

# Random group formation and the equal-group CR2 variance in finite populations

CausalSmith\*

September 10, 2026

## Abstract

This paper studies design-based variance estimation for random group-formation experiments in finite populations. Units are assigned to disjoint groups of a fixed common size  $M$ , the group-size parameter, and treatment is randomized across realized groups. The target is the finite-population partition-average marginal effect, defined by averaging composition-indexed potential outcomes over the uniform  $M$ -slice of possible groups. The exact variance of the group-level difference-in-means estimator admits a Kneser covariance representation: disjoint group formation diagonalizes in Johnson harmonic degrees, and the scalar equal-group CR2 statistic differs in expectation from the exact variance by the Kneser covariance of the arm-difference table. In triangular arrays with growing group count  $G_n$ , the number of realized groups, stable treatment fraction, bounded potential outcomes, and grouped-unit sampling fraction  $N_n/n$  converging to  $\rho$ , the limiting grouped-unit fraction, the exact design variance  $\sigma_n^2$  of the difference-in-means estimator satisfies, after scaling by  $G_n$ ,

$$G_n \sigma_n^2 = R_n - \rho E_{\tau,1,n} + o(1),$$

where  $R_n$  is the independent-group leading variance scale and  $E_{\tau,1,n}$  is the degree-one Johnson energy of the arm contrast. Consequently, the equal-group CR2 statistic estimates  $R_n/G_n$ , is asymptotically conservative for the exact variance in the stated one-sided probability sense, and is ratio-consistent exactly when the dense correction is negligible relative to the corrected variance scale. At zero sampling density, ratio consistency holds under  $N_n/n \rightarrow 0$ , including regimes with  $N_n^2/n \rightarrow \infty$ . The paper also identifies the scalar statistic returned by a versioned `clubSandwich` CR2 calculation, gives an eight-unit witness with exact variance  $1/7$  and expected CR2 value  $2/7$ , and proves a positive-density lower bound: for a fixed group size  $M \geq 2$ , a strictly positive limiting grouped-unit fraction  $\rho > 0$ , an outcome envelope  $B \geq 1$ , and a variance floor  $c_\sigma$  below the separated independent-arm limit  $1/[Mp(1-p)] - 2\rho/M$ , every measurable one-realization variance statistic has nonvanishing worst-case relative error over the bounded schedule class with those parameters.

## 1 Introduction

Random group formation is common in experiments where outcomes depend on who interacts with whom. Classrooms, teams, peer groups, neighborhoods, and online clusters can all be

---

\*Machine-generated by the CausalSmith research pipeline. Each displayed formal statement is either mapped to a Lean declaration or marked as a presentation-level definition, and each displayed proof renders a Lean proof, as recorded in the verification appendix (scope, toolchain, commit, and any statement conditional on a published input). Cited work otherwise supplies framing and positioning.

randomized objects rather than fixed sampling units. In such designs, the realized groups are typically disjoint: once a unit is placed in one group, it is removed from the pool available to the others. This finite-population dependence changes the variance target for a group-level treatment contrast, even when treatment is randomized across groups after the partition is formed.

This paper gives a design-based account of that variance target for homogeneous random partitions. The finite population has size  $n$ , the common group size is fixed at  $M$ , and  $G_n$  disjoint groups are drawn uniformly before a balanced treatment assignment is applied across groups. Potential outcomes are deterministic functions of the treatment arm and the realized group. The estimand is the partition-average marginal effect, obtained by averaging the contrast of group-level potential-outcome tables over the uniform  $M$ -slice of possible groups. This places random group formation inside the finite-population potential-outcome tradition of [Splawa-Neyman et al. \[1990\]](#), [Fisher \[1935\]](#), [Cox \[1958\]](#), [Rubin \[1974\]](#), and [Holland \[1986\]](#).

The central object is the covariance induced by drawing disjoint groups. On the uniform  $M$ -slice, the normalized Kneser disjointness operator averages a group table over all groups disjoint from a given group. The Johnson association scheme diagonalizes this operator degree by degree. [Theorem 1](#) uses that diagonalization to express the covariance of two centered group tables as a sum of Johnson harmonic inner products multiplied by finite- $n$  Kneser eigenvalues. Applied to the arm-difference table, this identity identifies the exact finite-population covariance term generated by group formation.

The exact design variance follows from combining this Kneser covariance with balanced treatment assignment. [Theorem 2](#) establishes that the group-level difference in means is unbiased for the partition-average marginal effect and that its exact variance decomposes into an independent-group component plus ordered-disjoint covariance terms. The same theorem shows that the expected scalar equal-group CR2 statistic differs from the exact variance by exactly the negative of the Kneser covariance contrast. Thus the discrepancy between the usual cluster calculation and the finite-population randomization variance is determined by the geometry of the full potential group tables.

The asymptotic analysis studies triangular arrays with fixed  $M$ , growing group count, stable treated-group fraction  $p_n \rightarrow p \in (0, 1)$ , bounded potential outcomes, nondegenerate scaled exact variance, and grouped-unit sampling fraction  $N_n/n \rightarrow \rho \in [0, 1]$ . Under these conditions, [Theorem 3](#) gives the expansion

$$G_n \sigma_n^2 = R_n - \rho E_{\tau,1,n} + o(1),$$

where  $R_n$  is the independent-group leading variance scale and  $E_{\tau,1,n}$  is the degree-one Johnson energy of the arm-difference table. The degree-one term has a direct finite-population interpretation: it is the component of the treatment-control group-table contrast carried by marginal unit participation, and disjoint sampling makes that component contribute at first order whenever the limiting grouped-unit fraction is positive.

This expansion yields the CR2 phase frontier. [Theorem 4](#) shows that the scalar equal-group CR2 statistic estimates  $R_n/G_n$  in probability, while the exact variance uses the corrected scale  $(R_n - \rho E_{\tau,1,n})/G_n$ . The theorem gives a one-sided asymptotic conservativeness statement and characterizes ratio consistency by the relative-energy condition

$$\frac{\rho E_{\tau,1,n}}{R_n - \rho E_{\tau,1,n}} \rightarrow 0.$$

At zero sampling density, [Theorem 5](#) gives ratio consistency under  $N_n/n \rightarrow 0$ , including the benchmark  $N_n = M \lfloor n^{3/4}/M \rfloor$ , for which  $N_n^2/n \rightarrow \infty$ . This connects the dense calculation to the sparse intuition in [Fu et al. \[2026\]](#) while identifying the sampling fraction that governs the scalar variance ratio in the present fixed- $M$  random-partition setting.

The paper also relates the population calculation to applied regression software. In the equal-group design, [Proposition 1](#) identifies the treatment-coordinate entry returned by `clubSandwich::vcovCR` with type `CR2`, for the stated unweighted intercept-plus-treatment regression and versioned `sandwich` bread convention, with the scalar statistic analyzed in the design theory. [Proposition 2](#) then gives a finite eight-unit, four-pair, two-treated-group schedule in which the exact variance is  $1/7$  and the expected `CR2` statistic is  $2/7$ . The witness makes the dense correction visible in a small balanced design.

Finally, the one-realization result gives an identification-theoretic interpretation of the dense correction. [Proposition 3](#) constructs same-sign and independent-arm Rademacher schedule priors whose one-realization observed-data mixtures agree while their scaled exact variances converge to separated limits at positive density. [Theorem 6](#) converts this comparison into a uniform lower bound over the dense bounded schedule class: for some positive  $\varepsilon$ , every measurable statistic of the realized partition, assignment, and observed outcomes has worst-case relative error probability at least one half asymptotically. The obstruction is the degree-one cross-arm component of the potential group tables, which enters the exact variance but is not uniformly recoverable from a single realized grouping and assignment.

The formal statements and internal algebraic derivations are machine-checked in Lean 4; the verification appendix records the precise scope. The paper proceeds as follows. [Section 2](#) places the results in the design-based, survey-sampling, cluster-robust, and interference literatures. [Section 3](#) defines the random group-formation design, the partition-average marginal effect, the Johnson decomposition, and the scalar `CR2` statistic. [Section 4](#) proves the Kneser covariance identity and the finite-sample PAME variance formula. [Section 5](#) derives the projection limit, the `CR2` phase frontier, and the zero-density ratio result. [Section 6](#) connects the scalar statistic to the versioned `clubSandwich` calculation and evaluates the eight-unit witness. [Section 7](#) proves the one-realization obstruction, [Section 8](#) summarizes the scope and extensions, and [Section A](#) collects the proofs and the verification note.

## 2 Related work

The paper belongs to the design-based tradition in causal inference, where treatment assignment is the source of randomness and potential outcomes are fixed features of a finite population [[Splawa-Neyman et al., 1990](#), [Fisher, 1935](#), [Cox, 1958](#), [Rubin, 1974](#), [Holland, 1986](#)]. Its population target is a group-composition estimand: for a population of size  $n$  and common group size  $M$ , the analysis averages potential outcomes over the uniform  $M$ -slice of possible groups. This places random group formation inside the same finite-population logic that underlies classical randomization inference, while making the composition of a treated or control group part of the assignment mechanism.

The asymptotic perspective is closest to survey-sampling triangular arrays, where the population may grow, the realized sample may occupy a nonvanishing fraction of it, and finite-population corrections can survive in the limit [[Hájek, 1960](#), [Cochran, 1977](#), [Ohlsson, 1989](#), [Fuller, 2009](#)]. The grouped-unit sampling fraction  $N_n/n$  plays that role here. Under the dense design studied in [Definition 4](#), the number of realized groups  $G_n$  diverges while  $M$  is fixed, the

treated-group fraction  $p_n$  converges to an interior limit, and  $N_n/n$  converges to a limit  $\rho$  in the unit interval. The resulting variance formula retains the finite-population geometry induced by drawing disjoint groups, so the large-sample object is governed by both the usual independent-group variance scale and a correction determined by overlap in potential group tables.

The closest theorem-level comparison is [Fu et al. \[2026\]](#). For group interaction experiments with fixed or randomly formed groups, their Section 3 and Theorems 1–2 establish both asymptotic normality of the group-level contrast and ratio consistency of the CR2 variance estimator, under boundedness, their Assumption 1, and the sparse whole-tuple condition  $N_n^2/n \rightarrow 0$ . That condition makes accidental reuse of population units across independently sampled tuples asymptotically negligible, so the finite-population disjointness geometry disappears from the limit and a single sparsity rate carries both conclusions. The present paper differs on three counts. It works in the homogeneous fixed- $M$  random group-formation subcase and keeps the exact disjointness geometry visible rather than assuming it away. Its subject is the variance object itself: the results characterize the exact design variance and the probability limit of the equal-group CR2 statistic, and [Section 8](#) records what a complete inference theory would still require. And it replaces the whole-tuple rate by the grouped-unit sampling fraction: [Theorem 4](#) shows that what governs the CR2 variance ratio in this subcase is  $N_n/n$ , through the degree-one Johnson correction, and [Theorem 5](#) delivers ratio consistency under  $N_n/n \rightarrow 0$  alone, in particular on schedules with  $N_n^2/n \rightarrow \infty$  that their condition excludes. [Theorem 1](#) expresses that geometry through the Johnson-scheme decomposition of slice functions and the normalized Kneser disjointness operator, while [Theorem 2](#) converts it into the exact variance of the group-level difference-in-means estimator. In the dense regime, [Theorem 3](#) yields the expansion  $G_n \sigma_n^2 = R_n - \rho E_{\tau,1,n} + o(1)$ , where  $R_n$  is the independent-group leading variance scale and  $E_{\tau,1,n}$  is the degree-one Johnson energy of the arm-difference table. At zero sampling density, [Theorem 4](#) gives the corresponding convergence of the scalar CR2 ratio, aligning the dense calculation with the sparse intuition when  $N_n/n \rightarrow 0$ .

The algebra behind these comparisons draws on the Johnson association scheme and Kneser graph spectrum [[Delsarte, 1973](#), [Brouwer et al., 1989](#), [Filmus, 2016](#), [Brouwer et al., 2018](#)]. Those tools are used here as finite-population variance geometry. The group table  $h_{z,n}$  records, for each treatment arm and possible group, the group-average potential outcome; its centered version and Johnson projections separate additive unit-level variation from higher-order composition variation. This decomposition is what makes the dense correction interpretable: the term involving  $\rho$  is attached to the first Johnson component of the arm contrast, the part of the table most directly affected by removing units from the remaining pool after a group has been drawn.

The variance-estimation part of the paper connects to the cluster-robust covariance literature initiated by [White \[1980\]](#) and [Liang and Zeger \[1986\]](#), and to finite-sample and leverage adjustments for clustered least squares [[Bell and McCaffrey, 2002](#), [Cameron and Miller, 2015](#), [Abadie et al., 2023](#)]. The scalar statistic in [Definition 12](#) is the equal-group analogue of the CR2 adjustment used in applied software. Two design-based points of contact deserve emphasis. [Pustejovsky and Tipton \[2018\]](#) develop bias-reduced linearization and Satterthwaite-type testing for fixed-effects models, and the CR2 adjustment they analyze is derived under an independent-cluster working covariance model; that working model is exactly what random group formation violates here, since drawing disjoint groups from a finite population correlates the realized group tables. The present results quantify the resulting gap rather than repairing the adjustment: [Theorem 2](#) shows that the expected scalar CR2 statistic differs from the exact design variance by precisely the Kneser covariance contrast, which is the term the independent-cluster work-

ing model sets to zero. [Su and Ding \[2021\]](#) give a design-based comparison of individual-level, cluster-average, and cluster-total regression analyses of cluster-randomized experiments, showing how the estimand and the robust standard error depend on cluster-size weighting and on the model-assisted structure imposed. Their clusters are fixed by the design; the object studied here is the additional variance geometry contributed by making the clusters themselves random and disjoint, at a common size  $M$  that removes the cluster-size weighting question from the comparison. [Proposition 1](#) gives a software-facing identity for the unweighted intercept-plus-treatment regression and the `clubSandwich` CR2 covariance calculation, tying the population calculation to the implementation line in [Pustejovsky and Tipton \[2018\]](#), [Zeileis \[2004\]](#), [Zeileis \[2006\]](#), [Zeileis et al. \[2020\]](#), [Pustejovsky \[2026\]](#), and [Zeileis and Lumley \[2026\]](#). This identity is useful because it pins down the scalar object consumed by standard regression software before the asymptotic comparison is made.

The design also sits near the literature on interference, exposure mappings, and group-formation experiments [[Sobel, 2006](#), [Hudgens and Halloran, 2008](#), [Tchetgen and VanderWeele, 2012](#), [Manski, 2013](#), [Liu and Hudgens, 2014](#), [Aronow and Samii, 2017](#), [Ugander et al., 2013](#), [Eckles et al., 2017](#), [Athey et al., 2018](#), [Li et al., 2019](#), [Xu and Basse, 2021](#), [Wang et al., 2025](#), [Samii et al., 2023](#), [Gao and Ding, 2025](#)]. That work clarifies how causal estimands change when a unit’s outcome can depend on peers, assignments in a neighborhood, or a realized exposure condition. Here the exposure condition is the realized group itself: the potential outcome  $Y_{i,n}(z, A)$ , first introduced formally in [Assumption 4](#), depends on the arm  $z$  and the group  $A$ . The random-partition design then supplies a tractable law over group compositions, making the population average marginal effect  $\tau_n$  and its estimator  $\hat{\tau}_n$  design-based quantities.

Finally, the one-realization lower bound is related to finite-population variance nonidentification and conservative variance-estimation results [[Aronow et al., 2014](#), [Basse and Airolidi, 2018](#), [Harshaw et al., 2026](#), [Park and Wager, 2026](#)] and sits alongside design-based randomization and permutation inference for interference settings [[Basse et al., 2024](#), [Puelz et al., 2022](#)]. [Proposition 3](#) constructs two randomized schedule priors with identical one-realization observables and separated scaled variances, and [Theorem 6](#) turns that separation into a lower bound for arbitrary one-realization variance statistics. Together with the exact variance and CR2 results, this positions the paper as a finite-population account of what random group formation contributes to variance geometry, what the standard scalar software statistic targets in the equal-group design, and what information a single realized partition contains about the design variance.

### 3 Setup and assumptions

Throughout,  $n$  indexes the finite population size and  $M$  is the fixed common group size. Row-dependent quantities carry an  $n$  subscript, all limits are along the triangular array, and generic finite constants may change from line to line. Expectations, inner products, and norms below are finite-population objects taken over the uniform slice of possible groups.

Three row-level quantities recur in the conditions below and are fixed once here. The realized group count is  $G_n$ ; the number of grouped units is

$$N_n = MG_n,$$

the total number of population units that end up inside a realized group; and  $\sigma_n^2 = \text{Var}(\hat{\tau}_n)$  is the exact design variance of the group-level difference-in-means estimator of [Definition 7](#), taken over the joint law of the realized groups and the treatment assignment. The asymptotic

analysis rests on the number of realized groups and on the treatment, sampling, boundedness, and variance-scale restrictions imposed on the triangular array.

**Assumption 1.** `[ass:group-count-growth]` (Group-count growth) *The number of realized groups diverges:*

$$G_n \rightarrow \infty.$$

[Assumption 1](#) is the growing number of randomized groups condition, a standard route for design-based asymptotics with many randomized clusters [\[Fu et al., 2026\]](#). It is the dimension along which the random partition supplies repeated group-level observations.

**Definition 1.** `[synth_11]` (Row treatment fraction) *For a triangular schedule array with row group count  $G_n$  and treated-group count  $G_{1n}$ , where  $0 < G_{1n} < G_n$ , the row treatment fraction is*

$$p_n = \frac{G_{1n}}{G_n}.$$

The number of treated groups  $G_{1n}$  determines the row treatment fraction in [Definition 1](#); the number of control groups  $G_{0n}$  is the complementary count used in the difference-in-means estimator below.

**Assumption 2.** `[ass:stable-treatment-fraction]` (Stable treatment fraction) *The treated-group fraction  $p_n$  converges to an interior limit:*

$$p_n \rightarrow p \in (0, 1).$$

[Assumption 2](#) is the stable complete-randomization arm fractions condition [\[Li and Ding, 2017\]](#). The interior limiting treated-group fraction  $p$  keeps both treatment arms asymptotically represented in the randomized group assignment.

**Assumption 3.** `[ass:sampling-fraction]` (Sampling fraction limit) *The grouped-unit fraction  $N_n/n$  converges:*

$$\frac{N_n}{n} \rightarrow \rho \in [0, 1].$$

[Assumption 3](#) is the finite-population sampling-fraction limit condition [\[Ohlsson, 1989\]](#). It records the limiting density of grouped units in the population, allowing the same notation to cover sparse and dense grouping regimes.

**Assumption 4.** `[ass:bounded-potential-outcomes]` (Bounded potential outcomes) *There is a finite uniform outcome envelope  $B$  such that*

$$\sup_{n,i,z,A} |Y_{i,n}(z, A)| \leq B < \infty.$$

[Assumption 4](#) is the uniformly bounded composition potential outcomes condition [\[Fu et al., 2026\]](#). The deterministic potential outcome  $Y_{i,n}(z, A)$  may depend on the unit, arm, and realized group, while the uniform envelope  $B$  supplies a common finite bound across rows.

**Assumption 5.** `[ass:scaled-variance-non degeneracy]` (Scaled variance nondegeneracy) *For a feasible triangular schedule array with common group size  $M \geq 2$ , positive group counts  $G_n$ , treated-group counts satisfying  $0 < G_{1n} < G_n$ , grouped size  $MG_n \leq n$ , and deterministic potential-outcome schedules, there is a scaled-variance lower-bound constant  $c_\sigma$  such that  $c_\sigma > 0$  and*

$$c_\sigma \leq \liminf_{n \rightarrow \infty} G_n \sigma_n^2.$$

[Assumption 5](#) is the nondegenerate scaled design variance condition [Li and Ding, 2017]. The scaled-variance lower-bound constant  $c_\sigma$  fixes the normalization used by the asymptotic variance comparisons.

We next define the finite-population design and the function spaces on which its variance geometry is expressed.

**Definition 2.** [synth\_2] (Treatment-arm set) *The treatment-arm set is the two-point set*

$$\mathcal{Z} = \{0, 1\}.$$

The treatment-arm set  $\mathcal{Z}$  is used throughout for the two arms of the group-level potential-outcome tables.

**Definition 3.** [synth\_1] (Feasible schedule array) *A feasible triangular schedule array with common group size  $M$  consists of the following data:*

- **(Group size.)** *A natural number  $M$  satisfying  $2 \leq M$ .*
- **(Row sizes.)** *For each row  $r \in \mathbb{N}$ , a population size  $n_r \in \mathbb{N}$ , a number of groups  $G_r \in \mathbb{N}$ , and a number of treated groups  $G_{1r} \in \mathbb{N}$ .*
- **(Feasibility.)** *For every row  $r$ ,*

$$0 < G_r, \quad MG_r \leq n_r, \quad 0 < G_{1r}, \quad G_{1r} < G_r, \quad G_{1r} \leq G_r.$$

- **(Potential-outcome schedule.)** *For every row  $r$ , a deterministic potential-outcome schedule*

$$Y_r : \Omega_{n_r, M} \rightarrow \{i \in \{0, \dots, n_r - 1\} : i \in A\} \rightarrow \mathcal{Z} \rightarrow \mathbb{R},$$

*where  $A \in \Omega_{n_r, M}$  is the realized group argument of  $Y_r$ .*

[Definition 3](#) collects the rowwise feasibility restrictions used in the design: the group size is common, each row has positive groups and a nontrivial treated count, and the potential-outcome schedule is deterministic.

**Definition 4.** [def:random-partition-design] (Random partition design  $(\mathbf{A}_n, \mathbf{Z}_n)$ ) *Let  $\mathcal{U}_n$  be the finite population of  $n$  labeled units, and let*

$$\Omega_{n, M} = \{A \subseteq \mathcal{U}_n : |A| = M\}.$$

*The group tuple  $\mathbf{A}_n = (A_{1n}, \dots, A_{G_n n})$  is drawn uniformly from all ordered  $G_n$ -tuples of pairwise-disjoint elements of  $\Omega_{n, M}$ . Independently, the assignment vector  $\mathbf{Z}_n$  is drawn uniformly from all binary vectors with exactly  $G_{1n}$  treated groups.*

[Definition 4](#) defines the homogeneous random group-formation design. The ordered group tuple  $\mathbf{A}_n$  carves  $G_n$  pairwise-disjoint groups of common size  $M$  out of the finite population  $\mathcal{U}_n$ , and the balanced group-treatment assignment vector  $\mathbf{Z}_n$  is drawn independently conditional on the realized groups. The feasibility condition  $N_n = MG_n \leq n$  allows the grouped units to occupy any fraction of the population. The realized groups therefore form a disjoint packing of  $\mathcal{U}_n$ : they are pairwise disjoint and each has size  $M$ , and they cover the whole population exactly when  $N_n = n$ . The general case, in which  $n - N_n$  units lie outside every realized group and contribute no observation, is what gives the sampling fraction  $N_n/n$  its range and is where the sparse and intermediate-density regimes studied below live. The design's name, *random partition*, refers throughout to this packing mechanism; the saturated case  $N_n = n$  is the special case in which it is literally a partition.

**Definition 5.** [def:group-table] (Group-level arm table  $h_{z,n}$ ) For each treatment arm  $z \in \mathcal{Z}$  and each group  $A \in \Omega_{n,M}$ , define the group-level arm table by

$$h_{z,n}(A) = \frac{1}{M} \sum_{i \in A} Y_{i,n}(z, A).$$

The group-level arm table  $h_{z,n}$  converts unit-level composition potential outcomes into the group mean observed when group  $A$  receives arm  $z$ .

