ALLOCATION-DEPENDENT DISTORTION FOR FAIR CLUSTERING ENSEMBLES

Anonymous authors
Paper under double-blind review

ABSTRACT

Clustering ensembles combine multiple base partitions into a final consensus. When protected groups are present, the consensus must also account for how each group is distributed across the output clusters relative to the marginal cluster masses. We study allocation-dependent distortion, in which the conditional proportion of a group assigned to a cluster directly weights the distortion of that same group–cluster cell. The resulting Fair k-means Consensus Clustering (FKCC) objective incorporates group representation into consensus aggregation without enforcing profile alignment as a hard constraint. We show that the complete objective can move profile discrepancy in either direction. Fixed-distortion analysis characterizes the allocation response and identifies when quadratic group profiles coincide or remain close; it also explains the special role of the quadratic coefficient within the positive-power family. Center elimination yields an exact finite objective with monotone one-object descent. An algebraic finite–empirical identity holds for feasible induced assignments, while uniform estimation is established separately under a fixed-representation sampling model. Experiments on nine datasets show strong clustering utility and consistent matched reductions in profile discrepancy on the tested benchmarks. Anonymous code is available at `https://anonymous.4open.science/r/method_code-F99B`.

1 INTRODUCTION

Clustering ensembles combine multiple base partitions into a final consensus, providing a common aggregation stage for partition-level information produced by different clustering procedures (Strehl & Ghosh, 2002; Fred & Jain, 2005; Monti et al., 2003). When samples also carry protected-group attributes, the consensus must account for how these groups are represented in the output. We study a group-representation criterion under which each protected group's distribution over the output clusters matches the marginal cluster-mass distribution. Equivalently, the protected group and the final cluster are independent. Because the common distribution may be nonuniform, group-representation fairness is distinct from equal cluster capacity.

Fair clustering has incorporated protected-group information through representation constraints, group-wise service objectives, regularization, and distribution alignment (Chierichetti et al., 2017; Bera et al., 2019; Makarychev & Vakilian, 2021; Zhou et al., 2024; Kim et al., 2025). Fair clustering ensembles have likewise been studied through fair graph partitioning, count-based regularization, and proportional-representation constraints (Chhabra et al., 2023; Zhou et al., 2025; Li et al., 2025; Chakraborty et al., 2025). These formulations provide several ways to specify or promote fairness in clustering. We study a different question: *how should group allocation enter the consensus distortion itself?*

Consider the center-based consensus setting studied here. If the average distortion of a protected group within each output cluster is held fixed, assigning more members of that group to a cluster increases the corresponding total distortion through the assigned sample mass. At the same time, group-representation fairness is determined by the group's allocation across all output clusters. The same allocation therefore has two roles: it determines how much sample distortion is accumulated in each group–cluster cell, and it determines the conditional cluster distribution used to assess representation. This leads to the central question of this work: *should group allocation affect consensus distortion only through sample mass, or should it also directly weight the corresponding distortion?*

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

We address this question with *allocation-dependent distortion*. In FKCC, the conditional proportion of group t assigned to output cluster q directly weights the distortion of the same group–cluster cell. Thus protected-group allocation enters the consensus objective rather than an exact feasibility constraint, and the attained profile alignment depends on the distortion structure and the final assignment.

Appendix C.1 gives two raw-association constructions: FKCC has lower profile discrepancy in one and higher discrepancy in the other. Thus the effect of allocation-dependent weighting on Δ_{prof} depends on the instance. These constructions characterize global objective preferences, whereas the coordinate-descent solver converges to a one-object local optimum.

With positive cell-average distortions held fixed, mass-only allocation is linear, whereas a coefficient p_{tq}^{α} with $\alpha > 0$ gives $p_{tq}^{\star} \propto \bar{\ell}_{tq}^{-1/\alpha}$. If $\bar{\ell}_{tq} = c_t d_q$, the quadratic profiles are identical across groups. Corollary 1 further shows that bounded multiplicative deviations from this pattern give a quantitative bound on $\Delta_{\mathrm{prof}}^{\star}$. For arbitrary observed positive distortions, Eq. (7) can be evaluated group by group to obtain the corresponding fixed-distortion profiles directly.

Quadratic weighting also has an exact coefficient-level property. For $k \geq 3$, it is the unique positive power in this family whose aggregate separates into squared group-profile discrepancy and a marginal-profile term. Binary outputs admit a cubic exception; requiring one exponent to retain the decomposition across arbitrary cluster numbers singles out $\alpha = 2$. The identity characterizes the coefficient aggregate, which contains both group-profile discrepancy and marginal-mass nonuniformity.

For the complete finite objective, we derive a pointwise scale-normalized upper bound on group-allocation discrepancy and show that it is sharp on an association-derived positive-distortion family. Empirical evaluation therefore reports Δ_{prof} directly.

The finite–empirical identity is algebraic for any realized representation and feasible induced assignment. Uniform estimation is developed under a fixed-representation sampling model, distinct from the same-sample ensemble construction used in Section 5.

The finite objective admits exact sufficient-statistics updates. Accepting only decreasing one-object moves gives monotone descent and finite termination at a one-object local optimum. Experiments on nine datasets evaluate FKCC against representative clustering-ensemble and fair-clustering baselines.

The main contributions are summarized as follows:

- We introduce allocation-dependent distortion, allowing current conditional group allocation to directly weight the corresponding consensus distortion.
- We characterize complete finite preferences in both directions, derive the fixed-distortion response, and give an exact condition for identical quadratic profiles together with a quantitative bound for near alignment. Within the positive-power family, $\alpha = 2$ uniquely retains the exact squared-Euclidean coefficient decomposition across arbitrary cluster numbers; the complete-objective discrepancy bound is sharp.
- We derive an exact center-eliminated finite objective and sufficient-statistics updates with monotone one-object coordinate descent. Proposition 5 gives an algebraic finite–empirical identity for feasible induced assignments; under a separately specified fixed-representation sampling model, we establish uniform estimation over group-specific partition classes.

2 Related Work

Clustering ensembles. Clustering ensembles combine multiple base partitions into a final consensus. Existing methods commonly use either co-association representations, which convert repeated co-occurrence across base partitions into a similarity representation, or objective-based aggregation, which directly optimizes agreement between the consensus partition and the input partitions (Fred & Jain, 2005; Liu et al., 2015; Wu et al., 2015). KCC provides an association-space, center-based consensus framework; the direct theoretical and matched reference used here is its unweighted squared-Euclidean form. Later work introduces locally weighted aggregation and scalable bipartite

constructions (Huang et al., 2018; 2020). Recent work has also studied generalization properties of ensemble clustering (Zhang et al., 2025). These weighting mechanisms characterize the reliability or contribution of input ensemble components. FKCC instead studies coefficients induced by the current group allocation of the output assignment, within the same fixed-representation, center-based aggregation setting.

Fair clustering. Fair clustering incorporates protected-group information through several criteria and algorithmic mechanisms. Representative formulations include balance and proportional-representation requirements (Chierichetti et al., 2017; Bera et al., 2019), fair spectral clustering (Kleindessner et al., 2019), variational objectives (Ziko et al., 2021), and worst-group service costs (Makarychev & Vakilian, 2021). More recent methods integrate fairness more directly with the clustering model. Fair Kernel K-Means combines kernel k-means with a fairness regularizer, while Fair Clustering via Alignment alternates protected-group distribution alignment with center optimization (Zhou et al., 2024; Kim et al., 2025). These works show that fairness can enter clustering through constraints, regularization, distribution matching, and group-dependent service objectives. FKCC focuses specifically on group-representation fairness and on the role of conditional group allocation inside the consensus distortion.

