\mathbb{E}_{B1}(T_e C_f)$ and $\mathbb{E}_{B1}(C_e T_f)$ need to be calculated, and these depend on the geometric configuration of the edges. We considered various cases of *ordered* pairs. The following cases are considered: an ordered pair of identical treatment edges (case 1), an ordered pair of identical control edges (case 2), an ordered pair of distinct treatment edges adjacent to a vertex (case 3), an ordered pair of distinct control edges adjacent to a vertex (case 4), an ordered pair of a treatment edge and a control edge adjacent to a vertex (case 5), and an ordered pair of a control edge and a treatment edge adjacent to a vertex (case 6). Finally, for non-intersecting edges, 4 cases are considered: an ordered pair of distinct treatment edges (case 7), an ordered pair of distinct control edges (case 8), an ordered pair of a treatment and a control edge (case 9), and an ordered pair of a control edge and a treatment edge (case 10).

> In principle some of the above cases could be combined, by considering a mixture of ordered and unordered pairs of edges. It seems clearer to use ordered pairs throughout.

Case 1: An ordered pair of identical treatment edges. For an edge e of $\mathbb{G}$ that corresponds to a treatment win

$$\mathbb{E}_{B1}(T_e^2) = \frac{1}{N^N} \sum_{k_1+\ldots k_N=N} k_1^2 k_2^2 \binom{N}{k_1, k_2, \ldots, k_N}$$
$$= \frac{(N-1)(4N^2-9N+6)}{N^3},$$

where we have used equation (S17) from section S7. For this same edge, $C_e = 0$ for all bootstrap samples, and hence $\mathbb{E}_{B1}(C_e^2) = \mathbb{E}_{B1}(T_e C_e) = \mathbb{E}_{B1}(C_e T_e) = 0$.

Case 2: An ordered pair of identical control edges. For an edge e of $\mathbb{G}$ that corresponds to a treatment win

$$\mathbb{E}_{B1}(C_e^2) = \frac{1}{N^N} \sum_{k_1+\ldots k_N=N} k_1^2 k_2^2 \binom{N}{k_1, k_2, \ldots, k_N}$$
$$= \frac{(N-1)(4N^2-9N+6)}{N^3}.$$

For this same edge, $T_e = 0$ for all bootstrap samples, and hence $\mathbb{E}_{B1}(T_e^2) = \mathbb{E}_{B1}(T_e C_e) = \mathbb{E}_{B1}(C_e T_e) = 0$.

Case 3: An ordered pair (e, f) of distinct treatment edges, each adjacent to a vertex v:

- Let the edges be $e = (v_1, v_2)$ and $f = (v_1, v_3)$, with corresponding vertex frequencies k_1, k_2, k_3. The number of possibilities for $T_e > 0$ is $k_1 k_2$, and for $T_f > 0$

is k_1k_3. Then the number of samples that produce both $T_e > 0$ and $T_f > 0$ is $k_1^2k_2k_3$, and

$$\begin{aligned}\mathbb{E}_{B1}(T_eT_f) &= \frac{1}{N^N}\sum_{k_1+\ldots k_N=N} k_1^2k_2k_3\binom{N}{k_1,k_2,\ldots,k_N} \\ &= \frac{1}{N^N}(N-1)(N-2)(2N-3)N^{N-3} \\ &= \frac{(N-1)(N-2)(2N-3)}{N^3},\end{aligned} \tag{S11}$$

where we have used formula (S18) from section S7.
Because both edges correspond to treatment wins, we have

$$\mathbb{E}_{B1}(C_eC_f) = \mathbb{E}_{B1}(T_eC_f) = \mathbb{E}_{B1}(C_eT_f) = 0.$$

The above computations were for a common treatment vertex. The computations would be exactly the same if the common vertex was a control vertex.