**Definition 6.** [synth\_3] (Uniform slice averaging) For integers  $M \leq n$ , let

$$\Omega_{n,M} := \{A \subseteq \{1, \dots, n\} : |A| = M\}.$$

The  $M$ -slice design is the complete-randomization design on  $\Omega_{n,M}$ , equivalently the uniform distribution over all  $M$ -element subsets of the  $n$ -unit population. For any real-valued table  $f : \Omega_{n,M} \rightarrow \mathbb{R}$ , write

$$\mathbb{E}_n f := \frac{1}{|\Omega_{n,M}|} \sum_{A \in \Omega_{n,M}} f(A).$$

The uniform  $M$ -slice  $\Omega_{n,M}$  is the population of possible groups, and the averaging operator  $\mathbb{E}_n$  turns any group table into its finite-population slice mean.

**Definition 7.** [def:pame-estimator] (PAME and estimator  $\tau_n, \hat{\tau}_n$ ) Define the finite-population partition-average marginal effect by

$$\tau_n = \mathbb{E}_n[h_{1,n} - h_{0,n}].$$

The difference-in-means estimator across treated and control groups is

$$\hat{\tau}_n = \frac{1}{G_{1n}} \sum_g Z_{ng} h_{1,n}(A_{ng}) - \frac{1}{G_{0n}} \sum_g (1 - Z_{ng}) h_{0,n}(A_{ng}).$$

[Definition 7](#) defines the finite-population partition-average marginal effect and its realized group-level difference-in-means estimator. The estimand averages the contrast of the two arm tables over the full slice, while the estimator uses the treatment labels and realized groups produced by [Definition 4](#).

**Definition 8.** [synth\_12] (Uniform-slice  $L^2$  space) For integers  $n$  and  $M$  with  $2M \leq n$ , let  $\mathcal{U}_n = \{1, \dots, n\}$  and

$$\Omega_{n,M} = \{A \subseteq \mathcal{U}_n : |A| = M\}$$

be the uniform  $M$ -slice. The space  $L^2(\Omega_{n,M})$  is the real vector space of functions  $f : \Omega_{n,M} \rightarrow \mathbb{R}$ , equipped with uniform-slice expectation

$$\mathbb{E}_n f = \binom{n}{M}^{-1} \sum_{A \in \Omega_{n,M}} f(A),$$

inner product

$$\langle f, g \rangle_n = \mathbb{E}_n[fg],$$

and induced norm  $\|f\|_n^2 = \langle f, f \rangle_n$ .

The inner product  $\langle \cdot, \cdot \rangle_n$  and induced norm  $\| \cdot \|_n$  in [Definition 8](#) provide the Hilbert-space notation used to decompose group tables by Johnson degree.

**Definition 9.** `[def:johnson-decomposition]` (Johnson orthogonal decomposition) *For integers  $n, M$  with  $2M \leq n$ , let*

$$\Omega_{n,M} := \{A \subseteq \{1, \dots, n\} : |A| = M\}$$

*be the uniform  $M$ -slice, let  $\mathbb{E}_n$  denote uniform averaging on  $\Omega_{n,M}$ , and let*

$$\langle f, g \rangle_n := \mathbb{E}_n[fg]$$

*be the uniform-slice inner product. For each  $k = 0, \dots, M$ , let  $\mathcal{H}_{n,k}$  be the canonical degree- $k$  Johnson harmonic subspace and let  $\Pi_{n,k}$  be the orthogonal projection onto  $\mathcal{H}_{n,k}$ :  $\Pi_{n,k}$  is linear, has range  $\mathcal{H}_{n,k}$ , fixes every element of  $\mathcal{H}_{n,k}$ , and satisfies*

$$\langle f - \Pi_{n,k}f, g \rangle_n = 0 \quad \text{for every } g \in \mathcal{H}_{n,k}.$$

*For every  $f \in L^2(\Omega_{n,M})$ , define its Johnson components by*

$$f_k := \Pi_{n,k}f, \quad k = 0, \dots, M.$$

*These components satisfy the centered decomposition*

$$\sum_{k=1}^M \Pi_{n,k}f(A) = f(A) - \mathbb{E}_n f \quad \text{for every } A \in \Omega_{n,M},$$

*and distinct degrees are orthogonal:*

$$\langle \Pi_{n,k}f, \Pi_{n,\ell}g \rangle_n = 0 \quad \text{whenever } k \neq \ell.$$

The Johnson harmonic subspace  $\mathcal{H}_{n,k}$  and projection  $\Pi_{n,k}$  organize a centered group table into orthogonal degrees. This is the representation-theoretic coordinate system for the exact variance identities below, following the Johnson and Kneser association-scheme perspective [[Filmus, 2016](#), [Brouwer et al., 2018](#)].

Two derived slice quantities enter the next two definitions. For each arm  $z$ , the centered arm table is

$$h_{z,n}^\circ = h_{z,n} - \mathbb{E}_n h_{z,n},$$

the group-level arm table of [Definition 5](#) with its slice mean removed, as recorded in [Definition 15](#); and the uniform-slice arm variance is its energy,

$$V_{z,n} = \|h_{z,n}^\circ\|_n^2 = \mathbb{E}_n[(h_{z,n} - \mathbb{E}_n h_{z,n})^2],$$

the variance of the arm- $z$  group table under a single uniform draw from the slice. These are the only two objects needed to read the covariance and correction definitions that follow.

**Definition 10.** `[def:kneser-covariance]` (Kneser covariance  $C_{ab,n}$ ) *For a function  $f$  on  $\Omega_{n,M}$ , define the normalized Kneser disjointness operator by*

$$(\mathbb{K}_n f)(A) = \binom{n-M}{M}^{-1} \sum_{C: C \cap A = \emptyset} f(C).$$

*For arm tables  $h_{a,n}$  and  $h_{b,n}$ , define the ordered-disjoint covariance by*

$$C_{ab,n} = \langle h_{a,n}^\circ, \mathbb{K}_n h_{b,n}^\circ \rangle_n.$$

**Definition 10** records the covariance generated by sampling disjoint groups. The normalized disjointness operator  $K_n$  averages over groups disjoint from a given group, while the centered arm table  $h_{z,n}^\circ$  and ordered-disjoint covariance  $C_{ab,n}$  isolate the finite-population dependence induced by the partition.

**Definition 11.** [def:dense-correction] (Dense correction terms  $E_{\tau,1,n}, R_n$ ) Define the degree-one Johnson energy of the arm-difference table by

$$E_{\tau,1,n} = \|\Pi_{n,1}(h_{1,n} - h_{0,n})\|_n^2.$$

Define the independent-group leading variance scale by

$$R_n = \frac{V_{1,n}}{p_n} + \frac{V_{0,n}}{1 - p_n}.$$

The degree-one energy  $E_{\tau,1,n}$  is the part of the arm-difference table carried by the linear Johnson component. The uniform-slice arm variance  $V_{z,n}$  enters [Definition 11](#) through the independent-group leading variance scale.

**Definition 12.** [def:cr2-statistic] (CR2 statistic  $\hat{V}_{\text{CR2},n}$ ) For each arm  $z$ , let  $s_{z,n}^2$  be the within-arm sample variance of the observed group means, computed with denominator  $G_{zn} - 1$ . Define the scalar equal-group CR2 variance statistic by

$$\hat{V}_{\text{CR2},n} = \frac{s_{1,n}^2}{G_{1n}} + \frac{s_{0,n}^2}{G_{0n}}.$$

The within-arm sample variance  $s_{z,n}^2$  in [Definition 12](#) is computed from observed group means within each assigned arm. The scalar CR2 statistic  $\hat{V}_{\text{CR2},n}$  is therefore the equal-group cluster-robust variance quantity analyzed in the later software identity and asymptotic results.

The class defined next is named for the dense regime that motivates it, and its parameter  $\rho$  ranges over the whole unit interval: arrays with  $\rho = 0$  belong to it, and the results below treat  $\rho = 0$  and  $\rho > 0$  as the two phases of one statement.

**Definition 13.** [def:dense-schedule-class] (Dense schedule class) The class  $\mathcal{C}_{\text{dense}}$  consists of triangular arrays  $(Y_{i,n}, G_n, G_{1n})_{n \geq 1}$ , with fixed common group size  $M$ , for which:

- **(Feasible rows.)** The common group size satisfies  $M \geq 2$ . In each row,  $G_n > 0$ ,  $N_n = MG_n \leq n$ ,  $0 < G_{1n} < G_n$ , and  $G_{1n} \leq G_n$ . The potential-outcome schedule assigns a real value  $Y_{i,n}(z, A)$  for every group  $A \in \Omega_{n,M}$ , every unit  $i \in A$ , and every arm  $z \in \mathcal{Z}$ .
- **(Dense growth.)** The group count satisfies [Assumption 1](#), so  $G_n \rightarrow \infty$ .
- **(Treatment fraction.)** The treated-group fraction satisfies [Assumption 2](#): for an interior limiting treated-group fraction  $p$ ,  $0 < p < 1$  and  $p_n \rightarrow p$ .
- **(Sampling fraction.)** The grouped-unit sampling fraction satisfies [Assumption 3](#): for a limiting grouped-unit sampling fraction  $\rho$ ,  $0 \leq \rho \leq 1$  and  $N_n/n \rightarrow \rho$ .
- **(Uniform envelope.)** The potential outcomes satisfy [Assumption 4](#): for a uniform envelope  $B > 0$ ,

$$\sup_{n, A \in \Omega_{n,M}, i \in A, z \in \mathcal{Z}} |Y_{i,n}(z, A)| \leq B.$$

- **(Scaled variance.)** The exact design variance satisfies [Assumption 5](#): for a scaled-variance lower-bound constant  $c_\sigma > 0$ ,

$$c_\sigma \leq \liminf_{n \rightarrow \infty} G_n \sigma_n^2.$$

Equivalently,

$$\mathcal{C}_{\text{dense}} = \left\{ (Y_{i,n}, G_n, G_{1n})_{n \geq 1} : G_n \rightarrow \infty, p_n \rightarrow p, N_n/n \rightarrow \rho, \sup_{n, A \in \Omega_{n,M}, i \in A, z \in \mathcal{Z}} |Y_{i,n}(z, A)| \leq B, \liminf_{n \rightarrow \infty} G_n \sigma_n^2 \geq c_\sigma \right\}$$

with the row feasibility, interiority, closed-interval, envelope, and variance-floor conditions listed above.

The dense schedule class  $\mathcal{C}_{\text{dense}}$  packages the conditions in [Assumptions 1](#) to [5](#). It is the common domain for the dense asymptotic statements, combining feasible rows, a stable treatment split, a limiting grouped-unit fraction, bounded potential outcomes, and a positive scaled design-variance floor.

## 4 Exact variance geometry

The setup in [Definitions 4](#), [5](#) and [7](#) makes the realized groups a sample of disjoint  $M$ -subsets. The variance calculation therefore depends on two finite-population operations: centering each arm table on the uniform  $M$ -slice and averaging a table over groups disjoint from a given group. This section records the exact algebra for those operations and then applies it to the design variance of  $\hat{\tau}_n$ , the group-level difference-in-means estimator.

We first fix the realized arm counts and the centered table notation used in the covariance formulas.

**Definition 14.** [\[synth\\_9\]](#) (Realized arm group counts) For  $z \in \mathcal{Z} = \{0, 1\}$ , let

$$G_{z,n} = \sum_{g=1}^{G_n} \mathbf{1}\{Z_{ng} = z\}$$

be the number of realized groups assigned to arm  $z$  at row  $n$ . Under the complete group-treatment assignment with exactly  $G_{1n}$  treated groups and  $G_{0n} = G_n - G_{1n}$  control groups, this notation gives  $G_{1,n} = G_{1n}$  and  $G_{0,n} = G_{0n}$ .

[Definition 14](#) aligns the random-assignment notation with the fixed arm counts from [Definition 4](#). In the balanced complete group assignment used throughout the paper, the realized counts equal the design counts, so later variance expressions may use  $G_{1n}$  and  $G_{0n}$  directly.

**Definition 15.** [\[synth\\_5\]](#) (Centered arm table) For integers  $n$  and  $M$  with  $M \leq n$ , let

$$\Omega_{n,M} = \{A \subseteq \{1, \dots, n\} : |A| = M\}.$$

Given a deterministic potential-outcome schedule  $Y_{i,n}(z, A)$  for units  $i \in A$ , arms  $z \in \mathcal{Z}$ , and groups  $A \in \Omega_{n,M}$ , define the centered arm- $z$  group table by

$$h_{z,n}^\circ(A) = h_{z,n}(A) - \mathbb{E}_n[h_{z,n}]$$

for each  $A \in \Omega_{n,M}$ , where

$$h_{z,n}(A) = \frac{1}{M} \sum_{i \in A} Y_{i,n}(z, A), \quad \mathbb{E}_n[h_{z,n}] = \frac{1}{|\Omega_{n,M}|} \sum_{B \in \Omega_{n,M}} h_{z,n}(B).$$

The centered table in [Definition 15](#) isolates the part of an arm-specific group mean that can covary across disjoint realized groups. The centering is with respect to the uniform slice, matching the estimand in [Definition 7](#) and the ordered-disjoint covariance in [Definition 10](#).

**Definition 16.** `[synth_4]` (Uniform-slice norm) *For integers  $M \leq n$ , let*

$$\Omega_{n,M} := \{A \subseteq \{1, \dots, n\} : |A| = M\}.$$

*For any real-valued function  $f : \Omega_{n,M} \rightarrow \mathbb{R}$ , define its squared uniform-slice norm by*

$$\|f\|_n^2 := \langle f, f \rangle_n := \mathbb{E}_n\{f(A)^2\},$$

*where  $\mathbb{E}_n$  denotes averaging over the uniform law on  $\Omega_{n,M}$ .*

[Definition 16](#) supplies the norm in which Johnson harmonic components are measured. Because the inner product averages over all candidate groups of size  $M$ , these energies describe finite-population variation before the random partition selects the realized disjoint groups.

The next result gives the spectral identity that turns disjointness dependence into degree-by-degree covariance terms. Its constants are the normalized Kneser eigenvalues, obtained from the classical Johnson association-scheme spectrum [\[Delsarte, 1973, Brouwer et al., 1989, 2018\]](#); the harmonic decomposition follows the standard slice analysis of Boolean functions [\[Filmus, 2016\]](#).

**Theorem 1.** `[thm:exact-kneser-identity]` (Exact Kneser covariance identity) *Let  $M, n \in \mathbb{N}$  satisfy  $2 \leq M$  and  $2M \leq n$ . Let  $Y_{i,n}(z, A)$  be a deterministic potential-outcome schedule for units  $i \in A$ , arms  $z \in \mathcal{Z}$ , and groups  $A \in \Omega_{n,M}$ , with group-level arm tables  $h_{z,n}$  as in [Definition 5](#). Suppose the following conditions hold.*

- **(Johnson projections.)** *For each  $k = 0, \dots, M$ ,  $\Pi_{n,k}$  is a linear projection on functions  $f : \Omega_{n,M} \rightarrow \mathbb{R}$  whose range is the degree- $k$  Johnson harmonic subspace  $\mathcal{H}_{n,k}$ , which fixes every element of  $\mathcal{H}_{n,k}$ , and whose residual  $f - \Pi_{n,k}f$  is orthogonal to  $\mathcal{H}_{n,k}$  under the uniform-slice inner product  $\langle u, v \rangle_n = \mathbb{E}_n[u(A)v(A)]$ .*
- **(Centered orthogonal decomposition.)** *For every  $f : \Omega_{n,M} \rightarrow \mathbb{R}$  and  $A \in \Omega_{n,M}$ ,*

$$\sum_{k=1}^M \Pi_{n,k}f(A) = f(A) - \mathbb{E}_nf(A),$$

*and the projected components are mutually orthogonal: if  $k \neq \ell$ , then*

$$\langle \Pi_{n,k}f, \Pi_{n,\ell}g \rangle_n = 0$$

*for all  $f, g : \Omega_{n,M} \rightarrow \mathbb{R}$ .*

- **(Kneser spectrum.)** For every  $k = 0, \dots, M$  and every  $f \in \mathcal{H}_{n,k}$ , the unnormalized Kneser adjacency operator has eigenvalue

$$(-1)^k \binom{n-M-k}{M-k}.$$

Then the normalized Kneser disjointness operator  $K_n$  acts on each projected component by

$$K_n(\Pi_{n,k}f)(A) = \lambda_{n,k} \Pi_{n,k}f(A), \quad \lambda_{n,k} = (-1)^k \frac{(M)_k}{(n-M)_k},$$

for every  $k = 0, \dots, M$ , every  $f : \Omega_{n,M} \rightarrow \mathbb{R}$ , and every  $A \in \Omega_{n,M}$ . Moreover, if  $f, g : \Omega_{n,M} \rightarrow \mathbb{R}$  are centered under the uniform  $M$ -slice, then for a uniform ordered disjoint pair  $(A, C)$ ,

$$\text{Cov}\{f(A), g(C)\} = \sum_{k=1}^M \lambda_{n,k} \langle \Pi_{n,k}f, \Pi_{n,k}g \rangle_n.$$

Finally, the ordered-disjoint covariance of the arm-difference table satisfies

$$C_{\tau\tau,n} = \sum_{k=1}^M \lambda_{n,k} \|\Pi_{n,k}(h_{1,n} - h_{0,n})\|_n^2.$$

[Theorem 1](#) shows that disjoint group formation has an exact finite-sample diagonalization in Johnson degrees. Each degree contributes its harmonic inner product multiplied by the signed finite- $n$  factor  $\lambda_{n,k}$ , so the covariance retained by sampling groups without replacement is completely summarized by the slice energies of the projected arm tables. Applied to the arm-difference table, the same identity gives  $C_{\tau\tau,n}$ , the ordered-disjoint covariance contrast that enters the variance comparison below.

The exact variance formula combines this disjointness covariance with the complete-randomization variation across treatment labels. The definitions in [Definitions 14](#) and [15](#) make the statement entirely finite-sample: the arm variances  $V_{z,n}$ , the ordered-disjoint covariances  $C_{ab,n}$ , and the independent-group scale  $R_n$  are all computed from the deterministic row schedule.

**Theorem 2.** [\[thm:exact-pame-variance\]](#) (Exact PAME variance) *Let  $n, M, G_n, G_{1n}$  be natural numbers and let  $G_{0n} = G_n - G_{1n}$ . Suppose that:*

- **(Feasible partition.)** *The two-stage random partition design in [Definition 4](#) is feasible:  $MG_n \leq n$ .*
- **(Group sizes.)**  *$2 \leq M$ ,  $2 \leq G_{1n}$ , and  $2 \leq G_{0n}$ .*
- **(Potential outcomes.)**  *$Y_{i,n}(z, A)$  is a deterministic potential-outcome schedule, defined for every  $A \in \Omega_{n,M}$ , every  $i \in A$ , and every  $z \in \mathcal{Z}$ .*

Set

$$V_{z,n} = \text{Var}_{A \sim \text{Unif}(\Omega_{n,M})} \{h_{z,n}(A)\}, \quad C_{ab,n} = \mathbb{E}_A \mathbb{E}_{C|C \cap A = \emptyset} [h_{a,n}^\circ(A) h_{b,n}^\circ(C)],$$

and

$$C_{\tau\tau,n} = C_{11,n} + C_{00,n} - 2C_{10,n}, \quad R_n = \frac{V_{1,n}}{G_{1n}/G_n} + \frac{V_{0,n}}{1 - G_{1n}/G_n}.$$

For the difference-in-means estimator  $\hat{\tau}_n$ , the finite-population PAME  $\tau_n$ , the exact design variance  $\sigma_n^2 = \text{Var}(\hat{\tau}_n)$ , and the scalar equal-group CR2 statistic  $\hat{V}_{\text{CR2},n}$  in [Definition 12](#),

$$\begin{aligned}\mathbb{E}[\hat{\tau}_n] &= \tau_n, \\ \sigma_n^2 &= \frac{R_n}{G_n} + C_{\tau\tau,n} - \frac{C_{11,n}}{G_{1n}} - \frac{C_{00,n}}{G_{0n}}, \\ \mathbb{E}[\hat{V}_{\text{CR2},n}] &= \frac{V_{1,n} - C_{11,n}}{G_{1n}} + \frac{V_{0,n} - C_{00,n}}{G_{0n}},\end{aligned}$$

and

$$\mathbb{E}[\hat{V}_{\text{CR2},n}] - \sigma_n^2 = -C_{\tau\tau,n}.$$

[Theorem 2](#) establishes three finite-sample facts for the two-stage design. First, the group difference in means is unbiased for the finite-population PAME. Second, the exact design variance separates an independent-group component,  $R_n/G_n$ , from the ordered-disjoint covariance terms induced by the random partition. Third, the scalar equal-group CR2 expectation differs from the exact variance by exactly  $-C_{\tau\tau,n}$ . Thus the sign and magnitude of the CR2 discrepancy are governed by the same Kneser covariance contrast characterized in [Theorem 1](#).

## 5 Dense and sparse variance behavior

The exact identities in [Theorems 1](#) and [2](#) reduce the large-population question to the behavior of Johnson energies under random grouping. In dense arrays, where a positive fraction of the finite population may be grouped, the first Johnson degree carries the leading correction to the independent-group variance scale. The higher degrees remain present in the finite-sample identity, but their aggregate contribution is uniformly of smaller order under the bounded-outcome and fixed-group-size conditions.

A word on how to read the statements below. Each theorem carries its Johnson-projection, orthogonal-decomposition, and Kneser-spectrum clauses explicitly, because the verified statements quantify over the projections rather than fixing them. Those three clauses are not extra economic assumptions: they hold automatically for the canonical Johnson projections on the uniform  $M$ -slice whenever  $2M \leq n$ , by [Lemmas 2](#) and [3](#), which are proved in [Section A](#). The substantive conditions are only the four array conditions of [Section 3](#): growing group count, stable treatment fraction, converging sampling fraction, and bounded potential outcomes. A reader may take the projection clauses as given and read each theorem as a statement about arrays satisfying those four conditions alone.

**Theorem 3.** `[thm:dense-projection-limit]` (Dense projection limit) *Fix a common group size  $M \geq 2$  and real numbers  $p$ ,  $\rho$ , and  $B$ . There is a constant  $C > 0$  such that, for every feasible triangular schedule array with common group size  $M$  and every rowwise choice of Johnson projections  $\Pi_{n,k}$  on  $L^2(\Omega_{n,M})$ , the following conditions imply the displayed limits below:*

- **(Group growth.)** *The array satisfies [Assumption 1](#).*
- **(Treatment fractions.)** *The array satisfies [Assumption 2](#) with limiting treated-group fraction  $p$ .*
- **(Sampling fractions.)** *The array satisfies [Assumption 3](#) with limiting grouped-unit fraction  $\rho$ .*

- **(Bounded outcomes.)** The array satisfies [Assumption 4](#) with uniform outcome envelope  $B$ .
- **(Johnson decomposition.)** In every row with  $2M \leq n$ , the projections give the centered orthogonal Johnson decomposition

$$\sum_{k=1}^M \Pi_{n,k} f(A) = f(A) - \mathbb{E}_n f$$

for every  $f$  and every  $A \in \Omega_{n,M}$ , and distinct Johnson degrees are orthogonal under  $\langle \cdot, \cdot \rangle_n$ .

- **(Kneser spectrum.)** In every row with  $2M \leq n$ , every  $f \in \mathcal{H}_{n,k}$  is an eigenfunction of the unnormalized Kneser adjacency operator with eigenvalue

$$(-1)^k \binom{n-M-k}{M-k}.$$

Then

$$\begin{aligned} G_n \sigma_n^2 - (R_n - \rho E_{\tau,1,n}) &\rightarrow 0, \\ G_n C_{\tau\tau,n} + \rho E_{\tau,1,n} &\rightarrow 0, \end{aligned}$$

and

$$\frac{C_{11,n}}{p_n} + \frac{C_{00,n}}{1-p_n} \rightarrow 0.$$

Moreover, for every row  $n$ , the pooled contribution of Johnson degrees  $k \geq 2$ ,

$$G_n \sum_{k=2}^M \lambda_{n,k} \|\Pi_{n,k}(h_{1,n} - h_{0,n})\|_n^2,$$

has absolute value at most  $C/n$ .