Fair clustering ensembles. Fair clustering ensembles combine protected-group requirements with consensus aggregation. Consensus Fair Clustering uses fair graph partitioning within a robustness framework (Chhabra et al., 2023). Fair Clustering Ensemble with Equal Cluster Capacity introduces a count-based regularizer that jointly promotes group fairness and cluster-capacity balance (Zhou et al., 2025). Other ensemble formulations combine co-association representations with fairness constraints (Li et al., 2025). A complementary line studies fair consensus from a combinatorial perspective, imposing proportional-representation constraints while seeking a consensus close to the input partitions, with extensions to multiple protected groups and streaming inputs (Chakraborty et al., 2025; 2026a;b).

These approaches impose or promote the desired group representation through constraints, regularization, graph partitioning, or combinatorial feasibility. In the center-based consensus setting considered here, FKCC instead places the current conditional group allocation directly in the coefficient of the corresponding group–cluster distortion, so allocation and consensus fit jointly determine each cell contribution.

3 Allocation-Dependent Consensus Clustering

3.1 Ensemble Representation and Group-Representation Fairness

Let $\Pi = \{\Pi^{(1)}, \dots, \Pi^{(m)}\}$ be an ensemble of m base partitions. Each object is represented by the concatenated one-hot memberships of the base clusters, yielding an association matrix $\mathbf{B}$. The consensus stage operates on a fixed representation $\mathbf{H} \in \mathbb{R}^{n \times d}$ derived from $\mathbf{B}$, whose i-th row is $\mathbf{h}_i^\top$.

Let $\mathbf{G} = [\mathbf{g}_1, \dots, \mathbf{g}_T] \in \{0,1\}^{n \times T}$ be the protected-group indicator matrix with $\mathbf{G}\mathbf{1}_T = \mathbf{1}_n$, and let $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_k] \in \{0,1\}^{n \times k}$ be the consensus assignment with $\mathbf{Y}\mathbf{1}_k = \mathbf{1}_n$. For group t and output cluster q, define $n_{tq} = \mathbf{g}_t^\top \mathbf{y}_q$, $n_t = \mathbf{g}_t^\top \mathbf{1}_n$, $p_{tq} = n_{tq}/n_t$, and $\rho_t = n_t/n$. Write $\mathbf{p}_t = (p_{t1}, \dots, p_{tk})^\top$ and $\bar{\mathbf{p}} = \sum_t \rho_t \mathbf{p}_t$, so that $p_{tq} = \widehat{P}(Q = q \mid G = t)$ and $\bar{p}_q = \widehat{P}(Q = q)$.

We measure disagreement among protected-group allocation distributions by

$$\Delta_{\text{prof}} = \sum_{t=1}^{T} \rho_t \|\mathbf{p}_t - \bar{\mathbf{p}}\|_2^2. \tag{1}$$

Because every protected group has positive sample mass, $\Delta_{\text{prof}} = 0$ if and only if $\mathbf{p}_t = \bar{\mathbf{p}}$ for every t, equivalently $G \perp Q$. Thus group-representation fairness requires the groups to share the same distribution over output clusters, but does not require uniform cluster masses. MNCE, used in the cross-method comparison, has the same independence endpoint $G \perp Q$ (Appendix A); because the two scores need not induce the same ordering away from independence, the matched analysis also reports Δ_{prof} directly.

3.2 Quadratic Allocation-Dependent Distortion

For centers $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_k]$, let $\ell_{tq} = \sum_i g_{it} y_{iq} \|\mathbf{h}_i - \mathbf{u}_q\|_2^2$ and, whenever $n_{tq} > 0$, $\bar{\ell}_{tq} = \ell_{tq}/n_{tq}$. Under the unweighted squared-Euclidean center-based consensus reference used here, allocation enters through assigned sample mass. FKCC additionally uses the same conditional allocation proportion to weight the corresponding distortion:

$$J(\mathbf{Y}, \mathbf{U}) = \sum_{q=1}^{k} \sum_{t=1}^{T} p_{tq}^2 \sum_{i=1}^{n} g_{it} y_{iq} \|\mathbf{h}_i - \mathbf{u}_q\|_2^2. \tag{2}$$

Because p_{tq} is induced by the current assignment, reassignment changes both the affected distortions and their coefficients. For every occupied cell, $n_{tq} = n_t p_{tq}$ and therefore

$$p_{tq}^2 \ell_{tq} = n_t p_{tq}^3 \bar{\ell}_{tq}. \tag{3}$$

This identity is the bridge to the allocation analysis in Section 4.1. Normalization by n_t makes $\mathbf{p}_t$ comparable across group sizes, while the objective still retains group-size information through the cell masses.

3.3 Finite-Sample Optimization

The solver operates on assignments with nonempty output clusters,

$$\mathcal{Y}_k = \left\{ \mathbf{Y} \in \{0,1\}^{n \times k} : \mathbf{Y}\mathbf{1}_k = \mathbf{1}_n,\ \mathbf{1}_n^\top \mathbf{y}_q \geq 1,\ \forall q \right\}. \tag{4}$$

For each cell, it caches $a_{tq} = \sum_i g_{it} y_{iq} \|\mathbf{h}_i\|_2^2$ and $\mathbf{r}_{tq} = \sum_i g_{it} y_{iq} \mathbf{h}_i$. For fixed $\mathbf{Y}$, let $\eta_q = \sum_t p_{tq}^2 n_{tq} > 0$. Minimization over the center gives $\mathbf{u}_q^\star = (\sum_t p_{tq}^2 \mathbf{r}_{tq})/\eta_q$, and substitution yields

$$F(\mathbf{Y}) = \sum_{q=1}^{k} \left[\sum_{t=1}^{T} p_{tq}^2 a_{tq} - \frac{\left\| \sum_{t=1}^{T} p_{tq}^2 \mathbf{r}_{tq} \right\|_2^2}{\eta_q} \right]. \tag{5}$$

We minimize F by one-object coordinate descent. Starting from any $\mathbf{Y}^{(0)} \in \mathcal{Y}_k$, each feasible destination is evaluated from the cached source and target statistics and accepted only when it strictly decreases F. Since $\mathcal{Y}_k$ is finite, accepted moves terminate after finitely many steps; a full sweep with no accepted move is one-object locally optimal. One sweep costs $O(nkd)$, and $\widehat{\tau}$ sweeps cost $O(nd + kTd + \widehat{\tau} nkd)$ including initialization. Detailed updates and pseudocode are given in the appendix.

4 Theoretical Analysis

We analyze how the complete finite objectives can differ, how allocation responds under fixed distortions, and how the complete FKCC objective relates to group-profile discrepancy.

4.1 Allocation Analysis

Finite examples in both directions. Allocation-dependent weighting can change the complete finite preference, with the direction of the resulting profile-discrepancy change depending on the instance.

Proposition 1 (Finite separation in both directions)**.** *There exist two native-association instances with equal-sized protected groups and $k = 2$. In the first, every global minimizer of the unweighted squared-Euclidean KCC reference has $\Delta_{\mathrm{prof}} = 1/8$, whereas every global FKCC minimizer has $\Delta_{\mathrm{prof}} = 0$. In the second, the reference has a unique global optimum up to cluster-label permutation with $\Delta_{\mathrm{prof}} = 0$, whereas every global FKCC minimizer has $\Delta_{\mathrm{prof}} = 1/18$.*

Appendix C.1 enumerates all seven and thirty-one bipartitions. Thus allocation-dependent weighting changes the objective preference while its effect on Δ_{prof} remains instance dependent.

Optimal allocation under fixed distortions. Fix group t and regard $\mathbf{p}_t \in \Delta_k$ as continuous, with $\bar{\ell}_{t1}, \ldots, \bar{\ell}_{tk} > 0$ treated as fixed per-sample distortion coefficients. The mass-only and quadratic allocation problems are

$$\min_{\mathbf{p}_t \in \Delta_k} \sum_{q=1}^{k} p_{tq} \bar{\ell}_{tq}, \qquad \min_{\mathbf{p}_t \in \Delta_k} \sum_{q=1}^{k} \bar{\ell}_{tq} p_{tq}^3. \tag{6}$$

The first selects the minimum-distortion face of the simplex, whereas the second is strictly convex under positive distortions.