Case 4: An ordered pair (e, f) of distinct control edges, each adjacent to a vertex v:

- The computations are the same as in case 3, except that now the contribution is to $\mathbb{E}_{B1}(C_eC_f)$. Accordingly

$$\begin{aligned}\mathbb{E}_{B1}(C_eC_f) &= \frac{(N-1)(N-2)(2N-3)}{N^3} \\ \mathbb{E}_{B1}(T_eT_f) &= \mathbb{E}_{B1}(T_eC_f) = \mathbb{E}_{B1}(C_eT_f) = 0\end{aligned}$$

Case 5: An ordered pair (e, f) of distinct edges, where e is a treatment edge and f is a control edge, each adjacent to a vertex v:

- The computations are the same as in case 3. Accordingly

$$\begin{aligned}\mathbb{E}_{B1}(T_eC_f) &= \frac{(N-1)(N-2)(2N-3)}{N^3} \\ \mathbb{E}_{B1}(T_eT_f) &= \mathbb{E}_{B1}(C_eT_f) = \mathbb{E}_{B1}(C_eC_f) = 0\end{aligned}$$

Case 6: An ordered pair (e, f) of distinct edges, where e is a control edge and f is a treatment edge, each adjacent to a vertex v:

- The computations are the same as in case 3. Accordingly

$$\begin{aligned}\mathbb{E}_{B1}(C_eT_f) &= \frac{(N-1)(N-2)(2N-3)}{N^3} \\ \mathbb{E}_{B1}(T_eT_f) &= \mathbb{E}_{B1}(T_eC_f) = \mathbb{E}_{B1}(C_eC_f) = 0\end{aligned}$$

Case 7: An ordered pair (e, f) of non-intersecting edges, both treatment edges:

- Let the edges be $e = (v_1, v_2)$ and $f = (v_3, v_4)$, with corresponding vertex frequencies k_1, k_2, k_3, k_4. The number of possibilities for $T_e > 0$ is k_1k_2, and for $T_f > 0$

is k_3k_4. Then the number of samples that produce both $T_e > 0$ and $T_f > 0$ is $k_1k_2k_3k_4$, and

$$\begin{aligned}\mathbb{E}_{B1}(T_eT_f) &= \frac{1}{N^N}\sum_{k_1+\ldots k_N=N} k_1k_2k_3k_4\binom{N}{k_1,k_2,\ldots,k_N} \\ &= \frac{1}{N^N}(N-1)(N-2)(N-3)N^{N-3} \\ &= \frac{(N-1)(N-2)(N-3)}{N^3},\end{aligned}$$

where we have used formula (S19) from section S7. Because both edges correspond to treatment wins, we have

$$\mathbb{E}_{B1}(C_eC_f) = \mathbb{E}_{B1}(T_eC_f) = \mathbb{E}_{B1}(C_eT_f) = 0.$$

Case 8: An ordered pair (e, f) of non-intersecting edges, both control edges. Following similar reasoning as in case 7 we have:

$$\begin{aligned}\mathbb{E}_{B1}(C_eC_f) &= \frac{(N-1)(N-2)(N-3)}{N^3}. \\ \mathbb{E}_{B1}(T_eT_f) &= \mathbb{E}_{B1}(T_eC_f) = \mathbb{E}_{B1}(C_eT_f) = 0.\end{aligned}$$

Case 9: An ordered pair (e, f) of non-intersecting edges, where e is a treatment edge and f is a control edge. Following similar reasoning as in case 7 we have:

$$\begin{aligned}\mathbb{E}_{B1}(T_eC_f) &= \frac{(N-1)(N-2)(N-3)}{N^3} \\ \mathbb{E}_{B1}(T_eT_f) &= \mathbb{E}_{B1}(C_eT_f) = \mathbb{E}_{B1}(C_eC_f) = 0.\end{aligned}$$