The theorem identifies the dense correction as the product of the grouped-unit sampling limit and the degree-one contrast energy  $E_{\tau,1,n}$ . Intuitively, degree one records the part of the arm-difference table that is visible through the marginal participation of individual population members in sampled groups. Sampling without replacement induces competition for those members across groups, and the limit in [Theorem 3](#) converts that competition into the correction  $-\rho E_{\tau,1,n}$  to the scaled exact variance. The final bound gives the corresponding uniform remainder statement for Johnson degrees at least two, so the first-order dense comparison is governed by the degree-one projection.

This projection expansion feeds directly into the behavior of the equal-group CR2 statistic. The statistic  $\hat{V}_{\text{CR2},n}$  is the usual scalar cluster-robust variance estimator for the difference in group means in this balanced design, matching the equal-group CR2 form used in finite-sample cluster-robust inference [[Pustejovsky and Tipton, 2018](#)]. The next result states its probability limit and the exact criterion under which the conservative dense correction vanishes relative to the target variance.

**Theorem 4.** [\[thm:cr2-phase-frontier\]](#) (CR2 phase frontier) *Let  $M \geq 2$ , and let  $A$  be a triangular array of feasible random-group experiment rows with population sizes  $n$ , group counts  $G_n > 0$ , treated-group counts  $0 < G_{1n} < G_n$ , grouped-unit counts  $N_n = MG_n \leq n$ , and deterministic potential-outcome schedules. Fix constants  $p, \rho, B, c_\sigma \in \mathbb{R}$ . Suppose that:*

- **(Dense array.)** The array  $A$  belongs to the dense bounded schedule class  $\mathcal{C}_{\text{dense}}$  in [Definition 13](#), with parameters  $p, \rho, B, c_\sigma$ .
- **(Johnson projections.)** For every row  $n$ , the spaces  $L^2(\Omega_{n,M})$  are equipped with linear projections  $\Pi_{n,k}$ ,  $0 \leq k \leq M$ , whose ranges are the degree- $k$  Johnson harmonic subspaces  $\mathcal{H}_{n,k}$ , which fix functions in their ranges and have residuals orthogonal to  $\mathcal{H}_{n,k}$  under  $\langle \cdot, \cdot \rangle_n$ .
- **(Orthogonal decomposition.)** For every row with  $2M \leq n$ , every  $f \in L^2(\Omega_{n,M})$ , and every  $A \in \Omega_{n,M}$ ,

$$\sum_{k=1}^M \Pi_{n,k} f(A) = f(A) - \mathbb{E}_n f,$$

and distinct Johnson components are orthogonal:

$$\langle \Pi_{n,k} f, \Pi_{n,\ell} g \rangle_n = 0 \quad \text{whenever } k \neq \ell.$$

- **(Kneser spectrum.)** For every row with  $2M \leq n$ , every  $0 \leq k \leq M$ , and every  $f \in \mathcal{H}_{n,k}$ , the unnormalized Kneser adjacency operator satisfies

$$\sum_{C \cap A = \emptyset} f(C) = (-1)^k \binom{n-M-k}{M-k} f(A).$$

Then, as  $n \rightarrow \infty$ ,

$$G_n \widehat{V}_{\text{CR2},n} - R_n \xrightarrow{p} 0$$

and

$$G_n (\widehat{V}_{\text{CR2},n} - \sigma_n^2) - \rho E_{\tau,1,n} \xrightarrow{p} 0.$$

Moreover, for every  $\varepsilon > 0$ ,

$$\Pr \left\{ \widehat{V}_{\text{CR2},n} / \sigma_n^2 < 1 - \varepsilon \right\} \longrightarrow 0,$$

and the ratio  $\widehat{V}_{\text{CR2},n} / \sigma_n^2$  converges in probability to 1 if and only if

$$\frac{\rho E_{\tau,1,n}}{R_n - \rho E_{\tau,1,n}} \longrightarrow 0.$$

[Theorem 4](#) says that CR2 estimates the independent-group scale  $R_n/G_n$  in probability, while the exact design variance subtracts the dense projection correction. Consequently, the statistic is asymptotically conservative for the exact variance in the one-sided probability sense stated in the theorem. Its ratio consistency is characterized by a single relative-energy condition: the dense correction  $\rho E_{\tau,1,n}$  must be negligible compared with the corrected scaled variance  $R_n - \rho E_{\tau,1,n}$ . This gives a diagnostic interpretation to the Johnson decomposition in [Definition 9](#): degree-one heterogeneity in the arm contrast is precisely the component that determines whether the usual CR2 standard error is sharp in dense random-group designs.

The sparse case follows by setting the grouped-unit sampling fraction to zero. In that regime, the same theorem conditions allow many realized groups while making the first-order finite-pool correction vanish at the scale relevant for the variance ratio. The final result records this implication and isolates the benchmark count sequence in which  $N_n/n \rightarrow 0$  even though  $N_n^2/n \rightarrow \infty$ .

**Theorem 5.** [thm:sparse-beyond-birthday] (Sparse birthday consistency) *Fix an integer  $M \geq 2$ . Then the following statements hold.*

- **(Sparse dense arrays.)** *For every feasible triangular schedule array with common group size  $M$ , if the array belongs to  $\mathcal{C}_{\text{dense}}$  in Definition 13 with limiting grouped-unit fraction  $\rho = 0$ , then*

$$\frac{\widehat{V}_{\text{CR2},n}}{\sigma_n^2} \xrightarrow{p} 1.$$

- **(Birthday counts.)** *With*

$$N_n^{\text{bd}} = M \left\lfloor \frac{n^{3/4}}{M} \right\rfloor,$$

*the benchmark counts satisfy  $N_n^{\text{bd}} \leq n$  for all sufficiently large  $n$ ,*

$$\frac{N_n^{\text{bd}}}{n} \rightarrow 0, \quad \frac{(N_n^{\text{bd}})^2}{n} \rightarrow \infty.$$

- **(Birthday-aligned arrays.)** *For every feasible triangular schedule array with common group size  $M$ , if its grouped-unit count satisfies*

$$N_n = M \left\lfloor \frac{n^{3/4}}{M} \right\rfloor$$

*row by row and the array satisfies Assumptions 1, 2, 4 and 5 for some  $p$ ,  $B$ , and  $c_\sigma$ , then*

$$\frac{\widehat{V}_{\text{CR2},n}}{\sigma_n^2} \xrightarrow{p} 1.$$

The sparse result separates the variance question from whole-design collision heuristics. What matters for the scalar PAME variance is the grouped-unit fraction  $N_n/n$ , and Theorem 5 establishes ratio consistency whenever that fraction converges to zero within the stated dense schedule class. The explicit sequence  $N_n = M \lfloor n^{3/4}/M \rfloor$  gives a concrete regime with vanishing sampling fraction and diverging  $N_n^2/n$ . Thus CR2 remains ratio-consistent for the exact design variance in a range where the experiment still uses many overlapping opportunities in the underlying finite population.

## 6 Software identity and finite witness

The preceding results analyze the scalar equal-group CR2 statistic in Definition 12. This section records how that scalar statistic is obtained from a standard software call in the equal-group design, then gives a fully finite calculation that fixes the sign and magnitude of the dense correction. The software statement is deliberately versioned: the relevant covariance entry is the treatment-coordinate entry returned by `clubSandwich` for an unweighted intercept-plus-treatment regression using the `sandwich` bread convention [Pustejovsky, 2026, Zeileis and Lumley, 2026, Zeileis, 2004, 2006, Zeileis et al., 2020, Pustejovsky and Tipton, 2018].

**Definition 17.** [def:eight-unit-witness] (Eight-unit witness  $\mathbf{a}_8$ ) *The eight-unit witness fixes*

$$n = N_n = 8, \quad M = 2, \quad G_n = 4, \quad G_{1n} = 2,$$

and uses the balanced sign vector

$$\mathbf{a}_8 = (a_1, \dots, a_8).$$

For every group  $A = \{j, k\}$  and every unit  $i \in A$ , the potential outcomes are

$$Y_{i,n}(0, A) = 0, \quad Y_{i,n}(1, A) = \frac{a_j + a_k}{2}.$$

**Proposition 1** identifies the population-design statistic with the coefficient variance delivered by the stated `vcovCR` call. The row-identification and outcome clauses align the regression rows with the realized groups in [Definition 4](#); the design-matrix and unweighted-fit clauses make the treatment coefficient equal to the group-level treated-control contrast in [Definition 7](#). The block-algebra and bread-scaling clauses then pin down the CR2 adjustment and the normalization used by the versioned software implementation.

The sign vector in [Definition 17](#) is balanced in the sense that half of its entries equal +1 and half equal -1; the verified witness takes the canonical representative

$$a_i = +1 \ (1 \leq i \leq 4), \quad a_i = -1 \ (5 \leq i \leq 8),$$

so that  $\sum_{i=1}^8 a_i = 0$ . The numbers reported below are computed for this representative. The finite calculation below specializes the design to eight units, four pairs, and two treated groups. It provides a small balanced schedule in which the exact variance components and the expected CR2 statistic can be evaluated without asymptotics.

**Proposition 1.** [`prop:clubswatch-consumer`] (CR2 software identity)<sup>1</sup> Suppose  $M \geq 2$ ,  $G_{1n} \geq 2$ , and  $G_{0n} \geq 2$ . Fix a deterministic potential-outcome schedule  $Y_{i,n}(z, A)$ , an ordered tuple  $\mathbf{A}_n$  of pairwise-disjoint realized groups of common size  $M$ , and a balanced group-treatment assignment vector  $\mathbf{Z}_n$  with  $G_{1n}$  treated groups and  $G_{0n} = G_n - G_{1n}$  control groups. Assume the software regression data satisfy the following conditions.

- **(Row identification.)** For each realized group, the  $M$  regression rows are in bijection with the units in that group.
- **(Design matrix.)** The fitted model matrix has exactly two columns, an intercept and the numeric binary group-treatment regressor:

$$X_{gij} = \begin{cases} 1, & j = 0, \\ 1\{Z_g = 1\}, & j = 1. \end{cases}$$

- **(Observed outcomes.)** The regression outcome in row  $i$  of group  $g$  is the realized potential outcome  $Y_{i,n}(Z_g, A_g)$ , and its within-group average equals the observed group mean.
- **(Unweighted fit.)** The fit is an unweighted full-column-rank least-squares fit, with residuals  $e_{gi}$ , and the average residual in each group equals the observed group mean minus the observed mean in that group's treatment arm.
- **(Block algebra.)** The stacked cross-product satisfies  $X^\top X = \sum_g X_g^\top X_g$ , the group hat matrices are

$$H_g = X_g(X^\top X)^{-1}X_g^\top,$$

each  $I - H_g$  is positive definite, and each CR2 adjustment  $A_g$  is symmetric positive definite with

$$A_g A_g (I - H_g) = I.$$

- **(Bread scaling.)** The matrix used by the CR2 covariance calculation is  $\tilde{B} = (G_n M)^{-1} B$ , where  $B$  is the unweighted `sandwich::bread(fit)` matrix.

Let  $\Sigma$  be the covariance matrix returned by `clubSandwich::vcovCR(fit, cluster = group, type = 'CR2')` in version 0.7.0, with `target`, `inverse_var`, and `form` omitted, applied to this unweighted intercept-plus-treatment fit using `sandwich` version 3.1-3 for the bread. Then the treatment-coordinate entry equals the scalar equal-group CR2 statistic in [Definition 12](#):

$$\Sigma_{11} = \hat{V}_{\text{CR2},n} = \frac{s_{1,n}^2}{G_{1n}} + \frac{s_{0,n}^2}{G_{0n}}.$$

The witness in [Definition 17](#) makes treatment outcomes depend only on the average sign within the realized pair, while control outcomes are identically zero. Balance of `a8` gives a concrete degree-one component in the treated arm table, so the exact finite-pool correction in [Theorem 2](#) has a visible numerical effect.

**Proposition 2.** [`prop:eight-unit-witness`] (Eight-unit witness moments) *For the balanced eight-unit witness schedule with  $n = N_n = 8$ ,  $M = 2$ ,  $G_n = 4$ , and  $G_{1n} = 2$ , suppose that:*

- **(Johnson projections.)** *There is a system of orthogonal projections  $\{\Pi_{8,k}\}_{k=0}^2$  onto the canonical Johnson harmonic spaces  $\mathcal{H}_{8,k}$  in  $L^2(\Omega_{8,2})$ .*
- **(Centered decomposition.)** *These projections decompose every  $f \in L^2(\Omega_{8,2})$  after removal of its uniform-slice mean, with different harmonic degrees orthogonal under  $\langle \cdot, \cdot \rangle_n$ .*
- **(Kneser spectrum.)** *The unnormalized Kneser adjacency operator acts on each  $\mathcal{H}_{8,k}$  by the eigenvalue*

$$(-1)^k \binom{8-2-k}{2-k}.$$

*Then the witness satisfies*

$$V_{1,n} = \frac{3}{7}, \quad E_{\tau,1,n} = \frac{3}{7}, \quad C_{11,n} = -\frac{1}{7}, \quad \sigma_n^2 = \frac{1}{7},$$

*and, for the random partition design in [Definition 4](#) and the scalar equal-group CR2 statistic in [Definition 12](#),*

$$\mathbb{E}[\hat{V}_{\text{CR2},n}] = \frac{2}{7}.$$

[Proposition 2](#) shows that the finite design variance is  $1/7$ , while the expected equal-group CR2 statistic is  $2/7$ . The same calculation reports  $V_{1,n} = E_{\tau,1,n} = 3/7$  and  $C_{11,n} = -1/7$ , matching the geometry in [Theorem 1](#): the treated table's degree-one energy generates a negative ordered-disjoint covariance, and that covariance enters the exact variance through the dense correction. Thus the witness supplies a concrete numerical check on both the Kneser covariance normalization and the software-facing CR2 statistic before the one-realization lower bound studied next.

---

<sup>1</sup>**Formalization scope.** This result uses the published conclusion [[Pustejovsky, 2026](#)] (Version 0.7.0 source package: R/lm.R lines 47–52 and 70–72; R/S3-methods.R lines 21–33 and 63–65; R/clubSandwich.R lines 167–175, 216–225, 237–287; R/CR-adjustments.R lines 5–7 and 22–42) and [[Zeileis and Lumley, 2026](#)] (Version 3.1-3 source package, R/bread.R lines 10–15). The cited source proof is not formalized here: Lean verifies its use and all remaining steps. If that cited conclusion is false, the portions of this result that depend on it are not certified.

## 7 One-realization lower bound

The finite witness in [Proposition 2](#) shows how the dense correction can move the exact variance away from the equal-group CR2 target in a small design. This section turns that calculation into a positive-density impossibility statement for one-realization variance recovery. The construction compares two randomized schedule priors whose induced observed-data mixtures agree for every statistic of the realized groups, assignments, and observed outcomes, while their scaled exact variances concentrate at separated limits.

We first record the observed object and the two sign priors used in the comparison.

**Definition 18.** [\[synth\\_8\]](#) (One-realization observation) *A one-realization observation  $\mathcal{O}_n$  at population size  $n$ , common group size  $M$ , realized group count  $G_n$ , and treated-group count  $G_{1n}$  is a triple*

$$\mathcal{O}_n = (\mathbf{A}_n, \mathbf{Z}_n, Y_n^{\text{obs}})$$

*with the following components:*

- **(Partition.)**  $\mathbf{A}_n = (A_{g,n})_{g=1}^{G_n}$  is an ordered tuple with  $A_{g,n} \in \Omega_{n,M}$  for every  $g$ , and  $A_{g,n} \cap A_{g',n} = \emptyset$  whenever  $g \neq g'$ .
- **(Treatment allocation.)**  $\mathbf{Z}_n$  selects a subset of the realized groups with exactly  $G_{1n}$  elements.
- **(Observed outcomes.)** For every realized group  $g$  and every unit  $i \in A_{g,n}$ ,  $Y_n^{\text{obs}}$  assigns a real observed outcome  $Y_{i,g,n}^{\text{obs}} \in \mathbb{R}$ .

[Definition 18](#) fixes the sigma-field available to a one-realization procedure: the sampled partition, the group treatment allocation, and the observed unit outcomes within sampled groups. It matches the information used by the design estimator in [Definition 7](#) and by the equal-group CR2 statistic in [Definition 12](#), but it treats the variance statistic itself as an arbitrary measurable function of the realized data.

**Definition 19.** [\[synth\\_7\]](#) (Independent sign prior) *For a population size  $n$ ,  $\mathbb{H}_n^{\text{ind}}$  is the product finite design on arrays  $h : \{1, \dots, n\} \times \{0, 1\} \rightarrow \{0, 1\}$  under which every unit-arm coordinate is drawn independently from the fair two-point law:*

$$\mathbb{H}_n^{\text{ind}} = \bigotimes_{(i,z) \in \{1, \dots, n\} \times \{0, 1\}} \text{Bernoulli}\left(\frac{1}{2}\right).$$

**Definition 20.** [\[synth\\_6\]](#) (Same-sign schedule prior) *For each  $n$ ,  $\mathbb{H}_n^{\text{same}}$  is the finite product design on Boolean schedules  $h : \{1, \dots, n\} \rightarrow \{0, 1\}$  whose coordinate design at every  $i \in \{1, \dots, n\}$  is the fair coin design:*

$$\mathbb{H}_n^{\text{same}} := \bigotimes_{i \in \{1, \dots, n\}} \text{Coin}\left(\frac{1}{2}\right).$$

Both priors are stated on Boolean coordinates, and a Boolean draw becomes a potential-outcome schedule through the sign map

$$\varepsilon(0) = -1, \quad \varepsilon(1) = +1.$$

Given a common-sign draw  $u : \{1, \dots, n\} \rightarrow \{0, 1\}$  from  $\mathbb{H}_n^{\text{same}}$ , the induced schedule is additive and arm-free,

$$Y_{i,n}(z, A) = \varepsilon(u_i) \quad \text{for every arm } z \text{ and every group } A \ni i,$$

so the same unit sign is carried into both arms. Given an arm-specific draw  $u : \{1, \dots, n\} \times \{0, 1\} \rightarrow \{0, 1\}$  from  $\mathbb{H}_n^{\text{ind}}$ , the induced schedule is

$$Y_{i,n}(z, A) = \varepsilon(u_{i,z}),$$

with the two arms drawn independently. Both induced schedules take values in  $\{-1, +1\}$  and are additive: the potential outcome of unit  $i$  does not depend on the composition of the group it lands in. Consequently every group table is an average of  $M$  independent signs, which is what produces the arm variances  $V_{z,n} \rightarrow 1/M$  quoted in the proposition. The independent prior in [Definition 19](#) randomizes unit signs separately by treatment arm; the same-sign prior in [Definition 20](#) uses a common unit sign across arms. Both priors generate bounded Boolean schedules, but they place different mass on the degree-one arm-difference energy in [Definition 11](#). That difference is invisible after observing only the assigned arm in a single realization.

**Proposition 3.** [[prop:rademacher-mixture-separation](#)] (Rademacher mixture separation) *Fix a common group size  $M \geq 2$ . Let  $A$  be a feasible triangular array with row population sizes  $n$ , group counts  $G_n$ , treated-group counts  $G_{1n}$ ,  $0 < G_{1n} < G_n$ ,  $MG_n \leq n$ , and deterministic schedules. Let  $p \in \mathbb{R}$  and let  $\rho$ , the limiting grouped-unit fraction, satisfy  $\rho > 0$ . Suppose:*

- **(Group growth.)** *The array satisfies [Assumption 1](#).*
- **(Treatment fractions.)** *The treated-group fractions satisfy [Assumption 2](#) with limit  $p$ .*
- **(Sampling fractions.)** *The grouped-unit fractions satisfy [Assumption 3](#) with limit  $\rho$ .*
- **(Johnson projections.)** *For every row  $n$ , the canonical Johnson projections  $\Pi_{n,k}$  on  $L^2(\Omega_{n,M})$  are linear orthogonal projections onto  $\mathcal{H}_{n,k}$ , fix  $\mathcal{H}_{n,k}$ , and have residuals orthogonal to  $\mathcal{H}_{n,k}$ .*
- **(Johnson decomposition.)** *For every row with  $2M \leq n$ , every  $f \in L^2(\Omega_{n,M})$ , and every  $A \in \Omega_{n,M}$ ,*

$$\sum_{1 \leq k \leq M} \Pi_{n,k} f(A) = f(A) - \mathbb{E}_n f,$$

*and distinct Johnson degrees are orthogonal under  $\langle \cdot, \cdot \rangle_n$ .*

- **(Kneser spectrum.)** *For every row with  $2M \leq n$ , every  $0 \leq k \leq M$ , and every  $f \in \mathcal{H}_{n,k}$ , the unnormalized Kneser adjacency operator satisfies*

$$\sum_{C \in \Omega_{n,M} : C \cap A = \emptyset} f(C) = (-1)^k \binom{n-M-k}{M-k} f(A).$$

*Then, for every row  $n$  and every statistic  $\varphi$  of the one-realization observation  $\mathcal{O}_n$ ,*

$$\mathbb{E}_{\mathbb{H}_n^{\text{same}}} \mathbb{E}_{\mathbf{Z}_n} \{\varphi(\mathcal{O}_n)\} = \mathbb{E}_{\mathbb{H}_n^{\text{ind}}} \mathbb{E}_{\mathbf{Z}_n} \{\varphi(\mathcal{O}_n)\}.$$

Moreover, under  $\mathbb{H}_n^{\text{same}}$ ,

$$V_{1,n} \xrightarrow{p} \frac{1}{M}, \quad V_{0,n} \xrightarrow{p} \frac{1}{M}, \quad E_{\tau,1,n} \xrightarrow{p} 0, \quad G_n \sigma_n^2 \xrightarrow{p} \frac{1}{Mp(1-p)}.$$

Under  $\mathbb{H}_n^{\text{ind}}$ ,

$$E_{\tau,1,n} \xrightarrow{p} \frac{2}{M}, \quad G_n \sigma_n^2 \xrightarrow{p} \frac{1}{Mp(1-p)} - \frac{2\rho}{M}.$$

The two scaled-variance limits are separated by a positive gap:

$$0 < \frac{2\rho}{M}, \quad \frac{2}{M} \leq \frac{1}{Mp(1-p)} - \frac{2\rho}{M}.$$

Consequently, for every fixed  $c_\sigma \in \mathbb{R}$  satisfying

$$0 < c_\sigma < \frac{1}{Mp(1-p)} - \frac{2\rho}{M},$$

the lower-tail probabilities satisfy

$$\Pr_{\mathbb{H}_n^{\text{same}}} \{G_n \sigma_n^2 \geq c_\sigma\} \rightarrow 1, \quad \Pr_{\mathbb{H}_n^{\text{ind}}} \{G_n \sigma_n^2 \geq c_\sigma\} \rightarrow 1.$$

**Proposition 3** isolates the information loss created by one realized treatment assignment. Every statistic of the one-realization observation  $\mathcal{O}_n$  has the same mixture expectation under  $\mathbb{H}_n^{\text{same}}$  and  $\mathbb{H}_n^{\text{ind}}$ , yet the exact-variance limit changes by  $2\rho/M$ . The equal observed-data mixture reflects the fact that each observed unit contributes only its assigned-arm sign; the variance separation comes from the unobserved cross-arm alignment, which enters [Theorem 2](#) through the degree-one term  $E_{\tau,1,n}$ . In positive-density designs, the grouped-unit fraction  $\rho$  keeps that term on the same scale as the leading variance.

The next result converts the mixture separation into a uniform lower bound over the dense schedule class. It conditions the two priors on high-probability deterministic supports, so the comparison is stated directly over arrays in  $\mathcal{C}_{\text{dense}}$  from [Definition 13](#).