Proposition 2 (Optimal allocation under quadratic weighting)**.** *If $\bar{\ell}_{tq} > 0$ for all q, then the second problem in Eq. (6) has the unique solution*

$$p_{tq}^{\star} = \frac{\bar{\ell}_{tq}^{-1/2}}{\sum_{j=1}^{k} \bar{\ell}_{tj}^{-1/2}}. \tag{7}$$

More generally, replacing p_{tq}^2 by p_{tq}^{α} with $\alpha > 0$ gives $p_{tq}^{\star} \propto \bar{\ell}_{tq}^{-1/\alpha}$. Thus smooth allocation is a property of positive powers generally. Corollary 1 further quantifies this structure for any observed positive distortion matrix: it defines the smallest multiplicative deviation from a group-by-cluster product form and bounds the resulting quadratic profile discrepancy.

Quadratic coefficient decomposition. Let $\Delta_{\mathrm{marg}} = \|\bar{\mathbf{p}} - k^{-1}\mathbf{1}_k\|_2^2$.

Proposition 3 (Quadratic allocation decomposition)**.** *For any $\mathbf{p}_t \in \Delta_k$,*

$$\sum_{t=1}^{T} \rho_t \|\mathbf{p}_t\|_2^2 = \frac{1}{k} + \Delta_{\mathrm{prof}} + \Delta_{\mathrm{marg}}. \tag{8}$$

Theorem 1 (Power-family characterization of the quadratic coefficient)**.** *Let $k \geq 3$, $T \geq 2$, and $\alpha > 0$. Suppose there exist $c > 0$ and a function M of the marginal profile such that, for every $\boldsymbol{\rho} \in \Delta_T^{\circ}$ and every collection of interior allocation profiles $\{\mathbf{p}_t \in \Delta_k^{\circ}\}_{t=1}^{T}$,*

$$\sum_{t} \rho_t \sum_{q} p_{tq}^{\alpha} = c\, \Delta_{\mathrm{prof}} + M(\bar{\mathbf{p}}). \tag{9}$$

Then $\alpha = 2$, $c = 1$, and $M(\bar{\mathbf{p}}) = \|\bar{\mathbf{p}}\|_2^2$ on Δ_k°. Conversely, $\alpha = 2$ satisfies Eq. (9).

For $k = 2$, the cubic power is an additional exact exception; requiring a single positive exponent to satisfy Eq. (9) across arbitrary cluster numbers uniquely selects $\alpha = 2$. This characterization is defined for the squared-Euclidean profile criterion and the coefficient power family. For $\alpha = 2$, Eq. (8) decomposes the coefficient aggregate into group-profile discrepancy and marginal-mass nonuniformity. Because the complete objective also couples these coefficients with cell distortions, objective descent does not imply monotone change in Δ_{prof}.

Bound for the complete objective. For a finite FKCC state, define

$$\lambda(\mathbf{Y}, \mathbf{U}) = \min_{t,q:\, n_{tq} > 0} \bar{\ell}_{tq}. \tag{10}$$

Proposition 4 (Bound on group-allocation discrepancy)**.** *For any $\mathbf{Y} \in \mathcal{Y}_k$ and centers $\mathbf{U}$ satisfying $\lambda(\mathbf{Y}, \mathbf{U}) > 0$,*

$$\Delta_{\mathrm{prof}}(\mathbf{Y}) \leq \Delta_{\mathrm{prof}}(\mathbf{Y}) + \Delta_{\mathrm{marg}}(\mathbf{Y}) \leq \sqrt{\frac{J(\mathbf{Y}, \mathbf{U})}{n\, \lambda(\mathbf{Y}, \mathbf{U})}} - \frac{1}{k}. \tag{11}$$

At the minimizing centers $\mathbf{U}^{\star}(\mathbf{Y})$, $J(\mathbf{Y}, \mathbf{U}^{\star}) = F(\mathbf{Y})$.

The bound is pointwise because $\lambda(\mathbf{Y}, \mathbf{U})$ varies with the state. Appendix C.7 constructs, for every $k \geq 2$, a strictly positive-distortion association-derived family attaining equality, establishing worst-case sharpness. Benchmark behavior is evaluated directly through Δ_{prof}. Appendix C also records a non-additivity result over the unrestricted Euclidean representation class.

4.2 Statistical Properties of the Aggregation Objective

Proposition 5 is algebraic once the representation and induced assignment are fixed. For sampling variability, we separately assume that the representation rule is fixed independently of the aggregation sample. Then, for each group t, $\mathbf{h}_{t1}, \ldots, \mathbf{h}_{tn_t}$ are modeled as i.i.d. samples from P_t, independently across groups, with fixed $n_t \geq 1$, $\rho_t = n_t/n$, and $\|\mathbf{h}\|_2 \leq B$ almost surely. The uniform-estimation theorem is stated under this fixed-representation sampling model and does not cover the same-sample construction used in Section 5.

Because FKCC uses group identity in its allocation-dependent coefficients, the population comparison class is allowed to be group-aware. Let $\boldsymbol{A} = (A^{(1)}, \ldots, A^{(T)})$, where $A^{(t)} = (A_1^{(t)}, \ldots, A_k^{(t)})$ is a measurable k-partition of the representation space used for group t; equivalently, the induced rule is $Q = A^{(G)}(\mathbf{h})$. The group-aware rule reflects that protected-group labels are observed during aggregation. Define

$$p_{tq} = P_t(A_q^{(t)}), \quad s_{tq} \quad = \mathbb{E}_t[\|\mathbf{h}\|_2^2 \mathbf{1}\{\mathbf{h} \in A_q^{(t)}\}], \quad \mathbf{m}_{tq} = \mathbb{E}_t[\mathbf{h}\, \mathbf{1}\{\mathbf{h} \in A_q^{(t)}\}]. \tag{12}$$

Empirical counterparts are denoted by hats. The population analogue of Eq. (2) is

$$R_0(\boldsymbol{A}, \mathbf{U}) = \sum_{q=1}^{k} \sum_{t=1}^{T} \rho_t p_{tq}^2 \mathbb{E}_t \left[\mathbf{1}\{\mathbf{h} \in A_q^{(t)}\} \|\mathbf{h} - \mathbf{u}_q\|_2^2 \right]. \tag{13}$$

Define $\boldsymbol{\mu}_q = \sum_t \rho_t p_{tq}^2 \mathbf{m}_{tq}$ and $\beta_q = \sum_t \rho_t p_{tq}^3$. Eliminating the centers gives

$$R(\boldsymbol{A}) = \sum_{q:\beta_q > 0} \left[\sum_{t=1}^{T} \rho_t p_{tq}^2 s_{tq} - \frac{\|\boldsymbol{\mu}_q\|_2^2}{\beta_q} \right], \tag{14}$$

with zero contribution when $\beta_q = 0$. Replacing the population statistics by their empirical counterparts defines $\widehat{R}(\boldsymbol{A})$.

Proposition 5 (Exact finite–empirical equality)**.** *If the sample assignment $\mathbf{Y}_{\boldsymbol{A}}$ induced by the group-aware rule belongs to $\mathcal{Y}_k$, then*

$$n\widehat{R}(\boldsymbol{A}) = F(\mathbf{Y}_{\boldsymbol{A}}). \tag{15}$$

Proposition 5 identifies the finite and empirical aggregation objectives for assignments induced by the comparison class. Coordinate descent is analyzed separately on the finite assignment set $\mathcal{Y}_k$ using the same center-eliminated objective.