Case 10: An ordered pair (e, f) of non-intersecting edges, where e is a control edge and f is a treatment edge. Following similar reasoning as in case 7 we have:

$$\begin{aligned}\mathbb{E}_{B1}(C_eT_f) &= \frac{(N-1)(N-2)(N-3)}{N^3} \\ \mathbb{E}_{B1}(T_eT_f) &= \mathbb{E}_{B1}(T_eC_f) = \mathbb{E}_{B1}(C_eC_f) = 0.\end{aligned}$$

The algorithm is now complete. We evaluate the variance (S4) by examining each ordered pair of edges (e, f), evaluating $\mathbb{E}_{B2}(T_eT_f)$, $\mathbb{E}_{B2}(T_eC_f)$, $\mathbb{E}_{B2}(C_eT_f)$, and $\mathbb{E}_{B2}(C_eC_f)$. This gives an $2E \times 2E$ matrix, where E is the number of edges in the graph $\mathbb{G}$. Pre- and post-multiply by the W terms. The variance matrix for the pair (W_T, W_C) is obtained by summing each of the four $E \times E$ blocks of this larger matrix.

Unfortunately, the algorithm as stated is $O(N^4)$, because the number of edges is $O(N^2)$, and we have explicitly computed all the entries of an $2E \times 2E$ matrix. For a practical algorithm we need to reduce the entire computation to $O(N^2)$.

S6.3. Practical algorithm in the one-sample bootstrap case

The key to a practical computation is to look at cases 1-6 at each vertex separately. Then sum those values to include in the four entries of (S4). Finally, compute the numbers of pairs in each of cases 11 - 14, and include those terms also.

Case 1: For each treatment edge e adjacent to v, the contribution to the square terms in the variance V_{TT} is $w_e^2 \dfrac{(N-1)(4N^2-9N+6)}{N^3}$. Summing over all the treatment edges adjacent to v we have the contribution to the square terms

$$\begin{aligned} \frac{(N-1)(4N^2-9N+6)}{N^3} \#\mathbb{T}_v^s &\quad \text{for } V_{TT} \\ 0 &\quad \text{for } V_{TC}, V_{CT}, \text{and } V_{CC} \end{aligned}$$

It should be noted that when we add these terms from all vertices, each edge would be counted twice. An appropriate correction is needed in the actual computer program. The same comment applies to Case 2.

Case 2: For each control edge e adjacent to v, the contribution to the square terms in the variance V_{CC} is $w_e^2 \dfrac{(N-1)(4N^2-9N+6)}{N^3}$. Summing over all the control edges adjacent to v, and correcting for the fact that an edge will be counted at both ends, we have the contribution to the square terms

$$\begin{aligned} \frac{(N-1)(4N^2-9N+6)}{N^3} \#\mathbb{C}_v^s &\quad \text{for } V_{CC} \\ 0 &\quad \text{for } V_{TT}, V_{TC}, \text{and } V_{CT} \end{aligned}$$

Case 3: For an ordered pair (e, f) of distinct treatment edges adjacent to the vertex v, the contribution to the square terms in the variance V_{TT} is $w_e w_f \dfrac{(N-1)(N-2)(2N-3)}{N^3}$. Summing over all the the pairs, we have the contribution to the square terms

$$\begin{aligned} \frac{(N-1)(N-2)(2N-3)}{N^3} (\#\mathbb{T}_v^2 - \#\mathbb{T}_v^s) &\quad \text{for } V_{TT} \\ 0 &\quad \text{for } V_{TC}, V_{CT}, \text{and } V_{CC} \end{aligned}$$

The reason for the subtraction in $(\#\mathbb{T}_v^2 - \#\mathbb{T}_v^s)$ is that we are looking at pairs of *distinct* edges; the ordered pairs of identical edges were already counted in case 1.