**Theorem 6.** [thm:qv-diagonal-impossibility] (Diagonal variance obstruction) *Fix a common group size  $M$  with  $2 \leq M$ , and let  $A$  be a triangular schedule array with population sizes  $n$ , group counts  $G_n > 0$ , treated-group counts  $0 < G_{1n} < G_n$ ,  $MG_n \leq n$ , and deterministic potential-outcome rows. Let  $p, \rho, B, c_\sigma \in \mathbb{R}$ . Suppose:*

- **(Envelope and fractions.)**  $0 < \rho$ ,  $1 \leq B$ , and the array satisfies [Assumptions 1 to 3](#) with limiting treated-group fraction  $p$  and limiting grouped-unit fraction  $\rho$ .
- **(Johnson decomposition input.)** For every row  $n$ , there are orthogonal projections  $\Pi_{n,k}$  onto the degree- $k$  Johnson harmonic spaces  $\mathcal{H}_{n,k}$ ,  $k = 0, \dots, M$ , whose positive-degree components give the centered orthogonal decomposition of every  $f \in L^2(\Omega_{n,M})$ :

$$\sum_{k=1}^M \Pi_{n,k} f = f - \mathbb{E}_n f, \quad \langle \Pi_{n,k} f, \Pi_{n,\ell} g \rangle_n = 0 \quad \text{for } k \neq \ell.$$

- **(Kneser spectrum input.)** For every row  $n$  with  $2M \leq n$ , the Kneser adjacency operator acts on  $f \in \mathcal{H}_{n,k}$  by the eigenvalue

$$(-1)^k \binom{n-M-k}{M-k}, \quad k = 0, \dots, M.$$

- (*Variance lower bound.*) With

$$d_{\text{same}} = \frac{1}{Mp(1-p)}, \quad d_{\text{ind}} = \frac{1}{Mp(1-p)} - \frac{2\rho}{M},$$

the constant  $c_\sigma$  satisfies

$$0 < c_\sigma < d_{\text{ind}}.$$

Then there exist support sets  $\Gamma_n^{\text{same}}$  for the common-arm Rademacher prior  $\mathbb{H}_n^{\text{same}}$  and  $\Gamma_n^{\text{ind}}$  for the independent-arm Rademacher prior  $\mathbb{H}_n^{\text{ind}}$  such that both prior probabilities tend to one, every diagonal selection from either support induces an array in  $\mathcal{C}_{\text{dense}}$  with parameters  $(p, \rho, B, c_\sigma)$ , the conditioned one-realization laws satisfy

$$\text{TV}\left(\mathbb{H}_n^{\text{same}}(\mathcal{O}_n \in \cdot \mid \Gamma_n^{\text{same}}), \mathbb{H}_n^{\text{ind}}(\mathcal{O}_n \in \cdot \mid \Gamma_n^{\text{ind}})\right) \rightarrow 0,$$

and the scaled exact variances converge uniformly on the respective supports:

$$\sup_{u \in \Gamma_n^{\text{same}}} |G_n \sigma_n^2(u) - d_{\text{same}}| \rightarrow 0, \quad \sup_{u \in \Gamma_n^{\text{ind}}} |G_n \sigma_n^2(u) - d_{\text{ind}}| \rightarrow 0.$$

Moreover, there is an  $\varepsilon > 0$  such that every measurable one-realization variance statistic  $\hat{S}_n(\mathcal{O}_n)$  satisfies

$$\frac{1}{2} \leq \liminf_{n \rightarrow \infty} \sup_{Y \in \mathcal{C}_{\text{dense}}} \Pr_Y \left( \left| \frac{\hat{S}_n(\mathcal{O}_n)}{\sigma_n^2} - 1 \right| > \varepsilon \right).$$

[Theorem 6](#) establishes the one-realization lower bound in the positive-density class. The high-probability supports place both constructions inside  $\mathcal{C}_{\text{dense}}$  with the same envelope, fraction, and scaled-variance parameters, while total variation convergence makes the conditioned observed laws asymptotically indistinguishable. Uniform convergence of the scaled exact variances to  $d_{\text{same}}$  and  $d_{\text{ind}}$  supplies the separated targets, and the final display gives a single  $\varepsilon > 0$  for which every measurable statistic  $\hat{S}_n(\mathcal{O}_n)$  has worst-case relative error probability bounded below by one half.

The intuition is the same as in a two-point Le Cam comparison, adapted to deterministic triangular arrays. If the observed laws are asymptotically the same, a statistic based only on  $\mathcal{O}_n$  cannot reliably determine which variance target generated the data. Because the targets remain separated by the positive-density correction  $2\rho/M$ , any one-realization variance statistic must make a nonvanishing relative error on at least one member of the dense class. This identifies the degree-one cross-arm component in [Definition 11](#) as the obstruction to uniform exact-variance ratio consistency from a single realized grouping and assignment.

## 8 Discussion, scope, and extensions

The results above characterize a variance-estimation problem created by random group formation in dense finite populations. The exact identities in [Theorems 1](#) and [2](#) express the design variance through the Johnson decomposition and the Kneser disjointness geometry of the group table. Those identities separate the independent-group leading scale from the Kneser finite-population correction, whose dense-limit leading term is the degree-one arm-difference component in [Definition 11](#). In dense designs, [Theorem 3](#) turns that exact decomposition into an interpretable

limit: the scaled variance expansion contains the degree-one Johnson correction alongside the independent-group leading scale.

This characterization motivates the phase-frontier statement for the scalar CR2 statistic. The scalar CR2 statistic in [Definition 12](#) estimates the independent-group scale under the balanced equal-group design, and [Theorem 4](#) identifies the dense correction governing its scaled difference from the exact variance and the condition for ratio consistency. In sparse regimes, [Theorem 5](#) shows that overlap from group formation is asymptotically negligible at the variance scale considered there. Together, the dense and sparse statements give a single design-based account of when the usual cluster calculation tracks the exact randomization variance and when the finite-population group-formation term contributes to the target.

The software comparison in [Propositions 1](#) and [2](#) anchors this population geometry in the regression workflow used by applied researchers. The proposition identifies the equal-group scalar CR2 statistic with the corresponding `clubSandwich` calculation for the intercept-plus-treatment regression, in the setting of [Pustejovsky and Tipton \[2018\]](#). The eight-unit construction then gives a finite instance in which that statistic and the exact design variance differ. This finite witness complements the asymptotic analysis: the same degree-one structure that appears in [Theorem 3](#) is already visible in a small balanced design.

The lower bound in [Theorem 6](#) gives the identification-theoretic interpretation of the correction. For the dense schedule class in [Definition 13](#), the observed realization contains the realized groups, assignments, and member outcomes, while the exact variance also depends on how the unobserved counterfactual group tables co-vary across disjoint groups. [Proposition 3](#) supplies two dense constructions with asymptotically indistinguishable observed laws and separated exact variances, and [Theorem 6](#) converts that comparison into a uniform lower bound for every one-realization variance statistic. The result therefore locates the obstruction precisely: the observed data can support the usual cluster scale, but the dense finite-population correction includes a cross-arm component that is not uniformly recoverable from one realized grouping and assignment.

These conclusions also clarify the positive scope of the paper. The analysis is design-based and finite-population: potential outcomes are deterministic, randomness comes from group formation and balanced assignment, and the target is the exact variance of the group-level difference-in-means estimator. The asymptotic statements are triangular-array statements over the schedule class in [Definition 13](#), with bounded outcomes, stable treatment and sampling fractions, growing group counts, and nondegenerate scaled exact variance. Within that scope, the paper gives exact variance identities, dense and sparse asymptotic comparisons, a software identity for the scalar CR2 calculation, a finite witness, and a one-realization lower bound.

## 8.1 Limitations and future work

The results leave Gaussian approximation, studentized testing, Wald coverage, and finite-sample CR2 conservativeness as separate inferential questions. The variance identities and ratio limits characterize the target variance and the behavior of the scalar CR2 statistic, while a distributional theory for tests would require additional approximation arguments for the studentized estimator. Such arguments would naturally interact with the many-cluster refinements developed for CR2 procedures [[Pustejovsky and Tipton, 2018](#)] and with recent work on random group formation and cluster-robust inference [[Fu et al., 2026](#)].

Variable group sizes are another extension. The present analysis uses a common group size throughout, which makes the random partition a uniform slice design and brings the Johnson

association scheme directly into the variance calculation. Allowing heterogeneous group sizes would require the corresponding finite-population covariance operator and an analogue of the degree-one correction for the resulting design. That extension would preserve the same substantive question—how random group formation changes the variance target—while changing the combinatorial geometry used to express it.

## Appendices

### A Proofs and auxiliary lemmas

This appendix collects the auxiliary spectral statements used by the variance geometry and records how the formal arguments fit together. The starting point is the Johnson association scheme on the uniform slice, whose orthogonal decomposition supplies the coordinates used in [Definition 9](#) and whose Kneser disjointness operator supplies the covariance operator in [Definition 10](#). Classical accounts of these schemes appear in [Delsarte \[1973\]](#), [Brouwer et al. \[1989\]](#), and [Brouwer et al. \[2018\]](#); the harmonic-analysis notation follows the slice presentation in [Filmus \[2016\]](#).

**Definition 21.** [\[synth\\_10\]](#) (Johnson and Kneser covariance apparatus) *For  $0 \leq k \leq M$ , let  $\mathcal{H}_{n,k} \subseteq L^2(\Omega_{n,M})$  be the degree- $k$  Johnson harmonic subspace, and let  $\Pi_{n,k}$  be the  $\langle \cdot, \cdot \rangle_n$ -orthogonal projection onto  $\mathcal{H}_{n,k}$ . For  $f \in L^2(\Omega_{n,M})$ , the centered Johnson decomposition is*

$$f - \mathbb{E}_n f = \sum_{k=1}^M \Pi_{n,k} f.$$

*For  $2M \leq n$ , the unnormalized Kneser adjacency operator is*

$$(\text{Adj}_{n,M} f)(A) = \sum_{\substack{C \in \Omega_{n,M} \\ C \cap A = \emptyset}} f(C), \quad A \in \Omega_{n,M}.$$

*The normalized Kneser disjointness operator is*

$$(\mathbf{K}_n f)(A) = \binom{n-M}{M}^{-1} \sum_{\substack{C \in \Omega_{n,M} \\ C \cap A = \emptyset}} f(C).$$

*Its eigenvalue on  $\mathcal{H}_{n,k}$  is denoted by  $\lambda_{n,k}$ . For  $a, b \in \mathcal{Z}$ , define the ordered-disjoint covariance and the arm-difference covariance by*

$$C_{ab,n} = \langle h_{a,n}^\circ, \mathbf{K}_n h_{b,n}^\circ \rangle_n, \quad C_{\tau\tau,n} = C_{11,n} + C_{00,n} - 2C_{10,n}.$$

*The degree-one arm-difference energy is*

$$E_{\tau,1,n} = \|\Pi_{n,1}(h_{1,n} - h_{0,n})\|_n^2.$$

The objects in [Definition 21](#) restate, in one place, the spectral apparatus used throughout the exact and asymptotic variance arguments. The decomposition separates centered group tables into orthogonal Johnson degrees, and the Kneser operator averages over disjoint groups,

matching the dependence induced by random partitioning. With this notation fixed, two Kneser statements follow, and they are deliberately distinct. [Lemma 1](#) is the combinatorial fact: the *unnormalized* disjointness adjacency operator  $\text{Adj}_{n,M}$ , which sums a table over all groups disjoint from a given one, acts on the degree- $k$  harmonic space by the integer eigenvalue  $(-1)^k \binom{n-M-k}{M-k}$ . [Lemma 2](#) is the form the variance algebra consumes: it divides by the number  $\binom{n-M}{M}$  of disjoint groups to obtain the *averaging* operator  $K_n$  of [Definition 10](#), whose eigenvalues  $\lambda_{n,k} = (-1)^k (M)_k / (n-M)_k$  are the falling-factorial ratios that carry the  $n$ -dependence of the variance expansion. The second is obtained from the first by that normalization, which is why its proof is short; keeping them apart isolates the one combinatorial input of the paper from the bookkeeping that turns it into a covariance statement.

**Lemma 1.** [[lem:johnson-kneser-adjacency-eigenvalue](#)] (Johnson–Kneser adjacency eigenvalue) *Using the uniform  $M$ -slice  $\Omega_{n,M}$ , the slice function space  $L^2(\Omega_{n,M})$ , and the Johnson harmonic subspaces  $\mathcal{H}_{n,k}$  from [Definition 9](#), suppose that:*

- **(Population and group size.)**  $n$  and  $M$  are nonnegative integers, with  $n$  the population size and  $M$  the common group size, and  $2M \leq n$ .
- **(Harmonic degree.)**  $k \in \{0, \dots, M\}$ .
- **(Function.)**  $f \in \mathcal{H}_{n,k}$ .

Let  $\text{Adj}_{n,M}$  be the linear operator on  $L^2(\Omega_{n,M})$  defined by

$$(\text{Adj}_{n,M}f)(\Gamma) = \sum_{\Delta \in \Omega_{n,M}} \mathbf{1}\{\Gamma \cap \Delta = \emptyset\} f(\Delta).$$

Then  $f$  is an eigenfunction of  $\text{Adj}_{n,M}$  with eigenvalue

$$(-1)^k \binom{n-M-k}{M-k}.$$

Equivalently,

$$\text{Adj}_{n,M}f = (-1)^k \binom{n-M-k}{M-k} f$$

as an identity in  $L^2(\Omega_{n,M})$ .

*Proof of [Lemma 1](#).* Write

$$m_{\Upsilon}(\Gamma) := \mathbf{1}\{\Upsilon \subseteq \Gamma\}, \quad V_{\leq q} := \text{span}\{m_{\Upsilon} : |\Upsilon| \leq q\} \subseteq L^2(\Omega_{n,M}),$$

for finite  $\Upsilon \subseteq \{1, \dots, n\}$  and  $q \geq 0$ . Also write

$$\delta_{\Upsilon}(\Gamma) := \mathbf{1}\{\Upsilon \cap \Gamma = \emptyset\}.$$

The Johnson harmonic construction used in [Definition 9](#) has

$$\mathcal{H}_{n,0} = V_{\leq 0}, \quad \mathcal{H}_{n,j} = V_{\leq j} \cap V_{\leq j-1}^{\perp} \quad (j \geq 1),$$

where orthogonality is with respect to the uniform-slice inner product. The hypothesis  $2M \leq n$  gives  $M \leq n$ , so  $\Omega_{n,M}$  is nonempty.

1. First take  $k = 0$ . Since  $V_{\leq 0}$  is spanned by  $m_{\emptyset}$ , membership  $f \in \mathcal{H}_{n,0} = V_{\leq 0}$  gives a scalar  $\eta \in \mathbb{R}$  such that

$$f = \eta m_{\emptyset}.$$

For every  $\Gamma \in \Omega_{n,M}$ ,

$$(\text{Adj}_{n,M} m_{\emptyset})(\Gamma) = \#\{\Delta \in \Omega_{n,M} : \Gamma \cap \Delta = \emptyset\} = \binom{n-M}{M}.$$

Indeed, after  $\Gamma$  is fixed, the admissible  $\Delta$ 's are precisely the  $M$ -subsets of the  $n-M$  points outside  $\Gamma$ . Thus

$$\text{Adj}_{n,M} f = \eta \binom{n-M}{M} m_{\emptyset} = (-1)^0 \binom{n-M}{M} f,$$

which is the asserted formula in degree zero.

2. Now suppose  $k \geq 1$ . Define the scalar

$$\alpha_k := (-1)^k \binom{n-M-k}{M-k}$$

and the residual

$$\rho := \text{Adj}_{n,M} f - \alpha_k f.$$

We first show

$$\rho \in V_{\leq k-1}.$$

It is enough to check this on the spanning monomials  $m_{\Upsilon}$  with  $|\Upsilon| \leq k$ . If  $|\Upsilon| < k$ , then  $m_{\Upsilon} \in V_{\leq k-1}$ , and the adjacency operator preserves this lower filtration level:

$$\text{Adj}_{n,M} m_{\Upsilon} \in V_{\leq k-1}.$$

Hence  $\text{Adj}_{n,M} m_{\Upsilon} - \alpha_k m_{\Upsilon} \in V_{\leq k-1}$ .

It remains to consider  $|\Upsilon| = k$ . For every  $\Gamma \in \Omega_{n,M}$ ,

$$(\text{Adj}_{n,M} m_{\Upsilon})(\Gamma) = \binom{n-M-k}{M-k} \delta_{\Upsilon}(\Gamma).$$

When  $\Gamma \cap \Upsilon \neq \emptyset$ , no disjoint  $\Delta$  can contain  $\Upsilon$ . When  $\Gamma \cap \Upsilon = \emptyset$ , such  $\Delta$ 's are obtained by adjoining  $M-k$  further points from the complement of  $\Gamma \cup \Upsilon$ , which has size  $n-M-k$ . Inclusion-exclusion gives

$$\delta_{\Upsilon} = \sum_{\Theta \subseteq \Upsilon} (-1)^{|\Theta|} m_{\Theta}.$$

Therefore

$$\text{Adj}_{n,M} m_{\Upsilon} - \alpha_k m_{\Upsilon} = \binom{n-M-k}{M-k} \sum_{\Theta \subsetneq \Upsilon} (-1)^{|\Theta|} m_{\Theta} \in V_{\leq k-1},$$

because every proper subset  $\Theta \subsetneq \Upsilon$  has  $|\Theta| \leq k-1$ . By linearity over the span  $V_{\leq k}$ , the same conclusion holds for every  $f \in V_{\leq k}$ , in particular for the present  $f \in \mathcal{H}_{n,k} \subseteq V_{\leq k}$ .

3. Since  $f \in \mathcal{H}_{n,k} = V_{\leq k} \cap V_{\leq k-1}^\perp$ , for every  $g \in V_{\leq k-1}$ ,

$$\langle f, g \rangle_n = 0.$$

The adjacency operator preserves  $V_{\leq k-1}$ , so  $\text{Adj}_{n,M}g \in V_{\leq k-1}$ . It is also self-adjoint for the uniform-slice inner product:

$$\langle \text{Adj}_{n,M}u, v \rangle_n = \langle u, \text{Adj}_{n,M}v \rangle_n \quad (u, v \in L^2(\Omega_{n,M})),$$

which follows by interchanging the two finite sums and using symmetry of disjointness. Hence, for every  $g \in V_{\leq k-1}$ ,

$$\begin{aligned} \langle \rho, g \rangle_n &= \langle \text{Adj}_{n,M}f, g \rangle_n - \alpha_k \langle f, g \rangle_n \\ &= \langle f, \text{Adj}_{n,M}g \rangle_n - \alpha_k \langle f, g \rangle_n = 0. \end{aligned}$$

Thus  $\rho$  is orthogonal to  $V_{\leq k-1}$ .

4. Combining the previous two steps,  $\rho \in V_{\leq k-1}$  and  $\rho \perp V_{\leq k-1}$ . Taking  $g = \rho$  gives

$$\langle \rho, \rho \rangle_n = 0.$$

Since  $M \leq n$ ,

$$\langle \rho, \rho \rangle_n = \binom{n}{M}^{-1} \sum_{\Gamma \in \Omega_{n,M}} \rho(\Gamma)^2$$

with  $\binom{n}{M} > 0$ , so the uniform-slice inner product is positive definite. Therefore  $\rho = 0$ , i.e.

$$\text{Adj}_{n,M}f = \alpha_k f = (-1)^k \binom{n-M-k}{M-k} f.$$

This is the desired identity for  $k \geq 1$ , and together with the degree-zero case proves the lemma. □

**Lemma 2.** [lem:classical-kneser-spectrum] (Classical Kneser spectrum) *For any population size  $n$  and common group size  $M$ , suppose:*

- **(Overlap.)**  $2M \leq n$ .
- **(Harmonic degree.)**  $k$  is an integer with  $0 \leq k \leq M$ .
- **(Harmonic function.)**  $f : \Omega_{n,M} \rightarrow \mathbb{R}$  belongs to the degree- $k$  Johnson harmonic subspace  $\mathcal{H}_{n,k}$ .

*Then the unnormalized Kneser adjacency operator, which sums  $f$  over  $M$ -subsets disjoint from the argument, acts on  $f$  by*

$$(\text{Adj}_{n,M}f)(A) = (-1)^k \binom{n-M-k}{M-k} f(A)$$

*for every  $A \in \Omega_{n,M}$ .*

*Proof of Lemma 2.* Fix  $n, M$  with  $2M \leq n$ , fix  $0 \leq k \leq M$ , and let  $f \in \mathcal{H}_{n,k}$ . Write  $[n] := \{1, \dots, n\}$ .

1. For a slice function  $u : \Omega_{n,M} \rightarrow \mathbb{R}$ , use the unnormalized Kneser adjacency operator

$$(\text{Adj}_{n,M}u)(\Gamma) := \sum_{\Delta \in \Omega_{n,M}} \mathbf{1}\{\Gamma \cap \Delta = \emptyset\} u(\Delta), \quad \Gamma \in \Omega_{n,M}.$$

The hypotheses  $2M \leq n$ ,  $0 \leq k \leq M$ , and  $f \in \mathcal{H}_{n,k}$  are exactly the hypotheses needed to apply Lemma 1 to this operator. It gives

$$\text{Adj}_{n,M}f = (-1)^k \binom{n-M-k}{M-k} f$$

as an identity in the slice function space  $L^2(\Omega_{n,M})$ .

2. The preceding function identity says that, for every  $A \in \Omega_{n,M}$ ,

$$\sum_{\Delta \in \Omega_{n,M}} \mathbf{1}\{A \cap \Delta = \emptyset\} f(\Delta) = (-1)^k \binom{n-M-k}{M-k} f(A).$$

The left-hand side is exactly  $(\text{Adj}_{n,M}f)(A)$ , so

$$(\text{Adj}_{n,M}f)(A) = (-1)^k \binom{n-M-k}{M-k} f(A),$$

which is the asserted formula.

□

Together, Lemmas 1 and 2 give the spectral multiplier for disjointness averaging on each Johnson degree. After normalization by the number of disjoint  $M$ -subsets, these multipliers become the eigenvalues used in the covariance contrast of Theorem 1. The alternating sign in degree  $k$  is the algebraic source of the projection correction that later appears in Theorem 3.

The remaining classical input is the orthogonal expansion of functions on the slice. It ensures that the covariance calculation can be expressed degree by degree rather than as a direct summation over all ordered disjoint pairs.

**Lemma 3.** [lem:classical-johnson-decomposition] (Johnson orthogonal decomposition) *Let  $n$  and  $M$  be nonnegative integers with  $M \leq n$ . For the canonical Johnson harmonic projections  $\Pi_{n,k}$  on  $L^2(\Omega_{n,M})$ , whenever  $2M \leq n$ , every function  $f : \Omega_{n,M} \rightarrow \mathbb{R}$  and every  $A \in \Omega_{n,M}$  satisfy*

$$\sum_{k=1}^M \Pi_{n,k}f(A) = f(A) - \mathbb{E}_n f.$$

*Moreover, the projections are pairwise orthogonal: for all  $k, \ell \in \{0, \dots, M\}$ , all functions  $f, g : \Omega_{n,M} \rightarrow \mathbb{R}$ , and  $k \neq \ell$ ,*

$$\langle \Pi_{n,k}f, \Pi_{n,\ell}g \rangle_n = 0.$$

*Proof of Lemma 3.* Assume  $2M \leq n$ . Then  $M \leq n$ , so  $\Omega_{n,M}$  is the finite uniform slice with

$$|\Omega_{n,M}| = \binom{n}{M}.$$

For a slice function  $\varphi : \Omega_{n,M} \rightarrow \mathbb{R}$ , write

$$\text{mean}_n(\varphi) := \frac{1}{|\Omega_{n,M}|} \sum_{\omega \in \Omega_{n,M}} \varphi(\omega).$$

1. For  $0 \leq u \leq M$ , let  $\mathcal{F}_{\leq u} \subseteq L^2(\Omega_{n,M})$  be the span of the inclusion monomials

$$\omega \mapsto \prod_{i \in T} \mathbf{1}_{\{i \in \omega\}}, \quad T \subseteq \{1, \dots, n\}, \quad |T| \leq u.$$

The Johnson harmonic layers are constructed by

$$\mathcal{H}_{n,0} = \mathcal{F}_{\leq 0}, \quad \mathcal{H}_{n,u+1} = \mathcal{F}_{\leq u+1} \cap \mathcal{F}_{\leq u}^\perp.$$

Since  $\mathcal{F}_{\leq u} \subseteq \mathcal{F}_{\leq u+1}$ , finite-dimensional orthogonal decomposition gives

$$\mathcal{F}_{\leq u+1} = \mathcal{F}_{\leq u} \oplus \mathcal{H}_{n,u+1}.$$

Iterating from  $u = 0$  to  $M - 1$ ,

$$\mathcal{F}_{\leq M} = \bigoplus_{u=0}^M \mathcal{H}_{n,u}.$$

Moreover  $\mathcal{F}_{\leq M} = L^2(\Omega_{n,M})$ : for every fixed  $\omega_0 \in \Omega_{n,M}$ ,

$$\mathbf{1}_{\{\omega_0\}}(\omega) = \prod_{i \in \omega_0} \mathbf{1}_{\{i \in \omega\}},$$

because  $|\omega| = |\omega_0| = M$ . Thus point masses belong to  $\mathcal{F}_{\leq M}$ , and they span all slice functions.