Stability and uniform convergence. Boundedness gives $\|\mathbf{m}_{tq}\|_2 \leq B p_{tq}$ and therefore

$$\|\boldsymbol{\mu}_q\|_2 \leq B\beta_q. \tag{16}$$

Hence $0 \leq \|\boldsymbol{\mu}_q\|_2^2/\beta_q \leq B^2\beta_q \to 0$ as $\beta_q \to 0$. Let $\mathbf{p}_t = (p_{t1}, \ldots, p_{tk})^\top$, $\mathbf{s}_t = (s_{t1}, \ldots, s_{tk})^\top$, and $\mathbf{M}_t = [\mathbf{m}_{t1}, \ldots, \mathbf{m}_{tk}]$.

Theorem 2 (Global stability)**.** *For two admissible statistic collections, let R and R' denote their corresponding risks. Then*

$$|R - R'| \leq C \sum_{t=1}^{T} \rho_t \left(B^2 \|\mathbf{p}_t - \mathbf{p}_t'\|_2 + \|\mathbf{s}_t - \mathbf{s}_t'\|_2 + B\|\mathbf{M}_t - \mathbf{M}_t'\|_F \right) \tag{17}$$

for a universal constant $C > 0$.

Let $\mathcal{H}_t$ be a pointwise measurable class of k-partitions for group t, let $\mathbb{H} = \mathcal{H}_1 \times \cdots \times \mathcal{H}_T$, and let $\Pi_{\mathcal{H}_t}(N)$ denote the corresponding growth function.

Theorem 3 (Uniform convergence for group-aware comparison classes)**.** *For any $\delta \in (0, 1)$, with probability at least $1 - \delta$,*

$$\sup_{\boldsymbol{A} \in \mathbb{H}} |\widehat{R}(\boldsymbol{A}) - R(\boldsymbol{A})| \leq CB^2 \sum_{t=1}^{T} \rho_t \sqrt{\frac{\log \Pi_{\mathcal{H}_t}(n_t) + \log(6T/\delta)}{n_t}}, \tag{18}$$

where $C > 0$ is a universal constant.

The bound controls group-aware rules while retaining each group's own sample size and partition complexity. If all groups use the same structural class $\mathcal{H}_t = \mathcal{H}_0$, the bound takes the corresponding common complexity form with $\Pi_{\mathcal{H}_0}$ while still allowing the selected partition rule to differ across groups. Together with Proposition 5, this gives a uniform-estimation result for the aggregation objective under the fixed-representation sampling model specified above.

5 EXPERIMENTS

We evaluate FKCC through overall comparison, a matched KCC–FKCC analysis, numerical checks of the theory, and sensitivity to the allocation exponent.

5.1 EXPERIMENTAL PROTOCOL

We use nine datasets spanning 399–97,698 samples and 2–30 groups. D1, D4, D7, D8, and D9 use group attributes available in the original data, D6 uses person identity, and COIL20, MFEAT, and USPS3568 use fixed class-stratified binary groups balanced in expectation within each class. The latter provide controlled settings for evaluating group-allocation behavior; demographic interpretations are reserved for datasets with naturally occurring group attributes. Complete dataset information is given in Appendix A.

The main experiments construct 200 candidate partitions with coordinate-descent k-means. The base-cluster number is sampled from $[k, \lceil\sqrt{n}\rceil]$, $m = 20$ base partitions are selected per repetition, the output-cluster number is set to the ground-truth class number, and ten fixed random seeds are used.

We compare FKCC with PTASL, LWGP, SpFC, USENC, F3KM, FCMI, CFC, BFKM, FKKM, FMSC, FCE, GPEC, DREAM, and RANGE; KCC is used as the direct matched reference. Baseline parameters follow the original papers or released implementations. GPEC is reported on D1–D5, for which complete ten-run results are available; its remaining entries are marked as N/A.

We report ACC, NMI, and ARI for clustering utility and MNCE for group representation. The joint summary $F_{\mathrm{fair}} = 2\,\mathrm{ACC}\cdot\mathrm{MNCE}/(\mathrm{ACC}+\mathrm{MNCE})$ is reported together with its components. The matched analysis additionally uses Δ_{prof} and AvgTV. Full metric definitions and per-dataset results are given in Appendix A.

5.2 OVERALL COMPARISON

On D1–D5, the common evaluation subset for which all fifteen methods have complete ten-run results, FKCC ranks first in ACC, NMI, ARI, and F_{fair}, and third in MNCE. FKCC also obtains a lower average rank than each of the three recent ensemble baselines, GPEC, DREAM, and RANGE, on all five evaluation measures.

Across the nine datasets, FKCC achieves the highest available ACC and NMI on eight datasets, the highest ARI on all nine datasets, and the highest F_{fair} on all nine datasets. Its MNCE macro-average is 0.9809, indicating consistently high group representation across the tested datasets. Representative gains remain substantial: on D5, ACC increases from the strongest competing value of 0.7873 to 0.9032; on D4, NMI and ARI increase from 0.5448 and 0.5499 to 0.7747 and 0.8224, respectively. Complete ACC, NMI, ARI, MNCE, and F_{fair} results are reported in Appendix A.

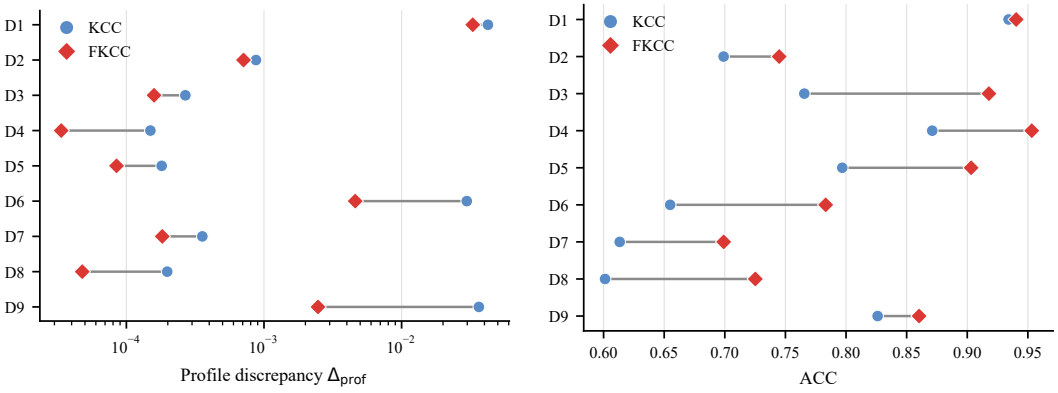
5.3 MATCHED KCC–FKCC COMPARISON

For each paired trial, KCC and FKCC use the same base partitions and representation. We first obtain the KCC solution and initialize FKCC from it, replacing the unit coefficients by p_{tq}^2. Across 90 paired trials, macro ACC increases from 0.7514 to 0.8364, NMI from 0.5215 to 0.5903, ARI from 0.5003 to 0.6037, MNCE from 0.9525 to 0.9809, and F_{fair} from 0.8324 to 0.8998. MNCE increases in $86/90$ pairs and ACC in $79/90$, with one ACC tie.

The two allocation measures show the same pattern: Δ_{prof} decreases on all nine datasets, with macro-average $0.012293 \to 0.004576$, while AvgTV decreases from 0.053599 to 0.032855. Across runs, Δ_{prof} decreases in $81/90$ pairs and AvgTV in $80/90$. The reductions are especially pro-

Table 1: Average ranks on the common complete subset D1–D5; lower is better. The full evaluation contains nine datasets. On the four larger datasets D6–D9, at least one baseline could not be completed, so a common fifteen-method ranking is not available. We therefore compute average ranks on D1–D5, where all methods have complete ten-run results. Ranks are computed from unrounded per-dataset means.