Case 4: For an ordered pair (e, f) of distinct control edges adjacent to the vertex v, the contribution to the square terms in the variance V_{CC} is $w_e w_f \dfrac{(N-1)(N-2)(2N-3)}{N^3}$. Summing over all the the pairs, we have the contribution to the square terms

$$\begin{aligned} \frac{(N-1)(N-2)(2N-3)}{N^3} (\#\mathbb{C}_v^2 - \#\mathbb{C}_v^s) &\quad \text{for } V_{CC} \\ 0 &\quad \text{for } V_{TT}, V_{TC}, \text{and } V_{CT} \end{aligned}$$

Case 5: For an ordered pair (e, f) of distinct edges, where e is a treatment edge and f is a control edge, each adjacent to the vertex v, the contribution to the square terms in the variance V_{TC} is $w_e w_f \dfrac{(N-1)(N-2)(2N-3)}{N^3}$. Summing over all the the pairs, we have the contribution to the square terms

$$\begin{aligned} \frac{(N-1)(N-2)(2N-3)}{N^3} (\#\mathbb{T}_v \#\mathbb{C}_v) &\quad \text{for } V_{TC} \\ 0 &\quad \text{for } V_{TT}, V_{CT}, \text{and } V_{CC} \end{aligned}$$

Case 6: For an ordered pair (e, f) of distinct edges, where e is a control edge and f is a treatment edge, each adjacent to the vertex v, the contribution to the square terms in the variance V_{CT} is $w_e w_f \dfrac{(N-1)(N-2)(2N-3)}{N^3}$. Summing over all the the pairs, we have the contribution to the square terms

$$\begin{aligned} \frac{(N-1)(N-2)(2N-3)}{N^3}(\#\mathbb{T}_v \#\mathbb{C}_v) &\quad \text{for } V_{CT} \\ 0 &\quad \text{for } V_{TT}, V_{TC}, \text{and } V_{CC} \end{aligned}$$

Case 7: For an ordered pair (e, f) of non-intersecting treatment edges, the contribution to the square terms in the variance V_{TT} is $w_e w_f \dfrac{(N-1)(N-2)(N-3)}{N^3}$. The sum of the $w_e w_f$ terms for all ordered pairs of treatment edges would be W_T^2. We have already included the terms $w_e w_f$ for single edges (Case 1) and for pairs of intersecting edges (Case 3). Accordingly the remaining variance computation is

$$\begin{aligned} \left[W_T^2 - \sum_v \#\mathbb{T}_v^2\right] \frac{(N-1)(N-2)(N-3)}{N^3} &\quad \text{for } V_{TT} \\ 0 &\quad \text{for } V_{TC}, V_{CT}, \text{and } V_{CC} \end{aligned}$$

Case 8: For an ordered pair (e, f) of non-intersecting control edges, the contribution to the square terms in the variance V_{CC} is $w_e w_f \dfrac{(N-1)(N-2)(N-3)}{N^3}$. The sum of the $w_e w_f$ terms for all ordered pairs of control edges would be W_C^2. We have already included the terms $w_e w_f$ for single edges (Case 2) and for pairs of intersecting edges (Case 4). Accordingly the remaining variance computation is

$$\begin{aligned} \left[W_C^2 - \sum_v \#\mathbb{C}_v^2\right] \frac{(N-1)(N-2)(N-3)}{N^3} &\quad \text{for } V_{CC} \\ 0 &\quad \text{for } V_{TT}, V_{TC}, \text{and } V_{CT} \end{aligned}$$

Case 9: For an ordered pair (e, f) of non-intersecting edges, where e is a treatment edge and f is a control edge, the contribution to the square terms in the variance V_{TC} is $w_e w_f \dfrac{(N-1)(N-2)(N-3)}{N^3}$. The sum of the $w_e w_f$ terms for all ordered pairs would be $W_T W_C$. We have already included the terms $w_e w_f$ for pairs of intersecting edges (Case 5). Accordingly the remaining variance computation is

$$\begin{aligned} \left[W_T W_C - \sum_v \#\mathbb{T}_v \#\mathbb{C}_v\right] \frac{(N-1)(N-2)(N-3)}{N^3} &\quad \text{for } V_{TC} \\ 0 &\quad \text{for } V_{TT}, V_{CT}, \text{and } V_{CC} \end{aligned}$$