Therefore the orthogonal projections onto the layers satisfy, for every  $\varphi : \Omega_{n,M} \rightarrow \mathbb{R}$ ,

$$\varphi = \sum_{u=0}^M \Pi_{n,u} \varphi.$$

The degree-zero space  $\mathcal{H}_{n,0}$  is the space of constant functions. Its orthogonal projection is the constant function with value  $\text{mean}_n(\varphi)$ , since for every real scalar  $a_{\text{const}}$ ,

$$\langle \varphi - (\omega \mapsto \text{mean}_n(\varphi)), (\omega \mapsto a_{\text{const}}) \rangle_n = a_{\text{const}} (\text{mean}_n(\varphi) - \text{mean}_n(\varphi)) = 0.$$

Hence

$$\Pi_{n,0} \varphi = \omega \mapsto \text{mean}_n(\varphi).$$

Evaluating the full decomposition at  $A \in \Omega_{n,M}$  and subtracting the degree-zero component gives

$$\sum_{u=1}^M \Pi_{n,u} \varphi(A) = \varphi(A) - \text{mean}_n(\varphi).$$

Applied to  $\varphi = f$ , this is the centered projection identity with the algebraic uniform mean.

2. The averaging operator  $\mathbb{E}_n$  is the uniform-slice expectation, so for every  $f : \Omega_{n,M} \rightarrow \mathbb{R}$ ,

$$\mathbb{E}_n f = \frac{1}{|\Omega_{n,M}|} \sum_{\omega \in \Omega_{n,M}} f(\omega) = \text{mean}_n(f).$$

Substituting this equality into the preceding display yields

$$\sum_{u=1}^M \Pi_{n,u} f(A) = f(A) - \mathbb{E}_n f.$$

When  $M = 0$ , the same formula reads  $0 = f(A) - \mathbb{E}_n f$ , because the slice has one point and the degree-zero projection is  $f$ .

3. It remains to check distinct-degree orthogonality. For any  $k \in \{0, \dots, M\}$  and any slice function  $f$ ,

$$\Pi_{n,k} f \in \mathcal{H}_{n,k},$$

because  $\Pi_{n,k}$  is the orthogonal projection onto  $\mathcal{H}_{n,k}$ . If  $k < \ell$ , then

$$\mathcal{H}_{n,\ell} \subseteq \mathcal{F}_{\leq \ell-1}^\perp, \quad \mathcal{H}_{n,k} \subseteq \mathcal{F}_{\leq k} \subseteq \mathcal{F}_{\leq \ell-1},$$

and therefore every element of  $\mathcal{H}_{n,k}$  is orthogonal to every element of  $\mathcal{H}_{n,\ell}$ . The case  $\ell < k$  is symmetric. Thus, whenever  $k \neq \ell$ ,

$$\langle \Pi_{n,k} f, \Pi_{n,\ell} g \rangle_n = 0.$$

This proves the pairwise orthogonality assertion.

□

**Lemma 3** is the bookkeeping device behind the exact variance formula. Applying the decomposition to the centered arm tables in [Definition 5](#) lets the Kneser covariance in [Definition 10](#) act separately on each harmonic component. This is what turns the two-stage random partition design in [Definition 4](#) into the exact identity in [Theorem 2](#).

The subsequent verified arguments use these ingredients in a fixed order. First, the finite-sample variance algebra combines the PAME estimator in [Definition 7](#) with the Kneser covariance identity to obtain [Theorem 2](#). Next, the dense schedule class in [Definition 13](#) specializes that identity to the projection limit in [Theorem 3](#) and to the CR2 comparison in [Theorem 4](#); the sparse construction gives the complementary behavior in [Theorem 5](#). The software calculation in [Proposition 1](#) identifies the scalar CR2 statistic with the stated `clubSandwich` call, and [Proposition 2](#) evaluates the finite eight-unit witness from [Definition 17](#). Finally, the randomized schedule construction in [Proposition 3](#) feeds the one-realization lower bound in [Theorem 6](#).

The finite-sample and software statements are algebraic consequences of the displayed definitions and of the versioned software conventions cited in [Pustejovsky \[2026\]](#) and [Zeileis and Lumley \[2026\]](#). The surrounding econometric motivation connects these calculations to random group formation and cluster-robust variance practice as in [Fu et al. \[2026\]](#).

**Verification note.** Every theorem, proposition, and assumption displayed in this paper, and every internal algebraic derivation supporting them, is machine-checked in Lean 4, with no unproved hypotheses and no remaining proof obligations. The displayed definitions divide into two kinds. Those carrying a Lean declaration – the schedule array, the arm set, slice averaging, the slice norm, the centered arm table, the two Rademacher priors, the one-realization observation, the row treatment fraction, and all numbered `def`: objects – are the objects the verified statements quantify over. Three are presentation-level definitions written for the reader and carry no Lean declaration of their own: the uniform-slice space  $L^2(\Omega_{n,M})$  in [Definition 8](#), the realized arm group counts in [Definition 14](#), and the unnormalized adjacency operator  $\text{Adj}_{n,M}$  in [Definition 21](#). They name notation that the verified statements express directly in terms of the slice and its inner product, and nothing in the verified development depends on them. The trust boundary is narrower than is usual for results of this kind: the classical Johnson decomposition and Kneser spectrum recorded in [Lemmas 1 to 3](#) are not imported as cited facts but are themselves proved in the verified development, so the citations to [Delsarte \[1973\]](#), [Brouwer et al. \[1989\]](#), [Filmus \[2016\]](#), and [Brouwer et al. \[2018\]](#) attribute these classical results rather than assume them. What remains outside the machine-checked scope is the bibliographic attribution itself and the versioned software conventions of [Pustejovsky \[2026\]](#) and [Zeileis and Lumley \[2026\]](#), which enter [Proposition 1](#) as stated implementation conventions.

## B Proofs of the main results

*Proof of [Theorem 1](#).* 1. Let

$$(\text{A}_n q)(A) = \sum_{C: C \cap A = \emptyset} q(C)$$

be the unnormalized Kneser adjacency operator on functions  $q : \Omega_{n,M} \rightarrow \mathbb{R}$ . For  $k = 0, \dots, M$  and  $f : \Omega_{n,M} \rightarrow \mathbb{R}$ , the projection assumption gives  $\Pi_{n,k} f \in \mathcal{H}_{n,k}$ , so the Kneser spectrum assumption gives

$$\text{A}_n(\Pi_{n,k} f)(A) = (-1)^k \binom{n-M-k}{M-k} \Pi_{n,k} f(A).$$

By [Definition 10](#),

$$\text{K}_n(\Pi_{n,k} f)(A) = (-1)^k \frac{\binom{n-M-k}{M-k}}{\binom{n-M}{M}} \Pi_{n,k} f(A).$$

Since  $k \leq M$  and  $2M \leq n$ ,

$$\frac{\binom{n-M-k}{M-k}}{\binom{n-M}{M}} = \frac{(n-M-k)_{M-k}}{(M-k)!} \frac{M!}{(n-M)_M} = \frac{(M)_k}{(n-M)_k},$$

using the falling-factorial convention  $(r)_0 = 1$ . Hence

$$\text{K}_n(\Pi_{n,k} f)(A) = \lambda_{n,k} \Pi_{n,k} f(A).$$

2. Now let  $f, g : \Omega_{n,M} \rightarrow \mathbb{R}$  satisfy  $\mathbb{E}_n f = \mathbb{E}_n g = 0$ , and put

$$\mathcal{I}_+ = \{j \in \{0, \dots, M\} : j > 0\}.$$

The centered decomposition in [Definition 9](#) gives, for every  $A \in \Omega_{n,M}$ ,

$$f(A) = \sum_{j \in \mathcal{I}_+} \Pi_{n,j} f(A), \quad g(A) = \sum_{j \in \mathcal{I}_+} \Pi_{n,j} g(A).$$

For a uniform ordered disjoint pair  $(A, C)$ , each coordinate has the uniform  $M$ -slice marginal, and the product expectation equals the Kneser inner product:

$$\mathbb{E}[f(A)] = \mathbb{E}_n f, \quad \mathbb{E}[g(C)] = \mathbb{E}_n g,$$

and

$$\mathbb{E}[f(A)g(C)] = \langle f, \mathbf{K}_n g \rangle_n.$$

Therefore

$$\text{Cov}\{f(A), g(C)\} = \langle f, \mathbf{K}_n g \rangle_n.$$

Substituting the positive-degree decompositions and using linearity of  $\mathbf{K}_n$  and of  $\langle \cdot, \cdot \rangle_n$ ,

$$\begin{aligned} \langle f, \mathbf{K}_n g \rangle_n &= \left\langle \sum_{i \in \mathcal{I}_+} \Pi_{n,i} f, \mathbf{K}_n \sum_{j \in \mathcal{I}_+} \Pi_{n,j} g \right\rangle_n \\ &= \sum_{i \in \mathcal{I}_+} \sum_{j \in \mathcal{I}_+} \lambda_{n,j} \langle \Pi_{n,i} f, \Pi_{n,j} g \rangle_n \\ &= \sum_{j \in \mathcal{I}_+} \lambda_{n,j} \langle \Pi_{n,j} f, \Pi_{n,j} g \rangle_n, \end{aligned}$$

because distinct Johnson degrees are orthogonal. This proves the covariance identity.

3. For each arm  $z \in \{0, 1\}$ , define the centered arm table

$$h_{z,n}^\circ(A) = h_{z,n}(A) - \mathbb{E}_n h_{z,n}.$$

Then  $\mathbb{E}_n h_{z,n}^\circ = 0$ . By [Definition 10](#) and the preceding ordered-pair identity,

$$C_{ab,n} = \text{Cov}\{h_{a,n}^\circ(A), h_{b,n}^\circ(C)\}$$

for  $a, b \in \{0, 1\}$  under a uniform ordered disjoint pair  $(A, C)$ . Applying the covariance identity from the previous step to  $(h_{1,n}^\circ, h_{1,n}^\circ)$ ,  $(h_{0,n}^\circ, h_{0,n}^\circ)$ , and  $(h_{1,n}^\circ, h_{0,n}^\circ)$ , and using

$$C_{\tau\tau,n} = C_{11,n} + C_{00,n} - 2C_{10,n},$$

gives

$$C_{\tau\tau,n} = \sum_{j \in \mathcal{I}_+} \lambda_{n,j} [\langle \Pi_{n,j} h_{1,n}^\circ, \Pi_{n,j} h_{1,n}^\circ \rangle_n + \langle \Pi_{n,j} h_{0,n}^\circ, \Pi_{n,j} h_{0,n}^\circ \rangle_n - 2\langle \Pi_{n,j} h_{1,n}^\circ, \Pi_{n,j} h_{0,n}^\circ \rangle_n].$$

4. It remains to identify the summand with the squared norm of the projected arm contrast. For every positive degree  $j \in \mathcal{I}_+$  and constant  $r \in \mathbb{R}$ ,

$$\Pi_{n,j}(r) = 0.$$

Indeed, the centered decomposition applied to the constant function  $r$  gives

$$\sum_{\ell \in \mathcal{I}_+} \Pi_{n,\ell}(r) = 0.$$

Taking the slice inner product with  $\Pi_{n,j}(r)$  and using degree orthogonality leaves

$$\langle \Pi_{n,j}(r), \Pi_{n,j}(r) \rangle_n = 0,$$

hence  $\Pi_{n,j}(r) = 0$  by definiteness of the slice inner product. Therefore, by linearity of  $\Pi_{n,j}$ ,

$$\Pi_{n,j}h_{z,n}^\circ = \Pi_{n,j}(h_{z,n} - \mathbb{E}_n h_{z,n}) = \Pi_{n,j}h_{z,n}.$$

Again by linearity,

$$\Pi_{n,j}(h_{1,n} - h_{0,n}) = \Pi_{n,j}h_{1,n} - \Pi_{n,j}h_{0,n}.$$

Thus

$$\begin{aligned} \|\Pi_{n,j}(h_{1,n} - h_{0,n})\|_n^2 &= \langle \Pi_{n,j}h_{1,n} - \Pi_{n,j}h_{0,n}, \Pi_{n,j}h_{1,n} - \Pi_{n,j}h_{0,n} \rangle_n \\ &= \langle \Pi_{n,j}h_{1,n}, \Pi_{n,j}h_{1,n} \rangle_n + \langle \Pi_{n,j}h_{0,n}, \Pi_{n,j}h_{0,n} \rangle_n - 2\langle \Pi_{n,j}h_{1,n}, \Pi_{n,j}h_{0,n} \rangle_n. \end{aligned}$$

Substitution into the preceding display yields

$$C_{\tau\tau,n} = \sum_{j=1}^M \lambda_{n,j} \|\Pi_{n,j}(h_{1,n} - h_{0,n})\|_n^2,$$

which is the asserted contrast identity. □

*Proof of Theorem 2.* By the group-size hypotheses,  $G_{1n} > 0$ ,  $G_{1n} < G_n$ ,  $G_{0n} = G_n - G_{1n} > 0$ , and  $2 \leq G_n$ . The feasibility condition and  $2 \leq G_n$  also give  $2M \leq n$ . Write the treated-group fraction as

$$p_n = \frac{G_{1n}}{G_n}, \quad q_n = 1 - p_n = \frac{G_{0n}}{G_n}.$$

For a realized ordered partition  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_{G_n})$  of pairwise-disjoint elements of  $\Omega_{n,M}$ , define

$$x_{zg}(\boldsymbol{\alpha}) = h_{z,n}(\alpha_g), \quad d_g(\boldsymbol{\alpha}) = x_{1g}(\boldsymbol{\alpha}) - x_{0g}(\boldsymbol{\alpha}),$$

and, for any vector  $\mathbf{u} = (u_1, \dots, u_{G_n}) \in \mathbb{R}^{G_n}$ ,

$$\bar{u} = \frac{1}{G_n} \sum_{g=1}^{G_n} u_g, \quad S(\mathbf{u}) = \frac{1}{G_n - 1} \sum_{g=1}^{G_n} (u_g - \bar{u})^2.$$

Set

$$S_z(\boldsymbol{\alpha}) = S((x_{zg}(\boldsymbol{\alpha}))_{g=1}^{G_n}), \quad S_\tau(\boldsymbol{\alpha}) = S((d_g(\boldsymbol{\alpha}))_{g=1}^{G_n}).$$

These quantities are well defined because  $G_n \geq 2$ .

Let

$$\Delta_n(A) = h_{1,n}(A) - h_{0,n}(A), \quad \Delta_n^\circ(A) = \Delta_n(A) - \mathbb{E}_n \Delta_n(A).$$

Since  $h_{z,n}^\circ(A) = h_{z,n}(A) - \mathbb{E}_n h_{z,n}(A)$ , we have

$$\Delta_n^\circ = h_{1,n}^\circ - h_{0,n}^\circ.$$

Using the bilinearity of the uniform-slice inner product in [Definition 10](#) and the symmetry of the uniform ordered-disjoint-pair law under swapping the two groups,

$$\begin{aligned} \langle \Delta_n^\circ, \mathbf{K}_n \Delta_n^\circ \rangle_n &= C_{11,n} + C_{00,n} - C_{10,n} - C_{01,n} \\ &= C_{11,n} + C_{00,n} - 2C_{10,n} = C_{\tau\tau,n}. \end{aligned}$$

We next record the two partition-moment identities used below. For any  $\phi : \Omega_{n,M} \rightarrow \mathbb{R}$ , let

$$\mu_\phi = \mathbb{E}_n \phi(A), \quad \eta_\phi = \mathbb{E}_n \{\phi(A)^2\},$$

and let

$$\kappa_\phi = \mathbb{E}_A \mathbb{E}_{C|C \cap A = \emptyset} \{\phi(A)\phi(C)\}.$$

Then

$$\text{Var}_n(\phi) = \eta_\phi - \mu_\phi^2, \quad \langle \phi - \mu_\phi, \mathbf{K}_n(\phi - \mu_\phi) \rangle_n = \kappa_\phi - \mu_\phi^2.$$

Each coordinate of  $\alpha$  is uniform on  $\Omega_{n,M}$ , and each distinct ordered coordinate pair is a uniform ordered-disjoint pair. Therefore

$$\begin{aligned} \mathbb{E}_\alpha \mathbf{S}((\phi(\alpha_g))_{g=1}^{G_n}) &= \frac{1}{G_n - 1} \left\{ G_n \eta_\phi - \frac{1}{G_n} (G_n \eta_\phi + G_n(G_n - 1) \kappa_\phi) \right\} \\ &= \eta_\phi - \kappa_\phi \\ &= \text{Var}_n(\phi) - \langle \phi - \mu_\phi, \mathbf{K}_n(\phi - \mu_\phi) \rangle_n. \end{aligned}$$

Applying this with  $\phi = h_{1,n}$ ,  $\phi = h_{0,n}$ , and  $\phi = \Delta_n$  gives

$$\mathbb{E}_\alpha \mathbf{S}_1 = V_{1,n} - C_{11,n}, \quad \mathbb{E}_\alpha \mathbf{S}_0 = V_{0,n} - C_{00,n},$$

and

$$\mathbb{E}_\alpha \mathbf{S}_\tau = \text{Var}_n(\Delta_n) - C_{\tau\tau,n}.$$

The same exchangeability calculation gives the variance of the partition average. Indeed,

$$\text{Var}_\alpha \left\{ \frac{1}{G_n} \sum_{g=1}^{G_n} \phi(\alpha_g) \right\} = \frac{1}{G_n^2} \{ G_n \text{Var}_n(\phi) + G_n(G_n - 1) \langle \phi - \mu_\phi, \mathbf{K}_n(\phi - \mu_\phi) \rangle_n \}.$$

Thus, for  $\phi = \Delta_n$ ,

$$\text{Var}_\alpha \left\{ \frac{1}{G_n} \sum_{g=1}^{G_n} d_g(\alpha) \right\} = \frac{\text{Var}_n(\Delta_n) + (G_n - 1)C_{\tau\tau,n}}{G_n}.$$

Also, since each coordinate  $\alpha_g$  is uniform on  $\Omega_{n,M}$ ,

$$\mathbb{E}_\alpha \left\{ \frac{1}{G_n} \sum_{g=1}^{G_n} d_g(\alpha) \right\} = \mathbb{E}_n h_{1,n} - \mathbb{E}_n h_{0,n}.$$

Condition on  $\alpha$ . Under the balanced assignment in [Definition 4](#),

$$\mathbb{E}(Z_g \mid \alpha) = p_n, \quad \mathbb{E}(1 - Z_g \mid \alpha) = q_n,$$

and the usual complete-randomization second moments for distinct  $g \neq \ell$  are

$$\mathbb{E}(Z_g Z_\ell \mid \alpha) = \frac{G_{1n}(G_{1n} - 1)}{G_n(G_n - 1)},$$

$$\mathbb{E}\{(1 - Z_g)(1 - Z_\ell) \mid \alpha\} = \frac{G_{0n}(G_{0n} - 1)}{G_n(G_n - 1)},$$

$$\mathbb{E}\{Z_g(1 - Z_\ell) \mid \alpha\} = \frac{G_{1n}G_{0n}}{G_n(G_n - 1)}.$$

Expanding the difference-in-means estimator from [Definition 7](#) with these moments gives

$$\mathbb{E}(\hat{\tau}_n \mid \alpha) = \frac{1}{G_n} \sum_{g=1}^{G_n} d_g(\alpha),$$

and

$$\text{Var}(\hat{\tau}_n \mid \alpha) = \frac{S_1(\alpha)}{G_{1n}} + \frac{S_0(\alpha)}{G_{0n}} - \frac{S_\tau(\alpha)}{G_n}.$$

The same complete-randomization calculation for the denominator- $G_{zn} - 1$  within-arm sample variances in [Definition 12](#) yields

$$\mathbb{E}(\hat{V}_{\text{CR2},n} \mid \alpha) = \frac{S_1(\alpha)}{G_{1n}} + \frac{S_0(\alpha)}{G_{0n}}.$$

Taking expectations in the conditional mean identity and using the partition-average identity above,

$$\mathbb{E}[\hat{\tau}_n] = \mathbb{E}_n h_{1,n} - \mathbb{E}_n h_{0,n} = \tau_n,$$

where the last equality is [Definition 7](#).

By the finite tower variance identity for the two-stage design,

$$\sigma_n^2 = \mathbb{E}_\alpha \{\text{Var}(\hat{\tau}_n \mid \alpha)\} + \text{Var}_\alpha \{\mathbb{E}(\hat{\tau}_n \mid \alpha)\}.$$

Substituting the identities already proved gives

$$\begin{aligned} \sigma_n^2 &= \frac{V_{1,n} - C_{11,n}}{G_{1n}} + \frac{V_{0,n} - C_{00,n}}{G_{0n}} - \frac{\text{Var}_n(\Delta_n) - C_{\tau\tau,n}}{G_n} + \frac{\text{Var}_n(\Delta_n) + (G_n - 1)C_{\tau\tau,n}}{G_n} \\ &= \frac{V_{1,n} - C_{11,n}}{G_{1n}} + \frac{V_{0,n} - C_{00,n}}{G_{0n}} + C_{\tau\tau,n}. \end{aligned}$$

Since

$$\frac{R_n}{G_n} = \frac{1}{G_n} \left\{ \frac{V_{1,n}}{G_{1n}/G_n} + \frac{V_{0,n}}{G_{0n}/G_n} \right\} = \frac{V_{1,n}}{G_{1n}} + \frac{V_{0,n}}{G_{0n}},$$

this is exactly

$$\sigma_n^2 = \frac{R_n}{G_n} + C_{\tau\tau,n} - \frac{C_{11,n}}{G_{1n}} - \frac{C_{00,n}}{G_{0n}}.$$

Taking expectations in the conditional CR2 identity gives

$$\begin{aligned}\mathbb{E}[\widehat{V}_{\text{CR2},n}] &= \frac{\mathbb{E}_{\alpha} \mathbf{S}_1(\alpha)}{G_{1n}} + \frac{\mathbb{E}_{\alpha} \mathbf{S}_0(\alpha)}{G_{0n}} \\ &= \frac{V_{1,n} - C_{11,n}}{G_{1n}} + \frac{V_{0,n} - C_{00,n}}{G_{0n}}.\end{aligned}$$

Subtracting the variance identity from the CR2 expectation identity, the common terms cancel and therefore

$$\mathbb{E}[\widehat{V}_{\text{CR2},n}] - \sigma_n^2 = -C_{\tau\tau,n}.$$

□

*Proof of Theorem 3.* Set

$$\Gamma := 64M^2(M+1)B^2, \quad C := \Gamma + 1.$$

Then  $C > 0$ . Fix a feasible triangular array satisfying the four displayed assumptions, fix the rowwise Johnson projections, and bring into scope the two rowwise hypotheses that these projections give the centered orthogonal Johnson decomposition and that the corresponding Kneser spectrum formula holds. Index rows by  $r$ , and write  $n_r$  for the row population size,  $G_r$  for the group count,  $G_{1r}$  and  $G_{0r} = G_r - G_{1r}$  for the arm counts,  $N_r = MG_r$ , and  $p_r = G_{1r}/G_r$ . Since  $2 \leq M$ ,  $M > 0$ . The feasibility condition gives

$$G_r \leq MG_r = N_r \leq n_r.$$

Together with Assumption 1, this implies  $n_r \rightarrow \infty$ . Hence

$$\frac{1}{n_r} \rightarrow 0, \quad \frac{M}{n_r} \rightarrow 0, \quad 1 - \frac{M}{n_r} \rightarrow 1.$$

Define

$$\alpha_r := \frac{MG_r}{n_r - M}.$$

Because  $G_{1r} > 0$  and  $G_{1r} < G_r$ , every row has  $G_r \geq 2$ , and therefore  $2M \leq MG_r \leq n_r$ , so the denominator is positive. Also

$$\alpha_r = \frac{MG_r/n_r}{1 - M/n_r} = \frac{N_r/n_r}{1 - M/n_r}.$$

By Assumption 3 and the preceding display,  $\alpha_r \rightarrow \rho$ .