Metric	PTASL	LWGP	SpFC	USENC	F3KM	FCMI	CFC	BFKM	FKKM	FMSC	FCE	GPEC	DREAM	RANGE	FKCC
ACC	12.40	7.40	6.60	7.20	9.60	11.20	6.60	9.40	6.80	10.20	14.80	4.20	5.80	6.80	**1.00**
NMI	12.00	6.20	5.80	6.40	9.60	11.00	5.60	10.60	8.80	9.40	14.60	4.00	6.00	8.80	**1.20**
ARI	12.40	6.40	6.20	6.60	9.80	11.40	5.60	10.60	8.80	9.80	14.00	4.20	5.80	7.40	**1.00**
MNCE	11.20	11.60	7.80	12.40	9.40	6.80	11.60	7.40	2.20	7.40	**1.00**	6.80	9.80	10.40	4.20
F_{fair}	12.40	9.40	6.40	8.00	9.20	11.00	9.40	8.40	5.20	9.60	12.60	4.20	5.40	7.80	**1.00**



(a) Δ_{prof} (b) ACC

Figure 1: Matched changes from KCC to FKCC using the same base partitions and representation. Markers denote dataset means over ten paired runs. Panel (a) uses a logarithmic horizontal axis; lower Δ_{prof} and higher ACC are better.

nounced on D6 and D9, where Δ_{prof} changes from 0.02976 to 0.00460 and from 0.03645 to 0.00247, respectively.

These matched experiments characterize the complete joint objective, in which assignments, centers, and cell distortions change together; the fixed-distortion allocation is checked separately below. The same trend persists under heterogeneous base-partition generators; complete results are given in Appendix B.

5.4 Numerical Checks of the Analysis

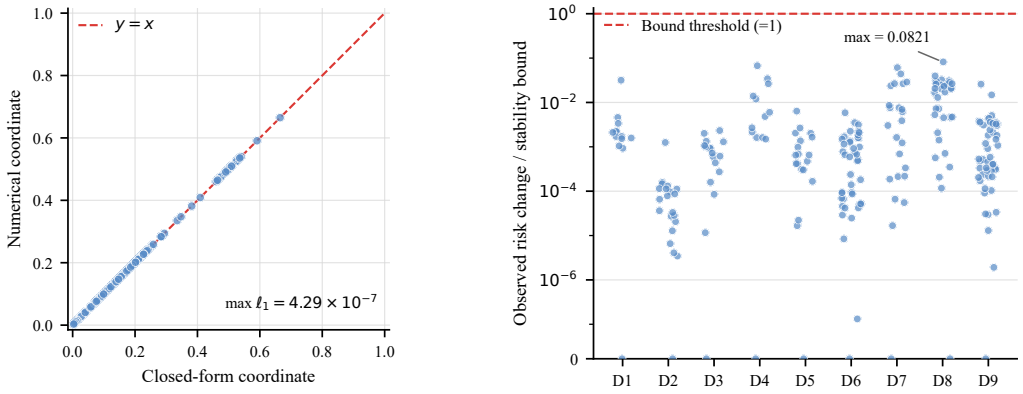
For the fixed-distortion analysis, each evaluation group in the nine datasets defines one simplex problem, giving 51 group-level instances in total. The numerical solutions agree with Proposition 2, with a maximum ℓ_1 discrepancy of 4.29×10^{-7}.

We additionally examine two boundary and stability consequences of the center-eliminated formulation. The objective remains finite in all nine zero-cell constructions. For the 239 examined statistic pairs, the observed risk-change-to-bound ratios are all below one, with a maximum of 0.0821.

The two panels visualize distinct consequences of the analysis: panel (a) compares the fixed-distortion closed form with numerical optimization, while panel (b) compares observed risk changes with the stability upper bound. Complete numerical details are reported in Appendix B.

5.5 Effect of the Allocation-Weight Exponent

We vary $\alpha \in \{0.25, 0.5, 1, 2, 4\}$ in p_{tq}^{α} while keeping the remaining protocol fixed; each exponent is used globally across all datasets.



(a) Closed-form versus numerical allocation values

(b) Observed risk changes relative to their bounds

Figure 2: Numerical checks of the analysis. (a) Closed-form and numerical allocation values for 51 fixed-positive-distortion group-level problems. (b) Ratios of observed risk changes to the corresponding stability bounds for 239 statistic pairs; the dashed line marks equality with the bound. The ratio axis uses a symmetric logarithmic scale to retain zero-valued ratios.

Table 2: Macro-averaged sensitivity to the allocation exponent on D1–D9; higher is better.

α	ACC	NMI	ARI	MNCE	F_{fair}
0.25	0.7735	0.5482	0.5358	0.9637	0.8511
0.50	0.7996	0.5664	0.5643	0.9649	0.8696
1.00	0.8254	0.5821	0.5906	0.9751	0.8905
2.00	**0.8364**	**0.5903**	**0.6037**	0.9809	**0.8998**
4.00	0.8301	0.5743	0.5879	**0.9877**	0.8990

All five macro-averages improve from $\alpha = 0.25$ to 2. At $\alpha = 4$, MNCE increases further to 0.9877, while ACC, NMI, ARI, and F_{fair} decrease. We therefore use $\alpha = 2$ as the single fixed exponent throughout the main experiments. Proposition 3 characterizes the coefficient-level decomposition associated with this quadratic choice. Additional sensitivity, convergence, runtime, ensemble-size, and matched-run results are given in Appendix B.

6 Conclusion

We studied allocation-dependent distortion for fair clustering ensembles. The complete objective can move profile discrepancy in either direction, while fixed-distortion analysis identifies when quadratic group profiles coincide or remain close. The quadratic coefficient admits an exact squared-Euclidean decomposition across arbitrary cluster numbers, characterizing the coefficient aggregate rather than an additive penalty in the complete objective. The scale-normalized complete-objective discrepancy bound is sharp on a positive-distortion family. Center elimination yields an exact finite objective and sufficient-statistics updates with monotone one-object coordinate descent and finite termination at a local optimum. Proposition 5 gives an algebraic finite–empirical identity for feasible induced assignments, and uniform estimation is established under the fixed-representation sampling model of Section 4.2. Across the tested benchmarks, FKCC combines strong clustering utility with lower matched group-allocation discrepancy.

AI use statement

Generative AI tools were used for language editing and grammatical polishing. All AI-assisted text was reviewed and revised by the authors. The authors take responsibility for the final content of this work.

REFERENCES

Suman Bera, Deeparnab Chakrabarty, Nicolas Flores, and Maryam Negahbani. Fair algorithms for clustering. *Advances in Neural Information Processing Systems*, 32, 2019.

Diptarka Chakraborty, Kushagra Chatterjee, Debarati Das, Tien Long Nguyen, and Romina Nobahari. Towards fair representation: Clustering and consensus. In *Proceedings of Thirty Eighth Conference on Learning Theory*, volume 291, pp. 838–853. PMLR, 30 Jun–04 Jul 2025.

Diptarka Chakraborty, Kushagra Chatterjee, Debarati Das, and Tien-Long Nguyen. Generalizing fair clustering to multiple groups: algorithms and applications. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 19934–19942, 2026a.

Diptarka Chakraborty, Kushagra Chatterjee, Debarati Das, and {Tien Long} Nguyen. A generic framework for fair consensus clustering in streams. In *AAMAS 2026 - Proceedings of the 25th International Conference on Autonomous Agents and Multiagent Systems*, pp. 2115–2123. Association for Computing Machinery, Inc, May 2026b.

Anshuman Chhabra, Peizhao Li, Prasant Mohapatra, and Hongfu Liu. Robust fair clustering: A novel fairness attack and defense framework. In *The Eleventh International Conference on Learning Representations*, 2023.

Flavio Chierichetti, Ravi Kumar, Silvio Lattanzi, and Sergei Vassilvitskii. Fair clustering through fairlets. *Advances in neural information processing systems*, 30, 2017.

Thomas M. Cover. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Transactions on Electronic Computers*, EC-14(3):326–334, 1965.