Case 10: For an ordered pair (e, f) of non-intersecting edges, where e is a treatment edge and f is a control edge, the contribution to the square terms in the variance V_{CT} is $w_e w_f \dfrac{(N-1)(N-2)(N-3)}{N^3}$. The sum of the $w_e w_f$ terms for all ordered pairs would be $W_T W_C$. We have already included the terms $w_e w_f$ for pairs of intersecting edges (Case 6). Accordingly the remaining variance computation is

$$\begin{aligned} \left[W_T W_C - \sum_v \#\mathbb{T}_v \#\mathbb{C}_v\right] \frac{(N-1)(N-2)(N-3)}{N^3} &\quad \text{for } V_{CT} \\ 0 &\quad \text{for } V_{TT}, V_{TC}, \text{and } V_{CC} \end{aligned}$$

S6.4. Complexity of the final bootstrap algorithm

For a specific vertex v, in order to compute the various vertex dependent terms in the variance formulas, ($\#\mathbb{T}_v, \#\mathbb{C}_v, \#\mathbb{T}_v^s$, and $\#\mathbb{C}_v^s$) each other vertex must be examined once. Thus the complexity of the computation at each vertex is $O(N)$, and computing at all vertices is thus $O(N^2)$. The total number of edges and the wins W_T and W_C are computed from these numbers in an additional $O(N)$ steps, and the final computations are $O(1)$. Accordingly the time complexity of the bootstrap algorithm is $O(N^2)$. No algorithm can have time complexity less than $O(N^2)$, because we must examine each entry of the original $N \times N$ matrix at least once.

We also note that the edges within trial arms (the so-called *neutral* edges), never entered into any of the computations. Accordingly these edges do not influence the variance computation. This is in contrast to the permutation situation, where the neutral edges do impact the variance.

S7. Multinomial coefficient identities

We look at various identities involving the multinomial coefficients. These are all elementary to derive, and they are summarized here for the convenience of the reader.

The multinomial theorem is

$$(x_1 + x_2 + \cdots + x_N)^N = \sum_{k_1+\ldots k_N=N} x_1^{k_1} x_2^{k_2} \ldots x_N^{k_N} \binom{N}{k_1, k_2, \ldots, k_N} \tag{S12}$$

Letting all $x_i = 1$ in (S12) we find

$$\sum_{k_1+\ldots k_N=N} \binom{N}{k_1, k_2, \ldots, k_N} = N^N. \tag{S13}$$

By taking derivatives of (S12) and then substituting all $x_i = 1$, we obtain a number of further identities.

$$\sum_{k_1+\ldots k_N=N} k_1 \binom{N}{k_1, k_2, \ldots, k_N} = N^N. \tag{S14}$$

$$\sum_{k_1+\ldots k_N=N} k_1^2 \binom{N}{k_1, k_2, \ldots, k_N} = (2N-1)N^{N-1} \tag{S15}$$

$$\sum_{k_1+\ldots k_N=N} k_1 k_2 \binom{N}{k_1, k_2, \ldots, k_N} = N(N-1)N^{N-2} = (N-1)N^{N-1}. \tag{S16}$$

$$\sum_{k_1+\ldots k_N=N} k_1^2 k_2^2 \binom{N}{k_1, k_2, \ldots, k_N} = (N-1)(4N^2-9N+6)N^{N-3} \tag{S17}$$

$$\sum_{k_1+\ldots k_N=N} k_1^2 k_2 k_3 \binom{N}{k_1, k_2, \ldots, k_N} = (N-1)(N-2)(2N-3)N^{N-3} \tag{S18}$$

$$\sum_{k_1+\ldots k_N=N} k_1 k_2 k_3 k_4 \binom{N}{k_1, k_2, \ldots, k_N} = (N-1)(N-2)(N-3)N^{N-3}, \tag{S19}$$