For  $S \in \Omega_{n_r, M}$ , Definition 5 and Assumption 4 give, for each arm  $z$ ,

$$|h_{z,r}(S)| = \left| \frac{1}{M} \sum_{i \in S} Y_{i,r}(z, S) \right| \leq \frac{1}{M} \sum_{i \in S} |Y_{i,r}(z, S)| \leq B.$$

Put

$$d_r(S) := h_{1,r}(S) - h_{0,r}(S).$$

Then

$$|d_r(S)| \leq |h_{1,r}(S)| + |h_{0,r}(S)| \leq 2B.$$

More generally, if  $|u(S)| \leq L$  on  $\Omega_{n_r, M}$ , then every Johnson projection satisfies

$$\|\Pi_{r,k}u\|_r^2 \leq L^2.$$

Indeed, writing  $v = \Pi_{r,k}u$ , orthogonal projection gives

$$\|u\|_r^2 = \|u - v\|_r^2 + \|v\|_r^2,$$

so  $\|v\|_r^2 \leq \|u\|_r^2 = \mathbb{E}_r[u(S)^2] \leq L^2$ . Applying this with  $u = d_r$ ,  $L = 2B$ , and  $k = 1$  yields

$$0 \leq E_{\tau,1,r} = \|\Pi_{r,1}d_r\|_r^2 \leq (2B)^2.$$

Consequently

$$(\rho - \alpha_r)E_{\tau,1,r} \rightarrow 0.$$

Let

$$\mathcal{I}_{\geq 2} := \{k \in \{0, \dots, M\} : 2 \leq k\}$$

and

$$H_r := G_r \sum_{k \in \mathcal{I}_{\geq 2}} \lambda_{r,k} \|\Pi_{r,k}d_r\|_r^2,$$

where  $\lambda_{r,k}$  is the normalized Kneser eigenvalue from [Theorem 1](#). We claim that, for every row,

$$|H_r| \leq \frac{\Gamma}{n_r}.$$

First suppose  $4M \leq n_r$ . Put  $\eta_r := 2M/n_r$ , so  $0 \leq \eta_r \leq 1$ . For any function  $u$  on  $\Omega_{n_r, M}$ , any finite degree set  $\mathcal{I} \subseteq \{0, \dots, M\}$ , any integer threshold  $q$  such that  $q \leq k$  for all  $k \in \mathcal{I}$ , and any envelope  $L$  satisfying  $|u(S)| \leq L$  for all  $S$ , the spectral-sum estimate is

$$\left| \sum_{k \in \mathcal{I}} \lambda_{r,k} \|\Pi_{r,k}u\|_r^2 \right| \leq |\mathcal{I}| \eta_r^q L^2.$$

Indeed, from

$$\lambda_{r,k} = (-1)^k \frac{(M)_k}{(n_r - M)_k},$$

each numerator factor is at most  $M$ , while each denominator factor is at least  $n_r/2$ ; hence

$$|\lambda_{r,k}| \leq \eta_r^k \leq \eta_r^q \quad (k \in \mathcal{I}).$$

The projection-energy bound proved above gives

$$\|\Pi_{r,k}u\|_r^2 \leq L^2,$$

and the triangle inequality therefore gives the displayed spectral-sum estimate.

Applying this estimate with  $\mathcal{I} = \mathcal{I}_{\geq 2}$ ,  $q = 2$ ,  $u = d_r$ , and  $L = 2B$  gives

$$\left| \sum_{k \in \mathcal{I}_{\geq 2}} \lambda_{r,k} \|\Pi_{r,k}d_r\|_r^2 \right| \leq (M+1) \left( \frac{2M}{n_r} \right)^2 (2B)^2.$$

Using  $G_r \leq n_r$ ,

$$|H_r| \leq n_r(M+1) \left( \frac{2M}{n_r} \right)^2 (2B)^2 = \frac{16M^2(M+1)B^2}{n_r} \leq \frac{\Gamma}{n_r}.$$

Now suppose  $4M > n_r$ . Since  $2M \leq n_r$ , the same eigenvalue formula gives

$$|\lambda_{r,k}| = \frac{(M)_k}{(n_r - M)_k} \leq 1 \quad (0 \leq k \leq M),$$

because  $M \leq n_r - M$ . Together with  $\|\Pi_{r,k} d_r\|_r^2 \leq (2B)^2$ , this yields

$$\left| \sum_{k \in \mathcal{I}_{\geq 2}} \lambda_{r,k} \|\Pi_{r,k} d_r\|_r^2 \right| \leq (M+1)(2B)^2.$$

Using  $G_r \leq n_r$  and  $n_r \leq 4M$ ,

$$|H_r| \leq n_r(M+1)(2B)^2 = 4n_r(M+1)B^2 \leq \frac{64M^2(M+1)B^2}{n_r} = \frac{\Gamma}{n_r}.$$

Since  $n_r \rightarrow \infty$ ,  $H_r \rightarrow 0$ .

By [Theorem 1](#), applied to  $d_r$ ,

$$C_{\tau\tau,r} = \sum_{k=1}^M \lambda_{r,k} \|\Pi_{r,k} d_r\|_r^2.$$

The degree-one eigenvalue is

$$\lambda_{r,1} = -\frac{M}{n_r - M}.$$

Splitting off degree one and multiplying by  $G_r$  gives

$$G_r C_{\tau\tau,r} = -\frac{MG_r}{n_r - M} E_{\tau,1,r} + G_r \sum_{k=2}^M \lambda_{r,k} \|\Pi_{r,k} d_r\|_r^2 = -\alpha_r E_{\tau,1,r} + H_r.$$

Therefore

$$G_r C_{\tau\tau,r} + \rho E_{\tau,1,r} = (\rho - \alpha_r) E_{\tau,1,r} + H_r \rightarrow 0.$$

This proves the second displayed limit.

Next fix an arm  $z$ . Let

$$h_{z,r}^\circ(S) := h_{z,r}(S) - \mathbb{E}_r h_{z,r}.$$

The preceding bound on  $h_{z,r}$  implies

$$|\mathbb{E}_r h_{z,r}| \leq B, \quad |h_{z,r}^\circ(S)| \leq 2B.$$

For all sufficiently large  $r$ ,  $4M \leq n_r$ . Applying [Theorem 1](#) to the centered table  $h_{z,r}^\circ$  and using the degree- $\geq 1$  spectral bound gives

$$|C_{zz,r}| \leq (M+1) \frac{2M}{n_r} (2B)^2.$$

Thus

$$C_{11,r} \rightarrow 0, \quad C_{00,r} \rightarrow 0.$$

Since [Assumption 2](#) gives  $p_r \rightarrow p \in (0, 1)$ , also  $1 - p_r \rightarrow 1 - p \in (0, 1)$ . Hence

$$\frac{C_{11,r}}{p_r} + \frac{C_{00,r}}{1 - p_r} \rightarrow 0.$$

This proves the third displayed limit.

It remains to identify the variance limit. The same stable-fraction assumption implies that, eventually,

$$p_r > \frac{p}{2}, \quad 1 - p_r > \frac{1 - p}{2}.$$

Together with  $G_r \rightarrow \infty$ , this yields

$$G_{1r} = p_r G_r \geq 2, \quad G_{0r} = (1 - p_r) G_r \geq 2$$

for all sufficiently large  $r$ . On those rows, [Theorem 2](#) gives

$$\sigma_r^2 = \frac{R_r}{G_r} + C_{\tau\tau,r} - \frac{C_{11,r}}{G_{1r}} - \frac{C_{00,r}}{G_{0r}}.$$

Multiplying by  $G_r$ , and using  $G_{1r} = p_r G_r$  and  $G_{0r} = (1 - p_r) G_r$ , gives the eventual identity

$$G_r \sigma_r^2 - (R_r - \rho E_{\tau,1,r}) = (G_r C_{\tau\tau,r} + \rho E_{\tau,1,r}) - \left( \frac{C_{11,r}}{p_r} + \frac{C_{00,r}}{1 - p_r} \right).$$

Both terms on the right tend to zero by the preceding two steps. Therefore

$$G_r \sigma_r^2 - (R_r - \rho E_{\tau,1,r}) \rightarrow 0,$$

which is the first displayed limit.

Finally, the rowwise bound already proved gives

$$\left| G_r \sum_{k=2}^M \lambda_{r,k} \|\Pi_{r,k} d_r\|_r^2 \right| \leq \frac{\Gamma}{n_r} \leq \frac{C}{n_r},$$

because  $C = \Gamma + 1$  and  $n_r > 0$ . Translating the row index  $r$  back to the theorem's notation gives the asserted bound for every row and completes the proof.  $\square$

*Proof of [Theorem 4](#).* 1. For  $z \in \{0, 1\}$ , set

$$m_{1,n} := G_{1n}, \quad m_{0,n} := G_{0n} := G_n - G_{1n}, \quad \pi_{1,n} := p_n, \quad \pi_{0,n} := 1 - p_n.$$

The dense schedule condition in [Definition 13](#) gives  $G_n \rightarrow \infty$ ,  $p_n \rightarrow p \in (0, 1)$ , and hence

$$m_{1,n} = p_n G_n \rightarrow \infty, \quad m_{0,n} = (1 - p_n) G_n \rightarrow \infty.$$

Consequently  $m_{1,n} \geq 2$  and  $m_{0,n} \geq 2$  for all sufficiently large  $n$ . Also, since  $M G_n \leq n$  and  $M \geq 2$ , the population size tends to infinity, so  $4M \leq n$  eventually.

2. Let  $g_n : \Omega_{n,M} \rightarrow \mathbb{R}$  be any deterministic table with  $|g_n(S)| \leq L$  for all  $S \in \Omega_{n,M}$ , where  $L > 0$ . For  $\omega = (\mathbf{A}_n, \mathbf{Z}_n)$  in the support of the design in [Definition 4](#), define the arm- $z$  selected mean by

$$\bar{g}_{z,n}(\omega) := \frac{1}{m_{z,n}} \sum_{j: Z_{jn}=z} g_n(A_{jn}), \quad \mu_{g,n} := \mathbb{E}_n g_n.$$

The selected mean has expectation  $\mu_{g,n}$ . Indeed, condition on any realized arm-position set of cardinality  $m_{z,n}$ . Under [Definition 4](#), each selected group is uniform on  $\Omega_{n,M}$ , and averaging over the independently chosen arm positions preserves this marginal law.

For the variance, define the centered table and the normalized disjointness operator by

$$g_n^\circ(S) := g_n(S) - \mu_{g,n}, \quad (\mathbf{K}_n u)(S) := \frac{1}{\#\{T \in \Omega_{n,M} : T \cap S = \emptyset\}} \sum_{T \cap S = \emptyset} u(T),$$

for  $S \in \Omega_{n,M}$ . For all sufficiently large rows,  $m_{z,n} \geq 2$  and  $2M \leq n$ . Conditional on the arm positions, two distinct selected groups form a uniform ordered disjoint pair, so

$$\text{Cov}\{g_n(A_{jn}), g_n(A_{\ell n})\} = \langle g_n^\circ, \mathbf{K}_n g_n^\circ \rangle_n, \quad j \neq \ell.$$

Expanding the square of the selected average gives

$$\begin{aligned} \text{Var}\{\bar{g}_{z,n}\} &= \frac{1}{m_{z,n}^2} \{m_{z,n} \text{Var}_n(g_n) + m_{z,n}(m_{z,n} - 1) \langle g_n^\circ, \mathbf{K}_n g_n^\circ \rangle_n\} \\ &= \frac{\text{Var}_n(g_n)}{m_{z,n}} + \left(1 - \frac{1}{m_{z,n}}\right) \langle g_n^\circ, \mathbf{K}_n g_n^\circ \rangle_n. \end{aligned}$$

It remains to bound the covariance term. Since  $|g_n(S)| \leq L$ , uniform averaging gives  $|\mu_{g,n}| \leq L$ , hence  $|g_n^\circ(S)| \leq 2L$  and  $\|g_n^\circ\|_n^2 \leq (2L)^2$ . Applying [Theorem 1](#) to the centered table gives

$$\langle g_n^\circ, \mathbf{K}_n g_n^\circ \rangle_n = \sum_{k=1}^M \lambda_{n,k} \|\Pi_{n,k} g_n^\circ\|_n^2.$$

For rows with  $4M \leq n$ , the falling-factorial expression for  $\lambda_{n,k}$  satisfies, for  $1 \leq k \leq M$ ,

$$|\lambda_{n,k}| \leq \left(\frac{2M}{n}\right)^k \leq \frac{2M}{n}.$$

The first inequality follows by bounding the numerator falling factorial by  $M^k$  and the denominator falling factorial by  $(n/2)^k$ ; the second uses  $2M/n \leq 1$ . Because each  $\Pi_{n,k}$  is an orthogonal projection,

$$\|\Pi_{n,k} g_n^\circ\|_n^2 \leq \|g_n^\circ\|_n^2 \leq (2L)^2.$$

Therefore, for all sufficiently large  $n$ ,

$$|\langle g_n^\circ, \mathbf{K}_n g_n^\circ \rangle_n| \leq (M+1) \frac{2M}{n} (2L)^2.$$

Since  $\text{Var}_n(g_n) \leq L^2$ ,  $m_{z,n} \rightarrow \infty$ , and  $n \rightarrow \infty$ , the displayed variance tends to zero. Chebyshev's inequality therefore gives

$$\bar{g}_{z,n} - \mathbb{E}_n g_n \xrightarrow{p} 0.$$

3. Apply the preceding conclusion first to  $g_n = h_{z,n}$ , using  $|h_{z,n}| \leq B$ , and then to  $g_n = h_{z,n}^2$ , using  $|h_{z,n}^2| \leq B^2$ . With

$$\bar{h}_{z,n}(\omega) := \frac{1}{m_{z,n}} \sum_{j: Z_{jn}=z} h_{z,n}(A_{jn}), \quad \bar{h}_{z,n}^2(\omega) := \frac{1}{m_{z,n}} \sum_{j: Z_{jn}=z} h_{z,n}(A_{jn})^2,$$

we have

$$\bar{h}_{z,n} - \mathbb{E}_n h_{z,n} \xrightarrow{p} 0, \quad \bar{h}_{z,n}^2 - \mathbb{E}_n h_{z,n}^2 \xrightarrow{p} 0.$$

The sample-variance identity gives, eventually,

$$s_{z,n}^2 = \frac{m_{z,n}}{m_{z,n} - 1} \left\{ \bar{h}_{z,n}^2 - \bar{h}_{z,n}^2 \right\}.$$

Because  $m_{z,n}/(m_{z,n} - 1) \rightarrow 1$ , and the selected means are bounded by  $B$ , the square map preserves convergence in probability here. Thus

$$s_{z,n}^2 \xrightarrow{p} \mathbb{E}_n h_{z,n}^2 - (\mathbb{E}_n h_{z,n})^2 = V_{z,n}.$$

4. By [Definition 12](#) and [Definition 11](#),

$$G_n \hat{V}_{\text{CR2},n} - R_n = \frac{s_{1,n}^2 - V_{1,n}}{p_n} + \frac{s_{0,n}^2 - V_{0,n}}{1 - p_n}$$

eventually. Since  $p_n \rightarrow p \in (0, 1)$ , both reciprocal factors are bounded deterministic sequences. Combining this boundedness with the two armwise convergences in the preceding step yields

$$G_n \hat{V}_{\text{CR2},n} - R_n \xrightarrow{p} 0.$$

5. Define, for every  $n$ ,

$$d_n := R_n - \rho E_{\tau,1,n}.$$

The first conclusion of [Theorem 3](#) supplies

$$G_n \sigma_n^2 - d_n \rightarrow 0.$$

Subtracting this deterministic convergence from the stochastic expansion just proved gives

$$G_n (\hat{V}_{\text{CR2},n} - \sigma_n^2) - \rho E_{\tau,1,n} = (G_n \hat{V}_{\text{CR2},n} - R_n) - (G_n \sigma_n^2 - d_n) \xrightarrow{p} 0.$$

6. The scaled-variance nondegeneracy condition in [Definition 13](#) gives

$$\frac{c_\sigma}{2} < G_n \sigma_n^2$$

eventually. Since  $G_n \sigma_n^2 - d_n \rightarrow 0$ , for all sufficiently large  $n$ ,

$$|G_n \sigma_n^2 - d_n| < \frac{c_\sigma}{4}, \quad \text{and hence} \quad \frac{c_\sigma}{4} < d_n.$$

Thus  $d_n$  is eventually positive and bounded away from zero.

7. Set

$$r_n := \frac{R_n}{d_n}.$$

From the first expansion and the eventual lower bound on  $d_n$ ,

$$\frac{G_n \widehat{V}_{\text{CR2},n}}{d_n} - \frac{R_n}{d_n} = (G_n \widehat{V}_{\text{CR2},n} - R_n) d_n^{-1} \xrightarrow{p} 0.$$

Similarly,

$$\frac{G_n \sigma_n^2}{d_n} - 1 = \frac{G_n \sigma_n^2 - d_n}{d_n} \rightarrow 0.$$

The deterministic target  $r_n$  is eventually bounded. To see this, bounded potential outcomes give  $V_{z,n} \leq B^2$ , while  $p_n \rightarrow p \in (0, 1)$  gives

$$\frac{p}{2} < p_n, \quad \frac{1-p}{2} < 1 - p_n$$

eventually. Hence, eventually,

$$0 \leq R_n \leq \frac{B^2}{p/2} + \frac{B^2}{(1-p)/2} =: \overline{R}, \quad |r_n| \leq \overline{R} \frac{4}{c_\sigma}.$$

Let

$$X_n(\omega) := \frac{G_n \widehat{V}_{\text{CR2},n}(\omega)}{d_n}, \quad Y_n := \frac{G_n \sigma_n^2}{d_n}.$$

The preceding displays give  $X_n - r_n \xrightarrow{p} 0$  and  $Y_n \rightarrow 1$ . Since  $Y_n$  is eventually bounded away from zero and  $r_n$  is eventually bounded,

$$\frac{X_n}{Y_n} - r_n = \frac{(X_n - r_n) + r_n(1 - Y_n)}{Y_n} \xrightarrow{p} 0.$$

Because  $X_n/Y_n = \widehat{V}_{\text{CR2},n}/\sigma_n^2$  on the same eventual rows, this gives

$$\frac{\widehat{V}_{\text{CR2},n}}{\sigma_n^2} \xrightarrow{p} r_n = \frac{R_n}{R_n - \rho E_{\tau,1,n}}.$$

8. Define

$$q_n := \frac{\rho E_{\tau,1,n}}{d_n} = \frac{\rho E_{\tau,1,n}}{R_n - \rho E_{\tau,1,n}}.$$

For all sufficiently large  $n$ ,  $d_n \neq 0$ , and hence

$$r_n = \frac{R_n}{d_n} = \frac{d_n + \rho E_{\tau,1,n}}{d_n} = 1 + q_n.$$

Moreover  $\rho \geq 0$  by the sampling-fraction condition, and  $E_{\tau,1,n} \geq 0$  because it is a squared norm. Since  $d_n > 0$  eventually,

$$r_n = 1 + q_n \geq 1$$

eventually.

9. Fix  $\varepsilon > 0$ . On every row for which  $r_n \geq 1$ ,

$$\left\{ \frac{\widehat{V}_{\text{CR2},n}}{\sigma_n^2} < 1 - \varepsilon \right\} \subseteq \left\{ \varepsilon \leq \left| \frac{\widehat{V}_{\text{CR2},n}}{\sigma_n^2} - r_n \right| \right\}.$$

The right-hand probability tends to zero because  $\widehat{V}_{\text{CR2},n}/\sigma_n^2 \xrightarrow{p} r_n$ . Therefore

$$\Pr\left\{ \widehat{V}_{\text{CR2},n}/\sigma_n^2 < 1 - \varepsilon \right\} \rightarrow 0.$$

10. It remains to prove the equivalence. Suppose first that

$$\frac{\widehat{V}_{\text{CR2},n}}{\sigma_n^2} \xrightarrow{p} 1.$$

Together with

$$\frac{\widehat{V}_{\text{CR2},n}}{\sigma_n^2} \xrightarrow{p} r_n,$$

uniqueness of deterministic probability limits gives  $r_n \rightarrow 1$ : indeed, if  $|r_n - 1| \geq \eta$ , the triangle inequality forces at least one of

$$\left| \frac{\widehat{V}_{\text{CR2},n}}{\sigma_n^2} - r_n \right| \geq \frac{\eta}{2}, \quad \left| \frac{\widehat{V}_{\text{CR2},n}}{\sigma_n^2} - 1 \right| \geq \frac{\eta}{2}$$

to occur. Since  $q_n = r_n - 1$  eventually, this yields  $q_n \rightarrow 0$ , namely

$$\frac{\rho E_{\tau,1,n}}{R_n - \rho E_{\tau,1,n}} \rightarrow 0.$$

11. Conversely, assume

$$\frac{\rho E_{\tau,1,n}}{R_n - \rho E_{\tau,1,n}} \rightarrow 0.$$

That is,  $q_n \rightarrow 0$ . Since  $r_n = 1 + q_n$  eventually,  $r_n \rightarrow 1$ . Retargeting the already proved convergence

$$\frac{\widehat{V}_{\text{CR2},n}}{\sigma_n^2} \xrightarrow{p} r_n$$

to the deterministic limit 1 gives

$$\frac{\widehat{V}_{\text{CR2},n}}{\sigma_n^2} \xrightarrow{p} 1.$$

This proves the asserted if and only if. □

*Proof of Theorem 5.* Since  $M \geq 2$ , in particular  $M > 0$ . We prove the three asserted conclusions in the order in which they are stated.