Ana LN Fred and Anil K Jain. Combining multiple clusterings using evidence accumulation. *IEEE transactions on pattern analysis and machine intelligence*, 27(6):835–850, 2005.

Dong Huang, Jian-Huang Lai, and Chang-Dong Wang. Robust ensemble clustering using probability trajectories. *IEEE Transactions on Knowledge and Data Engineering*, 28(5):1312–1326, 2016.

Dong Huang, Chang-Dong Wang, and Jian-Huang Lai. Locally weighted ensemble clustering. *IEEE transactions on cybernetics*, 48(5):1460–1473, 2018.

Dong Huang, Chang-Dong Wang, Jian-Sheng Wu, Jian-Huang Lai, and Chee-Keong Kwoh. Ultra-scalable spectral clustering and ensemble clustering. *IEEE Transactions on Knowledge and Data Engineering*, 32(6):1212–1226, 2020.

Kunwoong Kim, Jihu Lee, Sangchul Park, and Yongdai Kim. Fair clustering via alignment. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267, pp. 30363–30399. PMLR, 13–19 Jul 2025.

Matthäus Kleindessner, Samira Samadi, Pranjal Awasthi, and Jamie Morgenstern. Guarantees for spectral clustering with fairness constraints. In *International conference on machine learning*, pp. 3458–3467. PMLR, 2019.

Rongwen Li, Haiyang Hu, Liang Du, Jiarong Chen, Bingbing Jiang, and Peng Zhou. One-stage fair multi-view spectral clustering. In *Proceedings of the 32nd ACM International Conference on Multimedia*, pp. 1407–1416, 2024.

Yanqing Li, Ruixin Feng, and Caiming Zhong. A fair ensemble clustering method. *Symmetry*, 17 (12):2184, 2025.

Hao Lin, Hongfu Liu, Junjie Wu, Hong Li, and Stephan Günnemann. Algorithm 1038: Kcc: A matlab package for k-means-based consensus clustering. *ACM transactions on mathematical software*, 49(4):1–27, 2023.

Hongfu Liu, Tongliang Liu, Junjie Wu, Dacheng Tao, and Yun Fu. Spectral ensemble clustering. In *Proceedings of the 21st ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 715–724, 2015.

Yury Makarychev and Ali Vakilian. Approximation algorithms for socially fair clustering. In *Proceedings of Thirty Fourth Conference on Learning Theory*, volume 134, pp. 3246–3264. PMLR, 15–19 Aug 2021.

Stefano Monti, Pablo Tamayo, Jill Mesirov, and Todd Golub. Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. *Machine learning*, 52(1):91–118, 2003.

Renbo Pan, Caiming Zhong, and Jiangbo Qian. Balanced fair k-means clustering. *IEEE Transactions on Industrial Informatics*, 20(4):5914–5923, 2024.

Iosif Pinelis. Optimum bounds for the distributions of martingales in banach spaces. *The Annals of Probability*, 22(4):1679–1706, 1994.

Alexander Strehl and Joydeep Ghosh. Cluster ensembles—a knowledge reuse framework for combining multiple partitions. *Journal of machine learning research*, 3(Dec):583–617, 2002.

Junjie Wu, Hongfu Liu, Hui Xiong, Jie Cao, and Jian Chen. K-means-based consensus clustering: A unified view. *IEEE transactions on knowledge and data engineering*, 27(1):155–169, 2015.

Jiaxuan Xu, Lei Duan, Xinye Wang, and Liang Du. Large-scale robust enhanced ensemble clustering via outlier decoupling. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 39691–39700, 2026a.

Jiaxuan Xu, Lei Duan, Xinye Wang, Liang Du, Yidan Zhang, and Zhen Guo. Dynamic anchor-based one-step hypergraph ensemble clustering. *IEEE Transactions on Knowledge and Data Engineering*, 2026b.

Pengxin Zeng, Yunfan Li, Peng Hu, Dezhong Peng, Jiancheng Lv, and Xi Peng. Deep fair clustering via maximizing and minimizing mutual information: Theory, algorithm and metric. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 23986–23995, 2023.

Xu Zhang, Haoye Qiu, Weixuan Liang, Hui Liu, Junhui Hou, and Yuheng Jia. Generalization performance of ensemble clustering: From theory to algorithm. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267, pp. 75901–75927. PMLR, 13–19 Jul 2025.

Peng Zhou, Rongwen Li, and Liang Du. Fair kernel k-means: from single kernel to multiple kernel. *Advances in Neural Information Processing Systems*, 37:99686–99714, 2024.

Peng Zhou, Rongwen Li, Zhaolong Ling, Liang Du, and Xinwang Liu. Fair clustering ensemble with equal cluster capacity. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47 (3):1729–1746, 2025.

Shengkun Zhu, Quanqing Xu, Jinshan Zeng, Sheng Wang, Yuan Sun, Zhifeng Yang, Chuanhui Yang, and Zhiyong Peng. F3km: Federated, fair, and fast k-means. *Proceedings of the ACM on Management of Data*, 1(4):1–25, 2023.

Imtiaz Masud Ziko, Jing Yuan, Eric Granger, and Ismail Ben Ayed. Variational fair clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 11202–11209, 2021.

A Experimental Protocol and Complete Results

Throughout the result tables, N/A indicates that the corresponding complete ten-run result was unavailable because of memory or runtime requirements.

A.1 Datasets and Group Construction

Table 3 summarizes the nine datasets used in the experiments. The collection covers tabular data, image data, human-activity records, educational records, census data, and face representations, with sample sizes ranging from 399 to 97,698. The number of groups ranges from 2 to 30, allowing us to evaluate both binary-group and multi-group settings. Following the convention in fair clustering-ensemble evaluation Zhou et al. (2025), the last column reports the group source used for fairness-related metrics. For COIL20, MFEAT, and USPS3568, the group labels are controlled class-stratified synthetic binary groups: within each ground-truth class, group membership is sampled with equal probability once and then fixed for all experiments. This balanced-in-expectation design provides controlled, non-demographic settings for evaluating group-representation behavior. Dataset source links are provided below the table, and the processed data and preprocessing scripts are included in the anonymous code repository.

Table 3: Dataset overview.

ID	Dataset	# Samples	# Features	# Classes	Group Source
D1	CKD-Kidney[1]	399	23	2	Age (2)
D2	COIL20[2]	1440	1024	20	Class-balanced groups (2)
D3	MFEAT[3]	2000	240	10	Class-balanced groups (2)
D4	Billionaires[4]	2640	6	2	Gender (2)
D5	USPS3568[5]	3082	256	4	Class-balanced groups (2)
D6	HAR[6]	10299	561	6	Person identity (30)
D7	US Census 1990[7]	20000	6	2	Sex (2)
D8	OULAD[8]	32593	6	2	Disability (2)
D9	FairFace[9]	97698	512	2	Race (7)

[1] https://archive.ics.uci.edu/dataset/336/chronic+kidney+disease
[2] https://www.cs.columbia.edu/CAVE/software/softlib/coil-20.php
[3] https://archive.ics.uci.edu/dataset/72/multiple+features
[4] https://www.kaggle.com/datasets/nelgiriyewithana/billionaires-statistics-dataset
[5] https://www.kaggle.com/datasets/bistaumanga/usps-dataset
[6] https://archive.ics.uci.edu/dataset/240/human+activity+recognition+using+smartphones
[7] https://archive.ics.uci.edu/dataset/116/us+census+data+1990
[8] https://analyse.kmi.open.ac.uk/open_dataset
[9] https://github.com/joojs/fairface

A.2 Compared Methods and Implementation

We compare FKCC with fourteen representative baseline methods spanning three families.

- **PTASL** Huang et al. (2016): an ensemble clustering method that builds a sparse graph by filtering uncertain links and then uses probability-trajectory analysis to exploit random-walk-based global information.
- **LWGP** Huang et al. (2018): a locally weighted ensemble clustering method that estimates cluster-level reliability and constructs a locally weighted co-association graph for consensus partitioning.
- **SpFC** Kleindessner et al. (2019): a fair spectral clustering method that incorporates fairness constraints into normalized or unnormalized spectral clustering.
- **USENC** Huang et al. (2020): an ultra-scalable spectral ensemble clustering method that constructs a bipartite graph between objects and base clusters for efficient consensus clustering.
- **F3KM** Zhu et al. (2023): a federated, fair, and fast k-means method that solves fair clustering subproblems via ADMM without sharing raw data across clients.

- **FCMI** Zeng et al. (2023): a deep fair clustering method based on mutual information, designed to learn informative representations while encouraging compact, balanced, and fair clusters.
- **CFC** Chhabra et al. (2023): a robust fair clustering method that formulates consensus clustering as a fair graph-partitioning problem to defend against fairness attacks.
- **BFKM** Pan et al. (2024): a balanced fair k-means method that integrates fairness and cluster-size balance constraints into the k-means objective.
- **FKKM** Zhou et al. (2024): a fair kernel k-means framework that introduces a fairness regularization term into kernel k-means and extends it to multiple-kernel clustering.
- **FMSC** Li et al. (2024): a one-stage fair multi-view spectral clustering method that designs a graph-theoretic fairness regularizer and directly obtains discrete clustering assignments without post-processing.
- **FCE** Zhou et al. (2025): a fair clustering ensemble method that learns a fair consensus partition from multiple base clusterings while encouraging equal cluster capacity.
- **GPEC** Zhang et al. (2025): a theory-guided ensemble clustering method that learns weights for the base clusterings by reducing their bias while encouraging diversity, based on a generalization analysis of ensemble clustering.
- **DREAM** Xu et al. (2026b): a one-step hypergraph ensemble clustering method that dynamically learns anchors from an enhanced hypergraph and aligns local consensus information to obtain the final partition without additional post-processing.
- **RANGE** Xu et al. (2026a): a robust large-scale ensemble clustering method that enhances the bipartite representation of the base clusterings and decouples clean and outlier components in the anchor space.

PTASL, LWGP, USENC, GPEC, DREAM, and RANGE are clustering ensemble baselines that do not use protected-group information. SpFC, F3KM, FCMI, BFKM, FKKM, and FMSC are fair clustering baselines that incorporate fairness through different objective modifications or constraints. CFC and FCE are the closest fair clustering ensemble baselines because they introduce fairness directly into consensus clustering. KCC Wu et al. (2015); Lin et al. (2023) is additionally included as the direct reference in the matched KCC–FKCC comparison.

The experiments were run on a PC equipped with an Intel Core i7-12700 CPU and 64 GB RAM. The main clustering experiments were implemented in MATLAB R2020b.

A.3 Evaluation Metrics

Following the information-theoretic evaluation protocol for fair clustering (Zeng et al., 2023), we use MNCE to measure group-representation fairness inside clusters. Let G_t denote the set of samples from evaluation group t, and let C_q denote the set of samples assigned to cluster q. For a nonempty cluster C_q, the conditional entropy of the group labels is

$$H(G \mid C_q) = -\sum_{t=1}^{T} \frac{|G_t \cap C_q|}{|C_q|} \log \frac{|G_t \cap C_q|}{|C_q|}, \tag{19}$$

with the convention $0 \log 0 = 0$. Let

$$H(G) = -\sum_{t=1}^{T} \frac{|G_t|}{n} \log \frac{|G_t|}{n} \tag{20}$$

be the global group entropy. The minimal normalized conditional entropy is defined as

$$\mathrm{MNCE} = \min_{q: C_q \neq \emptyset} \frac{H(G \mid C_q)}{H(G)}. \tag{21}$$

A larger MNCE means that even the least entropic cluster preserves more of the dataset-level group entropy. Assuming $H(G) > 0$, let $a_q = P(Q = q)$, $\boldsymbol{\gamma}_q = P(G = \cdot \mid Q = q)$, and $\boldsymbol{\rho} = P(G = \cdot) = \sum_q a_q \boldsymbol{\gamma}_q$. Concavity of entropy gives

$$\min_{q: a_q > 0} H(\boldsymbol{\gamma}_q) \leq H(G \mid Q) \leq H(\boldsymbol{\rho}) = H(G). \tag{22}$$

Consequently, $0 \le \mathrm{MNCE} \le 1$. Moreover, $\mathrm{MNCE} = 1$ if and only if equality holds throughout the entropy inequalities; by strict concavity, this occurs exactly when $\boldsymbol{\gamma}_q = \boldsymbol{\rho}$ for every positive-mass cluster, equivalently $G \perp Q$. The same inequalities yield the quantitative relation

$$0 \le \frac{I(G;Q)}{H(G)} = 1 - \frac{H(G \mid Q)}{H(G)} \le 1 - \mathrm{MNCE}. \tag{23}$$

Thus MNCE has an exact independence interpretation at its maximum and controls normalized statistical dependence away from that maximum.

To summarize clustering utility and group-representation fairness in one score, we follow the harmonic-mean idea of the F_β metric in (Zeng et al., 2023). In the main fairness–utility table, ACC is used as the clustering-utility component, and the two components are assigned equal weight:

$$F_{\mathrm{fair}} = \frac{2 \cdot \mathrm{ACC} \cdot \mathrm{MNCE}}{\mathrm{ACC} + \mathrm{MNCE}}. \tag{24}$$

We use F_{fair} as a compact joint evaluation score whose value is high when both clustering accuracy and group-representation fairness are high. The component fairness and utility metrics are also reported separately. For each repetition, F_{fair} is computed from that run's ACC and MNCE before the mean and standard deviation are calculated across the ten runs.

For the matched allocation analysis, we additionally report the mass-weighted cluster-composition total variation

$$\mathrm{AvgTV} = \sum_{q:a_q>0} a_q \ \mathrm{TV}(\boldsymbol{\gamma}_q, \boldsymbol{\rho}) = \frac{1}{2} \sum_{t=1}^{T} \rho_t \|\mathbf{p}_t - \bar{\mathbf{p}}\|_1. \tag{25}$$

Lower values indicate cluster-wise group compositions closer to the dataset-level group proportions. By $\|\mathbf{x}\|_1 \le \sqrt{k}\|\mathbf{x}\|_2$ and Jensen's inequality,

$$\mathrm{AvgTV} \le \frac{1}{2}\sqrt{k\,\Delta_{\mathrm{prof}}}. \tag{26}$$

A.4 Complete Per-Dataset Results

For completeness, we report the five evaluation measures in the same order used in the main text: clustering utility (ACC, NMI, and ARI), group representation (MNCE), and the joint score F_{fair}. Bold and underlined entries are determined from the unrounded available means; displayed values are rounded to four decimals. Each reported entry summarizes ten runs; N/A denotes a complete ten-run result unavailable because of memory or runtime requirements.

Table 4: ACC results on the nine datasets.