1. **Sparse dense arrays.** Let  $A$  be a feasible triangular schedule array with common group size  $M$ , and suppose that  $A \in \mathcal{C}_{\text{dense}}$  in Definition 13 with limiting grouped-unit fraction  $\rho = 0$ . Equip each row with the canonical Johnson projections. The canonical decomposition and Kneser-spectrum hypotheses required by Theorem 4 are supplied by Lemmas 2 and 3. Applying Theorem 4 with  $\rho = 0$ , its ratio criterion becomes

$$\frac{0 \cdot E_{\tau,1,n}}{R_n - 0 \cdot E_{\tau,1,n}} = 0 \quad \text{for every row } n,$$

so the criterion converges to 0. Hence

$$\frac{\widehat{V}_{\text{CR2},n}}{\sigma_n^2} \xrightarrow{p} 1.$$

2. **Birthday counts.** Define

$$G_n^{\text{bd}} = \left\lfloor \frac{n^{3/4}}{M} \right\rfloor, \quad N_n^{\text{bd}} = M G_n^{\text{bd}}.$$

For  $n \geq 1$ ,

$$N_n^{\text{bd}} = M \left\lfloor \frac{n^{3/4}}{M} \right\rfloor \leq M \frac{n^{3/4}}{M} = n^{3/4} \leq n,$$

so  $N_n^{\text{bd}} \leq n$  eventually. For  $n > 0$ ,

$$\frac{N_n^{\text{bd}}}{n} = \frac{\lfloor n^{3/4}/M \rfloor}{n^{3/4}/M} n^{-1/4}.$$

Since  $n^{3/4}/M \rightarrow \infty$ , the floor ratio tends to 1, while  $n^{-1/4} \rightarrow 0$ . Therefore

$$\frac{N_n^{\text{bd}}}{n} \rightarrow 0.$$

For divergence of the squared count, the elementary floor bound gives

$$n^{3/4} - M \leq M \left\lfloor \frac{n^{3/4}}{M} \right\rfloor = N_n^{\text{bd}}.$$

Eventually  $n^{3/4} \geq 2M$ , and hence

$$\frac{n^{3/4}}{2} \leq N_n^{\text{bd}}.$$

Consequently, for all sufficiently large  $n$ ,

$$\frac{(N_n^{\text{bd}})^2}{n} \geq \frac{(n^{3/4}/2)^2}{n} = \frac{1}{4} n^{1/2} \rightarrow \infty.$$

3. **Birthday-aligned arrays.** Let  $A$  be a feasible triangular schedule array satisfying the rowwise birthday alignment, Assumptions 1, 2, 4 and 5. For row  $r \in \mathbb{N}$ , write  $G_r$  for the

group count,  $n_r$  for the population size, and  $N_r = MG_r$  for the grouped-unit count. From [Assumption 1](#),  $G_r \rightarrow \infty$ , and feasibility gives

$$G_r \leq MG_r = N_r \leq n_r.$$

Thus  $n_r \rightarrow \infty$ . The alignment condition says

$$N_r = M \left\lfloor \frac{n_r^{3/4}}{M} \right\rfloor = N_{n_r}^{\text{bd}},$$

and the birthday-count limit from the previous step therefore yields

$$\frac{N_r}{n_r} = \frac{N_{n_r}^{\text{bd}}}{n_r} \rightarrow 0.$$

Together with  $0 \leq \rho \leq 1$ , this is exactly [Assumption 3](#) with limiting grouped-unit fraction  $\rho = 0$ . The assumed group-count growth, stable treatment fraction, bounded potential outcomes, and scaled-variance nondegeneracy then place  $A$  in  $\mathcal{C}_{\text{dense}}$  with  $\rho = 0$  by [Definition 13](#). The first step applies and gives

$$\frac{\hat{V}_{\text{CR2},n}}{\sigma_n^2} \xrightarrow{p} 1.$$

□

*Proof of [Proposition 1](#).* 1. Write  $T_n \subseteq \{1, \dots, G_n\}$  for the realized treated-group set, so  $|T_n| = G_{1n}$ , and write  $C_n = T_n^c$ , so  $|C_n| = G_{0n} = G_n - G_{1n}$ . For a group  $g$ , define

$$G_{g,n} := \begin{cases} G_{1n}, & g \in T_n, \\ G_{0n}, & g \in C_n. \end{cases}$$

The assumptions  $G_{1n} \geq 2$  and  $G_{0n} \geq 2$  imply  $G_{1n} \leq G_n$ ,  $G_n > 0$ ,  $M > 0$ ,  $G_{0n} > 0$ , and the real denominators  $G_{1n}$  and  $G_{0n}$  are nonzero. The two versioned source identities give

$$\text{bread}(\text{fit}) = G_n M (X^\top X)^{-1}$$

for the unweighted full-rank least-squares fit [[Zeileis and Lumley, 2026](#), R/bread.R, lines 10–15], and

$$\Sigma = \tilde{B} \left\{ \sum_{g=1}^{G_n} X_g^\top A_g e_g e_g^\top A_g^\top X_g \right\} \tilde{B}$$

for the CR2 call with omitted `target`, `inverse_var`, and `form` [[Pustejovsky, 2026](#), R/lm.R, lines 47–52 and 70–72; R/S3-methods.R, lines 21–33 and 63–65; R/clubSandwich.R, lines 167–175, 216–225, and 237–287; R/CR-adjustments.R, lines 5–7 and 22–42]. The stated bread scaling therefore yields

$$\tilde{B} = (X^\top X)^{-1}.$$

2. The fitted design has group block

$$X_g = \mathbf{1}_M x_g^\top, \quad x_g = \begin{cases} (1, 1)^\top, & g \in T_n, \\ (1, 0)^\top, & g \in C_n, \end{cases}$$

where  $\mathbf{1}_M$  is the  $M$ -vector of ones. Summing the block cross-products gives

$$X^\top X = \sum_{g=1}^{G_n} X_g^\top X_g = M \begin{pmatrix} G_n & G_{1n} \\ G_{1n} & G_{1n} \end{pmatrix}.$$

Since the determinant is  $M^2 G_{1n} G_{0n} > 0$ , inversion gives the entries used below:

$$\tilde{B}_{10} = \tilde{B}_{01} = -\frac{1}{MG_{0n}}, \quad \tilde{B}_{11} = \frac{1}{MG_{0n}} + \frac{1}{MG_{1n}}, \quad \tilde{B}_{00} = \frac{1}{MG_{0n}}.$$

3. For every group  $g$ , the leverage block is

$$H_g = X_g \tilde{B} X_g^\top.$$

Using the preceding inverse entries and the two possible values of  $x_g$ ,

$$(H_g)_{ij} = \frac{1}{MG_{g,n}}, \quad 1 \leq i, j \leq M.$$

Thus

$$I_M - H_g = I_M - \frac{1}{G_{g,n}} P_M, \quad P_M := \frac{1}{M} \mathbf{1}_M \mathbf{1}_M^\top.$$

Let  $q = G_{g,n}$ . Since  $q \geq 2$ , the symmetric positive-definite matrix  $A_g$  satisfying  $A_g A_g (I_M - H_g) = I_M$  is determined by its action on the span of  $\mathbf{1}_M$  and on its orthogonal complement. For

$$u := \mathbf{1}_M, \quad v_j := e_j - \frac{1}{M} u,$$

one has

$$(I_M - H_g) v_j = v_j, \quad (I_M - H_g) u = \frac{q-1}{q} u.$$

The identity  $A_g A_g (I_M - H_g) = I_M$ , together with positive definiteness, therefore gives

$$A_g v_j = v_j, \quad A_g u = \sqrt{\frac{q}{q-1}} u.$$

Combining these two relations column by column,

$$(A_g)_{ij} = \mathbf{1}\{i = j\} + \frac{\sqrt{G_{g,n}/(G_{g,n} - 1)} - 1}{M}.$$

4. Summing the displayed entries down any column gives

$$\sum_{i=1}^M (A_g)_{ij} = \sqrt{\frac{G_{g,n}}{G_{g,n} - 1}}.$$

Consequently

$$\mathbf{1}_M^\top A_g e_g = \sum_{i=1}^M (A_g e_g)_i = \sqrt{\frac{G_{g,n}}{G_{g,n} - 1}} \sum_{j=1}^M e_{gj}.$$

The residual-mean condition in the statement says

$$\frac{1}{M} \sum_{j=1}^M e_{gj} = H_{g,n}^{\text{obs}} - \bar{H}_{z_g,n},$$

where

$$H_{g,n}^{\text{obs}} := \text{obsGroupMean}(Y, \mathbf{A}_n, \mathbf{Z}_n; g), \quad \bar{H}_{z_g,n} := \text{armObsMean}(Y, \mathbf{A}_n, \mathbf{Z}_n; z),$$

and  $z_g = 1$  on  $T_n$ ,  $z_g = 0$  on  $C_n$ . Hence

$$\sum_{j=1}^M e_{gj} = M \{H_{g,n}^{\text{obs}} - \bar{H}_{z_g,n}\}.$$

5. Define the treatment-coordinate adjusted cluster score by

$$q_{g,n} := \left[ \tilde{B} X_g^\top A_g e_g \right]_1.$$

Using  $X_g = \mathbf{1}_M x_g^\top$ , the inverse entries in Step 2, and the residual sum from Step 4 gives the two cases

$$q_{g,n} = \begin{cases} \sqrt{\frac{G_{1n}}{G_{1n} - 1}} \frac{H_{g,n}^{\text{obs}} - \bar{H}_{1,n}}{G_{1n}}, & g \in T_n, \\ -\sqrt{\frac{G_{0n}}{G_{0n} - 1}} \frac{H_{g,n}^{\text{obs}} - \bar{H}_{0,n}}{G_{0n}}, & g \in C_n. \end{cases}$$

6. Because  $\tilde{B}$  is symmetric, each summand in the CR2 sandwich is a rank-one outer product after premultiplication and postmultiplication by  $\tilde{B}$ . Indeed, with

$$P_g := \tilde{B} X_g^\top A_g,$$

$$\tilde{B} X_g^\top A_g e_g e_g^\top A_g^\top X_g \tilde{B} = P_g e_g e_g^\top P_g^\top = (P_g e_g)(P_g e_g)^\top.$$

Therefore

$$\Sigma = \sum_{g=1}^{G_n} \left( \tilde{B} X_g^\top A_g e_g \right) \left( \tilde{B} X_g^\top A_g e_g \right)^\top,$$

and its treatment-coordinate entry is

$$\Sigma_{11} = \sum_{g=1}^{G_n} q_{g,n}^2.$$

7. Squaring the two cases in Step 5 and using  $G_{1n}, G_{0n} \geq 2$  gives

$$\Sigma_{11} = \sum_{g \in T_n} \frac{(H_{g,n}^{\text{obs}} - \bar{H}_{1,n})^2}{(G_{1n} - 1)G_{1n}} + \sum_{g \in C_n} \frac{(H_{g,n}^{\text{obs}} - \bar{H}_{0,n})^2}{(G_{0n} - 1)G_{0n}}.$$

Equivalently,

$$\Sigma_{11} = \frac{1}{G_{1n}} \frac{\sum_{g \in T_n} (H_{g,n}^{\text{obs}} - \bar{H}_{1,n})^2}{G_{1n} - 1} + \frac{1}{G_{0n}} \frac{\sum_{g \in C_n} (H_{g,n}^{\text{obs}} - \bar{H}_{0,n})^2}{G_{0n} - 1}.$$

By [Definition 12](#), the two displayed fractions are  $s_{1,n}^2$  and  $s_{0,n}^2$ , respectively, and hence

$$\Sigma_{11} = \frac{s_{1,n}^2}{G_{1n}} + \frac{s_{0,n}^2}{G_{0n}} = \hat{V}_{\text{CR2},n}.$$

□

*Proof of [Proposition 2](#).* Let

$$\Sigma_8 : \Omega_{8,2} \rightarrow \mathbb{R}, \quad \Sigma_8(A) = \sum_{u \in A} a_u.$$

By [Definitions 5](#) and [17](#),

$$h_{1,8}(A) = \frac{\Sigma_8(A)}{2}, \quad h_{0,8}(A) = 0 \quad (A \in \Omega_{8,2}).$$

The vector  $\mathbf{a}_8$  has four +1 entries and four −1 entries, so its population mean is zero and its population variance is one. The usual fixed-size sample-mean identity for a uniform two-subset gives

$$\mathbb{E}_8 h_{1,8} = 0, \quad V_{1,8} = \text{Var}_{A \sim \text{Unif}(\Omega_{8,2})} \left\{ \frac{\Sigma_8(A)}{2} \right\} = \frac{8-2}{2(8-1)} \cdot 1 = \frac{3}{7}.$$

Since  $h_{0,8}$  is identically zero, its centered table is zero, hence

$$V_{0,8} = 0, \quad C_{00,8} = C_{10,8} = C_{01,8} = 0.$$

For the treated ordered-disjoint covariance, write the eight-unit population and its two-subset slice as

$$\mathcal{U}_8 = \{1, \dots, 8\}, \quad \Omega_{8,2} = \{A \subseteq \mathcal{U}_8 : |A| = 2\}.$$

Use ordered pairs

$$(A, C) \in \Omega_{8,2} \times \Omega_{8,2}, \quad A \cap C = \emptyset.$$

There are

$$\binom{8}{2} \binom{6}{2} = 420$$

such ordered pairs. For fixed  $A$ , each unit in the six-point complement of  $A$  appears in five two-subsets of that complement, and  $\sum_{u \in \mathcal{U}_8} a_u = 0$ . Therefore

$$\sum_{\substack{C \in \Omega_{8,2} \\ C \cap A = \emptyset}} \Sigma_8(C) = 5 \sum_{u \notin A} a_u = -5 \Sigma_8(A).$$

Also  $\sum_{A \in \Omega_{8,2}} \Sigma_8(A)^2 = 48$ , because the six same-positive pairs and six same-negative pairs contribute 4 each and the sixteen mixed pairs contribute 0. Thus

$$\sum_{\substack{A, C \in \Omega_{8,2} \\ A \cap C = \emptyset}} \Sigma_8(A) \Sigma_8(C) = -5 \sum_{A \in \Omega_{8,2}} \Sigma_8(A)^2 = -240.$$

Using  $\mathbb{E}_8 h_{1,8} = 0$ , this gives

$$C_{11,8} = \frac{1}{420} \sum_{\substack{A, C \in \Omega_{8,2} \\ A \cap C = \emptyset}} \frac{\Sigma_8(A)}{2} \frac{\Sigma_8(C)}{2} = \frac{-240}{1680} = -\frac{1}{7}.$$

Consequently

$$C_{\tau\tau,8} = C_{11,8} + C_{00,8} - 2C_{10,8} = -\frac{1}{7}.$$

Define the arm-difference table

$$\Delta_8 : \Omega_{8,2} \rightarrow \mathbb{R}, \quad \Delta_8(A) = h_{1,8}(A) - h_{0,8}(A).$$

Since  $h_{0,8} = 0$ ,  $\Delta_8 = h_{1,8}$ , so

$$\mathbb{E}_8 \Delta_8 = 0, \quad \text{Var}_{\mathbb{E}_8}(\Delta_8) = \frac{3}{7}.$$

Set

$$\eta_1 = \|\Pi_{8,1} \Delta_8\|_8^2, \quad \eta_2 = \|\Pi_{8,2} \Delta_8\|_8^2.$$

The centered Johnson decomposition assumed in the proposition gives

$$\Delta_8 = \Pi_{8,1} \Delta_8 + \Pi_{8,2} \Delta_8$$

and the two summands are orthogonal. Hence

$$\eta_1 + \eta_2 = \|\Delta_8\|_8^2 = \frac{3}{7}.$$

By [Theorem 1](#), instantiated with  $n = 8$  and  $M = 2$ ,

$$C_{\tau\tau,8} = \lambda_{8,1} \eta_1 + \lambda_{8,2} \eta_2.$$

The Kneser eigenvalues are

$$\lambda_{8,1} = -\frac{2}{6} = -\frac{1}{3}, \quad \lambda_{8,2} = \frac{2 \cdot 1}{6 \cdot 5} = \frac{1}{15}.$$

Using  $C_{\tau\tau,8} = -1/7$ ,

$$-\frac{1}{3} \eta_1 + \frac{1}{15} \eta_2 = -\frac{1}{7}.$$

Together with  $\eta_1 + \eta_2 = 3/7$ , this linear system gives

$$\eta_1 = \frac{3}{7}.$$

By Definition 11,  $\eta_1 = E_{\tau,1,8}$ , so

$$E_{\tau,1,8} = \frac{3}{7}.$$

It remains to evaluate the design variance and the CR2 expectation. Here

$$G_8 = 4, \quad G_{1,8} = 2, \quad G_{0,8} = 2, \quad p_8 = \frac{1}{2}.$$

By Definition 11,

$$R_8 = \frac{V_{1,8}}{p_8} + \frac{V_{0,8}}{1-p_8} = \frac{3/7}{1/2} + 0 = \frac{6}{7}.$$

Applying Theorem 2 to the feasible eight-unit design gives

$$\sigma_8^2 = \frac{R_8}{G_8} + C_{\tau\tau,8} - \frac{C_{11,8}}{G_{1,8}} - \frac{C_{00,8}}{G_{0,8}}.$$

Substituting the displayed values,

$$\sigma_8^2 = \frac{6/7}{4} - \frac{1}{7} - \frac{-1/7}{2} - 0 = \frac{1}{7}.$$

The same result in Theorem 2, now for the statistic in Definition 12, gives

$$\mathbb{E}[\widehat{V}_{\text{CR2},8}] = \frac{V_{1,8} - C_{11,8}}{G_{1,8}} + \frac{V_{0,8} - C_{00,8}}{G_{0,8}}.$$

Therefore

$$\mathbb{E}[\widehat{V}_{\text{CR2},8}] = \frac{3/7 + 1/7}{2} + 0 = \frac{2}{7}.$$

□

*Proof of Proposition 3.* 1. Fix a row  $n$  and a realized design  $w = (\mathbf{A}_n, \mathbf{Z}_n)$ . Under the independent-sign prior, define the selected sign field

$$\xi_i^w := \xi_{i,\alpha_w(i)}, \quad \alpha_w(i) := \mathbf{1}\{i \text{ is observed in a treated realized group under } w\}.$$

Because the realized groups are pairwise disjoint, every observed unit has  $\alpha_w(i)$  equal to its realized treatment arm. Hence the observation produced from the independent schedule is exactly the observation produced from the common-sign schedule with signs  $(\xi_i^w)_{i \in \mathcal{U}_n}$ . For fixed  $w$ , the vector  $(\xi_i^w)_i$  has the same product Rademacher law as the common-sign vector. Finite Fubini then gives, for every statistic  $\varphi$ ,

$$\mathbb{E}_{\mathbb{H}_n^{\text{same}}} \mathbb{E}_{\mathbf{Z}_n} \{\varphi(\mathcal{O}_n)\} = \mathbb{E}_{\mathbf{Z}_n} \mathbb{E}_{\mathbb{H}_n^{\text{same}}} \{\varphi(\mathcal{O}_n)\} = \mathbb{E}_{\mathbf{Z}_n} \mathbb{E}_{\mathbb{H}_n^{\text{ind}}} \{\varphi(\mathcal{O}_n)\} = \mathbb{E}_{\mathbb{H}_n^{\text{ind}}} \mathbb{E}_{\mathbf{Z}_n} \{\varphi(\mathcal{O}_n)\}.$$

2. For any deterministic vector  $\mathbf{x} = (x_i)_{i \in \mathcal{U}_n}$ , write

$$\bar{x}_n := \frac{1}{n} \sum_{i \in \mathcal{U}_n} x_i, \quad \text{PV}_n(\mathbf{x}) := \frac{1}{n-1} \sum_{i \in \mathcal{U}_n} (x_i - \bar{x}_n)^2, \quad m_{\mathbf{x}}(Q) := \frac{1}{M} \sum_{i \in Q} x_i$$

for  $Q \in \Omega_{n,M}$ . Sampling without replacement gives

$$\text{Var}_{Q \sim \text{Unif}(\Omega_{n,M})} \{m_{\mathbf{x}}(Q)\} = \left( \frac{1}{M} - \frac{1}{n} \right) \text{PV}_n(\mathbf{x}).$$

If  $x_i \in \{-1, 1\}$ , then

$$\text{PV}_n(\mathbf{x}) = \frac{n}{n-1} (1 - \bar{x}_n^2).$$

Under the common-sign prior,  $\mathbb{E}\bar{\xi}_n = 0$  and

$$\text{Var}(\bar{\xi}_n) = \frac{1}{n},$$

so  $\bar{\xi}_n \rightarrow 0$  in probability. Since [Assumption 1](#) implies  $n \rightarrow \infty$ ,

$$\left( \frac{1}{M} - \frac{1}{n} \right) \frac{n}{n-1} \rightarrow \frac{1}{M},$$

and therefore, for each arm  $z$ ,

$$V_{z,n} \xrightarrow{p} \frac{1}{M} \quad \text{under } \mathbb{H}_n^{\text{same}}.$$

3. The same calculation applies arm by arm under  $\mathbb{H}_n^{\text{ind}}$ . For each fixed arm  $z$ ,

$$\bar{\xi}_{z,n} := \frac{1}{n} \sum_{i \in \mathcal{U}_n} \xi_{i,z}, \quad \mathbb{E}\bar{\xi}_{z,n} = 0, \quad \text{Var}(\bar{\xi}_{z,n}) = \frac{1}{n},$$

so  $\bar{\xi}_{z,n} \rightarrow 0$  in probability and

$$V_{z,n} \xrightarrow{p} \frac{1}{M} \quad \text{under } \mathbb{H}_n^{\text{ind}}.$$

4. Under  $\mathbb{H}_n^{\text{same}}$ , the two arm tables coincide pointwise:

$$h_{1,n}(Q) - h_{0,n}(Q) = 0 \quad (Q \in \Omega_{n,M}).$$

By linearity of the Johnson projection,

$$E_{\tau,1,n} = \|\Pi_{n,1}(h_{1,n} - h_{0,n})\|_n^2 = \|\Pi_{n,1}0\|_n^2 = 0.$$

Thus  $E_{\tau,1,n} \xrightarrow{p} 0$  under  $\mathbb{H}_n^{\text{same}}$ .

5. Under  $\mathbb{H}_n^{\text{ind}}$ , put

$$d_i := \xi_{i,1} - \xi_{i,0}, \quad \bar{d}_n := \frac{1}{n} \sum_i d_i, \quad \bar{c}_n := \frac{1}{n} \sum_i \xi_{i,1} \xi_{i,0}.$$

The arm-difference table is the slice sample mean

$$h_{1,n}(Q) - h_{0,n}(Q) = m_{\mathbf{d}}(Q).$$

This function is a linear combination of degree-one inclusion indicators; after subtracting its slice mean, it is orthogonal to constants and lies in  $\mathcal{H}_{n,1}$ . Hence the degree-one projection equals the centered sample mean, and

$$E_{\tau,1,n} = \left( \frac{1}{M} - \frac{1}{n} \right) \text{PV}_n(\mathbf{d}).$$

Since

$$\text{PV}_n(\mathbf{d}) = \frac{n}{n-1} (2 - 2\bar{c}_n - \bar{d}_n^2),$$

while  $\bar{c}_n \rightarrow 0$ ,  $\bar{\xi}_{1,n} \rightarrow 0$ , and  $\bar{\xi}_{0,n} \rightarrow 0$  in probability, we get

$$E_{\tau,1,n} \xrightarrow{p} \frac{2}{M} \quad \text{under } \mathbb{H}_n^{\text{ind}}.$$

6. By [Assumption 2](#),

$$p_n \rightarrow p, \quad 0 < p < 1,$$

and hence

$$\frac{1}{p_n} \rightarrow \frac{1}{p}, \quad \frac{1}{1-p_n} \rightarrow \frac{1}{1-p}.$$

Consequently, for either prior,

$$\Theta_n := \frac{V_{1,n}}{p_n} + \frac{V_{0,n}}{1-p_n} \xrightarrow{p} \frac{1}{M} \left( \frac{1}{p} + \frac{1}{1-p} \right) = \frac{1}{Mp(1-p)}.$$

7. Since  $0 < G_{1n} < G_n$ , every row has  $G_n \geq 2$ , and therefore  $2M \leq MG_n \leq n$ . The degree-one Kneser eigenvalue in [Theorem 1](#) is

$$\lambda_{n,1} = -\frac{M}{n-M}.$$

Thus  $\lambda_{n,1} \rightarrow 0$ . Writing  $N_n = MG_n$  for the grouped-unit count and using [Assumption 3](#),

$$G_n \lambda_{n,1} = -\frac{MG_n}{n-M} = -\frac{N_n/n}{1-M/n} \rightarrow -\rho.$$

8. The exact rowwise variance identity is used on the eventual rows where both treatment arms contain at least two realized groups. These rows are eventual because  $p_n \rightarrow p \in (0, 1)$  and  $G_n \rightarrow \infty$ : eventually  $p_n > p/2$  and  $G_n > 4/p$ , forcing  $G_{1n} \geq 2$ , and eventually  $p_n < (1+p)/2$  and  $G_n > 4/(1-p)$ , forcing  $G_{0n} \geq 2$ .

On those rows, [Theorems 1](#) and [2](#) give

$$G_n \sigma_n^2 = \Theta_n + G_n \lambda_{n,1} E_{\tau,1,n} - \lambda_{n,1} \Theta_n = (1 - \lambda_{n,1}) \Theta_n + G_n \lambda_{n,1} E_{\tau,1,n}.$$

Combining the convergence statements above yields

$$G_n \sigma_n^2 \xrightarrow{p} \frac{1}{Mp(1-p)} \quad \text{under } \mathbb{H}_n^{\text{same}},$$

and

$$G_n \sigma_n^2 \xrightarrow{p} \frac{1}{Mp(1-p)} - \frac{2\rho}{M} \quad \text{under } \mathbb{H}_n^{\text{ind}}.$$

9. The separation is positive because  $M > 0$  and  $\rho > 0$ :

$$0 < \frac{2\rho}{M}.$$

Also [Assumption 3](#) gives  $\rho \leq 1$ , and  $p(1-p) \leq 1/4$  follows from  $(p-1/2)^2 \geq 0$ . Since  $p(1-p) > 0$ ,

$$\frac{1}{Mp(1-p)} - \frac{2\rho}{M} = \frac{1}{M} \left( \frac{1}{p(1-p)} - 2\rho \right) \geq \frac{1}{M} (4 - 2\rho) \geq \frac{2}{M}.$$

10. Finally, if  $X_n \rightarrow c$  in probability and  $t < c$ , then

$$\Pr\{X_n < t\} \leq \Pr\{|X_n - c| > c - t\} \rightarrow 0, \quad \Pr\{X_n \geq t\} \rightarrow 1.$$

Apply this with  $X_n = G_n \sigma_n^2$ . The same-sign limit exceeds the independent-sign limit by  $2\rho/M$ , so every

$$0 < c_\sigma < \frac{1}{Mp(1-p)} - \frac{2\rho}{M}$$

lies below both limits. Therefore

$$\Pr_{\mathbb{H}_n^{\text{same}}} \{G_n \sigma_n^2 \geq c_\sigma\} \rightarrow 1, \quad \Pr_{\mathbb{H}_n^{\text{ind}}} \{G_n \sigma_n^2 \geq c_\sigma\} \rightarrow 1.$$

□

*Proof of Theorem 6.* 1. Define

$$d_{\text{same}} = \frac{1}{Mp(1-p)}, \quad d_{\text{ind}} = d_{\text{same}} - \frac{2\rho}{M}.$$

By Proposition 3, the unconditioned same-sign and independent-arm one-realization mixtures have identical bounded-test expectations, and the scaled exact variances satisfy

$$G_n \sigma_n^2(Y_n^{\text{same}}) \xrightarrow{p} d_{\text{same}}, \quad G_n \sigma_n^2(Y_n^{\text{ind}}) \xrightarrow{p} d_{\text{ind}}$$

under their respective Rademacher schedule priors. The same result also gives  $0 < 2\rho/M$ , hence

$$d_{\text{ind}} < d_{\text{same}}.$$

The hypothesis on  $c_\sigma$  gives  $c_\sigma < d_{\text{ind}}$ , and therefore also  $c_\sigma < d_{\text{same}}$ .

2. We extract deterministic high-probability supports from these two convergence-in-probability statements. For  $j \in \{\text{same}, \text{ind}\}$ , let  $d_j$  denote the corresponding limit and let

$$X_{j,n}(\omega) = G_n \sigma_n^2(Y_n^{j,\omega})$$

for a schedule-prior sign configuration  $\omega$ . Since  $X_{j,n} \rightarrow d_j$  in probability, for each integer  $k \geq 1$  choose  $N_{j,k}$  such that, whenever  $n \geq N_{j,k}$ ,

$$\mathbb{H}_n^j \{|X_{j,n} - d_j| \geq k^{-1}\} < k^{-1}.$$

Let  $\ell_j(n)$  be the largest  $k \leq n$  with  $N_{j,k} \leq n$ , using  $\ell_j(n) = 1$  when the set is empty, and set

$$e_{j,n} = \ell_j(n)^{-1}, \quad \Gamma_n^j = \{\omega : |X_{j,n}(\omega) - d_j| < e_{j,n}\}.$$

Then  $\ell_j(n) \rightarrow \infty$ , so  $e_{j,n} \rightarrow 0$ , and

$$\mathbb{H}_n^j(\Gamma_n^j) \rightarrow 1, \quad \sup_{\omega \in \Gamma_n^j} |G_n \sigma_n^2(Y_n^{j,\omega}) - d_j| \rightarrow 0.$$

These are the supports  $\Gamma_n^{\text{same}}$  and  $\Gamma_n^{\text{ind}}$ .

3. Every full diagonal selection from either support belongs to  $\mathcal{C}_{\text{dense}}$ . Indeed, for the same-sign prior a selected row has the additive form

$$Y_{i,n}^{\text{same},\xi}(z, A) = \xi_i, \quad \xi_i \in \{-1, 1\},$$

and for the independent-arm prior it has the additive form

$$Y_{i,n}^{\text{ind},\zeta}(z, A) = \zeta_{i,z}, \quad \zeta_{i,z} \in \{-1, 1\}.$$

Thus  $|Y_{i,n}(z, A)| = 1 \leq B$ . The original design skeleton supplies the group-count, treated-fraction, and sampling-fraction conditions in [Definition 13](#). For any diagonal selection with  $\omega_n \in \Gamma_n^j$  for every  $n$ ,

$$G_n \sigma_n^2(Y_n^{j,\omega_n}) \rightarrow d_j,$$

because the support error is bounded by  $e_{j,n} \rightarrow 0$ . Since  $c_\sigma < d_j$ , this convergence gives

$$c_\sigma \leq \liminf_{n \rightarrow \infty} G_n \sigma_n^2(Y_n^{j,\omega_n}).$$

Hence all defining conditions of  $\mathcal{C}_{\text{dense}}$  hold for each such diagonal selection.

4. Since  $\mathbb{H}_n^j(\Gamma_n^j) \rightarrow 1$  for  $j \in \{\text{same}, \text{ind}\}$ , both conditioning masses are positive for all sufficiently large  $n$ . On those rows, let  $\mathbb{E}_{n,j}^\Gamma$  denote expectation under the one-realization mixture obtained by first conditioning  $\mathbb{H}_n^j$  on  $\Gamma_n^j$  and then applying the random partition and treatment design, and define

$$\Delta_n = \sup_{0 \leq \varphi \leq 1} |\mathbb{E}_{n,\text{same}}^\Gamma \varphi(\mathcal{O}_n) - \mathbb{E}_{n,\text{ind}}^\Gamma \varphi(\mathcal{O}_n)|.$$

For the finitely many remaining rows, take  $\Delta_n = 0$ ; this totalization convention has no effect on the limiting assertion. For any bounded test  $0 \leq \varphi \leq 1$ , conditioning a finite prior on a set  $\Gamma$  of positive prior probability changes its expectation by at most twice the complement probability:

$$|\mathbb{E}\{\varphi \mid \Gamma\} - \mathbb{E}\varphi| \leq 2\{1 - \mathbb{H}(\Gamma)\}.$$

Indeed, writing  $m = \mathbb{H}(\Gamma) > 0$ , the conditioned expectation is

$$\mathbb{E}\{\varphi \mid \Gamma\} = \frac{1}{m} \mathbb{E}\{\mathbf{1}_\Gamma \varphi\},$$

and hence

$$\begin{aligned} |\mathbb{E}\{\varphi \mid \Gamma\} - \mathbb{E}\varphi| &= \left| \frac{1-m}{m} \mathbb{E}\{\mathbf{1}_\Gamma \varphi\} - \mathbb{E}\{\mathbf{1}_{\Gamma^c} \varphi\} \right| \\ &\leq \frac{1-m}{m} \mathbb{E}\{\mathbf{1}_\Gamma \varphi\} + \mathbb{E}\{\mathbf{1}_{\Gamma^c} \varphi\} \\ &\leq \frac{1-m}{m} m + (1-m) = 2(1-m). \end{aligned}$$

This derivation also covers the case  $m = 1$ , where the unnormalized complement contribution is zero. Applying this bound to the two schedule priors on the eventual-positive rows and using the exact unconditioned equality from [Proposition 3](#) yields

$$\Delta_n \leq 2\{1 - \mathbb{H}_n^{\text{same}}(\Gamma_n^{\text{same}})\} + 2\{1 - \mathbb{H}_n^{\text{ind}}(\Gamma_n^{\text{ind}})\}.$$

The right-hand side tends to 0, while  $\Delta_n \geq 0$  because the zero test is admissible on the eventual-positive rows and by convention on the remaining rows. Therefore

$$\Delta_n \rightarrow 0.$$

Together with the previous steps, this proves the support, dense-class, conditioned-total-variation, and uniform scaled-variance parts of the diagonal certificate.

5. It remains to prove the all-statistic lower bound. Put

$$\varepsilon = \frac{d_{\text{same}} - d_{\text{ind}}}{8d_{\text{same}}}, \quad \eta = \frac{d_{\text{same}} - d_{\text{ind}}}{8}.$$

Because  $0 < c_\sigma < d_{\text{ind}} < d_{\text{same}}$ , both  $\varepsilon$  and  $\eta$  are positive. Fix an arbitrary one-realization variance statistic  $\widehat{S}_n(\mathcal{O}_n)$ , and define the threshold

$$\vartheta = \frac{d_{\text{same}} + d_{\text{ind}}}{2}.$$

For all sufficiently large  $n$ , the supports are nonempty, have positive conditioning probability, and satisfy  $e_{\text{same},n} < \eta$  and  $e_{\text{ind},n} < \eta$ . On such rows define the two wrong-decision tests

$$\psi_n^{\text{same}}(o) = \mathbf{1}\{G_n \widehat{S}_n(o) \leq \vartheta\}, \quad \psi_n^{\text{ind}}(o) = \mathbf{1}\{G_n \widehat{S}_n(o) > \vartheta\}.$$

If  $\xi \in \Gamma_n^{\text{same}}$  and

$$\left| \frac{\widehat{S}_n(o)}{\sigma_n^2(Y_n^{\text{same},\xi})} - 1 \right| \leq \varepsilon,$$

then the support bound gives

$$|G_n \sigma_n^2(Y_n^{\text{same},\xi}) - d_{\text{same}}| < \eta,$$

and hence

$$G_n \widehat{S}_n(o) > (1 - \varepsilon)(d_{\text{same}} - \eta) > \frac{d_{\text{same}} + d_{\text{ind}}}{2}.$$

Thus  $\psi_n^{\text{same}}(o) = 1$  implies relative error greater than  $\varepsilon$ . Similarly, if  $\zeta \in \Gamma_n^{\text{ind}}$  and the relative error is at most  $\varepsilon$ , then

$$G_n \widehat{S}_n(o) < (1 + \varepsilon)(d_{\text{ind}} + \eta) < \frac{d_{\text{same}} + d_{\text{ind}}}{2},$$

so  $\psi_n^{\text{ind}}(o) = 1$  implies relative error greater than  $\varepsilon$ . The two strict algebraic margins are exactly the displayed choices of  $\varepsilon$  and  $\eta$ .

6. Let

$$\mathfrak{W}_n(\widehat{S}, \varepsilon) = \sup_{Y \in \mathcal{C}_{\text{dense}}} \Pr_Y \left( \left| \frac{\widehat{S}_n(\mathcal{O}_n)}{\sigma_n^2(Y)} - 1 \right| > \varepsilon \right).$$

For each  $\xi \in \Gamma_n^{\text{same}}$ , the preceding implication bounds the same-sign row probability of  $\psi_n^{\text{same}} = 1$  by the corresponding relative-error probability. That row can be embedded into a full triangular array in  $\mathcal{C}_{\text{dense}}$ : keep the row  $Y_n^{\text{same},\xi}$  at index  $n$ , choose arbitrary support rows  $\Gamma_m^{\text{same}}$  for all sufficiently large  $m \neq n$ , and choose arbitrary bounded sign rows at the

finitely many remaining indices. The convergence and class argument from Step 3 applies because membership in the supports is eventual. Therefore

$$\mathbb{E}_{n,\text{same}}^\Gamma \psi_n^{\text{same}}(\mathcal{O}_n) \leq \mathfrak{W}_n(\hat{S}, \varepsilon).$$

The identical construction for  $\zeta \in \Gamma_n^{\text{ind}}$  gives

$$\mathbb{E}_{n,\text{ind}}^\Gamma \psi_n^{\text{ind}}(\mathcal{O}_n) \leq \mathfrak{W}_n(\hat{S}, \varepsilon).$$

Averaging over the conditioned finite supports uses only positivity of the conditioning masses.

7. Since  $\psi_n^{\text{same}} + \psi_n^{\text{ind}} = 1$  pointwise under the independent-arm observation channel,

$$\mathbb{E}_{n,\text{ind}}^\Gamma \psi_n^{\text{same}}(\mathcal{O}_n) = 1 - \mathbb{E}_{n,\text{ind}}^\Gamma \psi_n^{\text{ind}}(\mathcal{O}_n).$$

The definition of  $\Delta_n$ , applied to  $\psi_n^{\text{same}}$ , gives

$$|\mathbb{E}_{n,\text{same}}^\Gamma \psi_n^{\text{same}}(\mathcal{O}_n) - \mathbb{E}_{n,\text{ind}}^\Gamma \psi_n^{\text{same}}(\mathcal{O}_n)| \leq \Delta_n.$$

Combining the last two displays yields

$$1 - \Delta_n \leq \mathbb{E}_{n,\text{same}}^\Gamma \psi_n^{\text{same}}(\mathcal{O}_n) + \mathbb{E}_{n,\text{ind}}^\Gamma \psi_n^{\text{ind}}(\mathcal{O}_n).$$

Using Step 6 on both terms,

$$\frac{1 - \Delta_n}{2} \leq \mathfrak{W}_n(\hat{S}, \varepsilon)$$

for all sufficiently large  $n$ .

8. Finally,  $\Delta_n \rightarrow 0$ , so the left-hand side in the last display converges to  $1/2$ . Also  $\mathfrak{W}_n(\hat{S}, \varepsilon) \leq 1$ : if the set of dense-class risks is nonempty, this is the supremum of probabilities bounded by one, and if it is empty the supremum convention used here gives 0. Hence

$$\frac{1}{2} \leq \liminf_{n \rightarrow \infty} \mathfrak{W}_n(\hat{S}, \varepsilon).$$

Since  $\hat{S}$  was arbitrary, the asserted lower bound holds for every measurable one-realization variance statistic. □

## References

- Alberto Abadie, Susan Athey, Guido W. Imbens, and Jeffrey M. Wooldridge. When should you adjust standard errors for clustering? *Quarterly Journal of Economics*, 138(1):1–35, 2023. doi: 10.1093/qje/qjac038.
- Peter M. Aronow and Cyrus Samii. Estimating average causal effects under general interference, with application to a social network experiment. *Annals of Applied Statistics*, 11(4):1912–1947, 2017. doi: 10.1214/16-AOAS1005.
- Peter M. Aronow, Donald P. Green, and Donald K. K. Lee. Sharp bounds on the variance in randomized experiments. *Annals of Statistics*, 42(3):850–871, 2014. doi: 10.1214/13-AOS1200.

- Susan Athey, Dean Eckles, and Guido W. Imbens. Exact p-values for network interference. *Journal of the American Statistical Association*, 113(521):230–240, 2018. doi: 10.1080/01621459.2016.1241178.
- Guillaume Basse, Peng Ding, Avi Feller, and Panos Toulis. Randomization tests for peer effects in group formation experiments. *Econometrica*, 92(2):567–590, 2024. doi: 10.3982/ECTA20134.
- Guillaume W. Basse and Edoardo M. Airoidi. Limitations of design-based causal inference and A/B testing under arbitrary and network interference. *Sociological Methodology*, 48(1): 136–151, 2018. doi: 10.1177/0081175018782569.
- Robert M. Bell and Daniel F. McCaffrey. Bias reduction in standard errors for linear regression with multi-stage samples. *Survey Methodology*, 28(2):169–181, 2002. URL <https://www150.statcan.gc.ca/n1/en/catalogue/12-001-X20020029058>.
- Andries E. Brouwer, Arjeh M. Cohen, and Arnold Neumaier. *Distance-Regular Graphs*. Springer, Berlin, 1989. doi: 10.1007/978-3-642-74341-2.
- Andries E. Brouwer, Sebastian M. Cioabă, Ferdinand Ihringer, and Matt McGinnis. The smallest eigenvalues of hamming graphs, johnson graphs and other distance-regular graphs with classical parameters. *Journal of Combinatorial Theory, Series B*, 133:88–121, 2018. URL <https://arxiv.org/abs/1709.09011>.
- A. Colin Cameron and Douglas L. Miller. A practitioner’s guide to cluster-robust inference. *Journal of Human Resources*, 50(2):317–372, 2015. doi: 10.3368/jhr.50.2.317.
- William G. Cochran. *Sampling Techniques*. John Wiley & Sons, New York, 3 edition, 1977.
- David R. Cox. *Planning of Experiments*. John Wiley & Sons, New York, 1958.
- Philippe Delsarte. An algebraic approach to the association schemes of coding theory. Philips Research Reports Supplements 10, Philips Research Laboratories, Eindhoven, 1973.
- Dean Eckles, Brian Karrer, and Johan Ugander. Design and analysis of experiments in networks: Reducing bias from interference. *Journal of Causal Inference*, 5(1):20150021, 2017. doi: 10.1515/jci-2015-0021.
- Yuval Filmus. An orthogonal basis for functions over a slice of the boolean hypercube. *Electronic Journal of Combinatorics*, 23(1):P1.23, 2016. doi: 10.37236/4567. URL <https://arxiv.org/abs/1406.0142>.
- Ronald A. Fisher. *The Design of Experiments*. Oliver and Boyd, Edinburgh, 1935.
- Jiawei Fu, Cyrus Samii, and Ye Wang. Inference for group interaction experiments, 2026. URL <https://arxiv.org/abs/2607.02385>. Version 2; especially Sections 3–4, Theorems 1–2, Appendix A.1, and Section 8.
- Wayne A. Fuller. *Sampling Statistics*. John Wiley & Sons, Hoboken, NJ, 2009. doi: 10.1002/9780470523551.
- Mengsi Gao and Peng Ding. Causal inference in network experiments: Regression-based analysis and design-based properties. *Journal of Econometrics*, 252:106119, 2025. doi: 10.1016/j.jeconom.2025.106119.

- Jaroslav Hájek. Limiting distributions in simple random sampling from a finite population. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5:361–374, 1960.
- Christopher Harshaw, Joel A. Middleton, and Fredrik Sävje. Optimized variance estimation under interference and complex experimental designs. *Journal of the American Statistical Association*, 2026. doi: 10.1080/01621459.2026.2627027.
- Paul W. Holland. Statistics and causal inference. *Journal of the American Statistical Association*, 81(396):945–960, 1986. doi: 10.1080/01621459.1986.10478354.
- Michael G. Hudgens and M. Elizabeth Halloran. Toward causal inference with interference. *Journal of the American Statistical Association*, 103(482):832–842, 2008. doi: 10.1198/016214508000000292.
- Xinran Li and Peng Ding. General forms of finite population central limit theorems with applications to causal inference. *Journal of the American Statistical Association*, 112(520):1759–1769, 2017. doi: 10.1080/01621459.2017.1295865.
- Xinran Li, Peng Ding, Qian Lin, Dawei Yang, and Jun S. Liu. Randomization inference for peer effects. *Journal of the American Statistical Association*, 114(528):1651–1664, 2019. doi: 10.1080/01621459.2018.1512863.
- Kung-Yee Liang and Scott L. Zeger. Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1):13–22, 1986. doi: 10.1093/biomet/73.1.13.
- Lan Liu and Michael G. Hudgens. Large sample randomization inference of causal effects in the presence of interference. *Journal of the American Statistical Association*, 109(505):288–301, 2014. doi: 10.1080/01621459.2013.844698.
- Charles F. Manski. Identification of treatment response with social interactions. *Econometrica Journal*, 16(1):S1–S23, 2013. doi: 10.1111/j.1368-423X.2012.00368.x.
- Esbjörn Ohlsson. Asymptotic normality for two-stage sampling from a finite population. *Probability Theory and Related Fields*, 81:341–352, 1989. doi: 10.1007/BF00340058.
- Bryan Park and Stefan Wager. Neyman jackknife: Design-based variance estimation for causal inference under interference, 2026. URL <https://arxiv.org/abs/2604.24017>.
- David Puelz, Guillaume Basse, Avi Feller, and Panos Toulis. A graph-theoretic approach to randomization tests of causal effects under general interference. *Journal of the Royal Statistical Society: Series B*, 84(1):174–204, 2022. doi: 10.1111/rssb.12478.
- James E. Pustejovsky. *clubSandwich: Cluster-Robust (Sandwich) Variance Estimators with Small-Sample Corrections*, 2026. URL <https://cran.r-project.org/package=clubSandwich>. R package version 0.7.0, CRAN release 2026-05-04.
- James E. Pustejovsky and Elizabeth Tipton. Small-sample methods for cluster-robust variance estimation and hypothesis testing in fixed effects models. *Journal of Business & Economic Statistics*, 36(4):672–683, 2018. doi: 10.1080/07350015.2016.1247004.

- Donald B. Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5):688–701, 1974. doi: 10.1037/h0037350.
- Cyrus Samii, Ye Wang, Jonathan Sullivan, and Peter M. Aronow. Inference in spatial experiments with interference using the SpatialEffect package. *Journal of Agricultural, Biological and Environmental Statistics*, 28(1), 2023. doi: 10.1007/s13253-022-00517-y.
- Michael E. Sobel. What do randomized studies of housing mobility demonstrate?: Causal inference in the face of interference. *Journal of the American Statistical Association*, 101(476):1398–1407, 2006. doi: 10.1198/016214506000000636.
- Jerzy Splawa-Neyman, Dorota M. Dabrowska, and Terence P. Speed. On the application of probability theory to agricultural experiments. essay on principles. section 9. *Statistical Science*, 5(4):465–472, 1990. doi: 10.1214/ss/1177012031. English translation of the 1923 paper.
- Fangzhou Su and Peng Ding. Model-assisted analyses of cluster-randomized experiments. *Journal of the Royal Statistical Society: Series B*, 83(5):994–1015, 2021. doi: 10.1111/rssb.12468.
- Eric J. Tchetgen Tchetgen and Tyler J. VanderWeele. On causal inference in the presence of interference. *Statistical Methods in Medical Research*, 21(1):55–75, 2012. doi: 10.1177/0962280210386779.
- Johan Ugander, Brian Karrer, Lars Backstrom, and Jon Kleinberg. Graph cluster randomization: Network exposure to multiple universes. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 329–337, 2013. URL <https://arxiv.org/abs/1305.6979>.
- Ye Wang, Cyrus Samii, Haoge Chang, and Peter M. Aronow. Design-based inference for spatial experiments under unknown interference. *Annals of Applied Statistics*, 19(1), 2025. doi: 10.1214/24-AOAS1986.
- Halbert White. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4):817–838, 1980. doi: 10.2307/1912934.
- Hui Xu and Guillaume W. Basse. Randomization inference for composite experiments with spillovers and peer effects, 2021. URL <https://arxiv.org/abs/2103.00567>.
- Achim Zeileis. Econometric computing with HC and HAC covariance matrix estimators. *Journal of Statistical Software*, 11(10):1–17, 2004. doi: 10.18637/jss.v011.i10.
- Achim Zeileis. Object-oriented computation of sandwich estimators. *Journal of Statistical Software*, 16(9):1–16, 2006. doi: 10.18637/jss.v016.i09.
- Achim Zeileis and Thomas Lumley. *sandwich: Robust Covariance Matrix Estimators*, 2026. URL <https://cran.r-project.org/package=sandwich>. R package version 3.1-3, CRAN release 2026-08-03.
- Achim Zeileis, Susanne Köll, and Nathaniel Graham. Various versatile variances: An object-oriented implementation of clustered covariances in R. *Journal of Statistical Software*, 95(1): 1–36, 2020. doi: 10.18637/jss.v095.i01.