Method	D1	D2	D3	D4	D5	D6	D7	D8	D9
PTASL	0.7273 ±0.1012	0.5594 ±0.0465	0.6505 ±0.0771	0.5709 ±0.0192	0.3496 ±0.1197	0.3634 ±0.0790	0.5648 ±0.0207	0.5464 ±0.0547	N/A
LWGP	0.8223 ±0.0712	0.6528 ±0.0537	0.7812 ±0.0769	0.6553 ±0.1490	0.6963 ±0.0415	0.6132 ±0.0499	0.5045 ±0.0002	**0.7302** ±0.0020	0.5921 ±0.0608
SpFC	0.7719 ±0.0000	0.6919 ±0.0578	0.8644 ±0.0762	0.6068 ±0.0000	0.6950 ±0.0019	0.5928 ±0.0415	0.5404 ±0.0000	0.5208 ±0.0021	N/A
USENC	0.9306 ±0.0046	0.6557 ±0.0481	0.7590 ±0.0868	0.5384 ±0.0491	0.6926 ±0.0485	0.5930 ±0.0858	0.5044 ±0.0002	0.7172 ±0.0269	0.7807 ±0.0844
F3KM	0.9198 ±0.0000	0.6137 ±0.0463	0.6787 ±0.0773	0.5080 ±0.0000	0.6850 ±0.0616	0.5726 ±0.0318	0.5044 ±0.0000	0.6273 ±0.0002	0.5330 ±0.0000
FCMI	0.8251 ±0.1048	0.5929 ±0.0329	0.6454 ±0.0635	0.5878 ±0.0540	0.5530 ±0.0713	0.5751 ±0.1026	0.5748 ±0.0400	0.5612 ±0.0422	N/A
CFC	0.9373 ±0.0013	0.5771 ±0.0006	0.8225 ±0.0010	0.8708 ±0.0008	0.6379 ±0.0009	0.7763 ±0.0021	0.6260 ±0.0008	0.6030 ±0.0011	N/A
BFKM	0.8860 ±0.0020	0.1067 ±0.0038	0.7392 ±0.0753	0.5076 ±0.0000	0.7190 ±0.0712	0.6308 ±0.0621	0.5044 ±0.0000	0.6273 ±0.0000	0.5355 ±0.0000
FKKM	0.8847 ±0.0000	0.6405 ±0.0374	0.7266 ±0.0515	0.6981 ±0.0000	0.7118 ±0.0004	0.7375 ±0.0803	0.6886 ±0.0000	0.7198 ±0.0000	N/A
FMSC	0.8554 ±0.1183	0.6115 ±0.0510	0.7202 ±0.0677	0.5571 ±0.0346	0.6727 ±0.1088	0.5132 ±0.0503	0.5421 ±0.0254	0.5097 ±0.0081	N/A
FCE	0.6238 ±0.0547	0.4111 ±0.0330	0.2477 ±0.0182	0.5073 ±0.0053	0.2749 ±0.0051	0.2195 ±0.0116	0.5303 ±0.0150	0.5231 ±0.0181	N/A
GPEC	0.8739 ±0.0740	0.6984 ±0.0414	0.8476 ±0.0622	0.8395 ±0.0913	0.7738 ±0.0749	N/A	N/A	N/A	N/A
DREAM	0.8845 ±0.0717	0.7192 ±0.0295	0.8033 ±0.0650	0.7752 ±0.1032	0.6867 ±0.0525	0.6927 ±0.0586	0.5947 ±0.0449	0.5254 ±0.0097	0.5665 ±0.0432
RANGE	0.6241 ±0.0000	0.7218 ±0.0217	0.8266 ±0.0558	0.5174 ±0.0000	0.7873 ±0.0380	0.6797 ±0.0416	0.5692 ±0.0000	0.5279 ±0.0000	0.5300 ±0.0000
FKCC	**0.9404** ±0.0065	**0.7449** ±0.0211	**0.9179** ±0.0256	**0.9532** ±0.0144	**0.9032** ±0.0233	**0.7832** ±0.0085	**0.6991** ±0.0029	0.7252 ±0.0476	**0.8603** ±0.0015

Table 5: NMI results on the nine datasets.

Method	D1	D2	D3	D4	D5	D6	D7	D8	D9
PTASL	0.1881 ±0.1057	0.7197 ±0.0376	0.6777 ±0.0532	0.0586 ±0.0224	0.1156 ±0.1750	0.3662 ±0.1350	0.0016 ±0.0025	0.0139 ±0.0410	N/A
LWGP	0.3423 ±0.2231	0.7789 ±0.0280	0.7985 ±0.0315	0.1977 ±0.2239	0.6012 ±0.0172	0.6369 ±0.0246	0.0051 ±0.0001	0.1578 ±0.0024	0.0892 ±0.0776
SpFC	0.1924 ±0.0037	0.8227 ±0.0208	**0.8697** ±0.0359	0.0323 ±0.0000	0.5920 ±0.0005	0.6183 ±0.0254	0.0581 ±0.0000	0.0046 ±0.0012	N/A
USENC	0.6825 ±0.0136	0.7894 ±0.0229	0.7676 ±0.0453	0.0419 ±0.0628	0.5763 ±0.0253	0.6245 ±0.0651	0.0052 ±0.0001	0.1459 ±0.0233	0.3102 ±0.1114
F3KM	0.6520 ±0.0000	0.7539 ±0.0179	0.6919 ±0.0362	0.0094 ±0.0000	0.4833 ±0.0252	0.5958 ±0.0155	0.0052 ±0.0000	0.0713 ±0.0002	0.0016 ±0.0000
FCMI	0.3851 ±0.2698	0.7303 ±0.0228	0.6598 ±0.0412	0.0640 ±0.0694	0.3224 ±0.0673	0.4592 ±0.1189	0.0148 ±0.0144	0.0274 ±0.0288	N/A
CFC	0.7026 ±0.0015	0.6933 ±0.0020	0.7831 ±0.0013	0.5448 ±0.0007	0.5777 ±0.0009	0.7095 ±0.0011	0.1192 ±0.0005	0.1291 ±0.0007	N/A
BFKM	0.5710 ±0.0043	0.0666 ±0.0039	0.7201 ±0.0356	0.0098 ±0.0000	0.5053 ±0.0362	0.6502 ±0.0292	0.0052 ±0.0000	0.0714 ±0.0000	0.0020 ±0.0000
FKKM	0.4935 ±0.0000	0.7493 ±0.0196	0.7433 ±0.0163	0.2454 ±0.0000	0.4664 ±0.0007	0.6911 ±0.0431	0.0950 ±0.0000	0.1462 ±0.0000	N/A
FMSC	0.4925 ±0.2642	0.7492 ±0.0257	0.7462 ±0.0518	0.0138 ±0.0135	0.5386 ±0.1198	0.4048 ±0.0554	0.0065 ±0.0067	0.0004 ±0.0005	N/A
FCE	0.0595 ±0.0490	0.5211 ±0.0225	0.1634 ±0.0133	0.0002 ±0.0003	0.0025 ±0.0010	0.0155 ±0.0075	0.0034 ±0.0025	0.0025 ±0.0033	N/A
GPEC	0.5117 ±0.2253	0.8048 ±0.0156	0.8028 ±0.0301	0.4956 ±0.1802	0.6090 ±0.0595	N/A	N/A	N/A	N/A
DREAM	0.5390 ±0.2225	0.8069 ±0.0161	0.7778 ±0.0365	0.3940 ±0.2364	0.5110 ±0.0311	0.5956 ±0.0700	0.0831 ±0.0586	0.0447 ±0.0174	0.0419 ±0.0593
RANGE	0.0018 ±0.0000	0.8152 ±0.0092	0.7947 ±0.0210	0.0004 ±0.0000	0.5730 ±0.0495	0.6583 ±0.0287	0.0000 ±0.0000	0.0000 ±0.0000	0.0000 ±0.0000
FKCC	**0.7127** ±0.0208	**0.8259** ±0.0094	0.8499 ±0.0199	**0.7747** ±0.0519	**0.7379** ±0.0314	**0.7109** ±0.0095	**0.1209** ±0.0039	**0.1588** ±0.0463	**0.4212** ±0.0043

Table 6: ARI results on the nine datasets.

Method	D1	D2	D3	D4	D5	D6	D7	D8	D9
PTASL	0.2206 ±0.1676	0.4823 ±0.0590	0.5156 ±0.0973	0.0205 ±0.0128	0.1019 ±0.1555	0.3117 ±0.1184	0.0015 ±0.0045	0.0177 ±0.0530	N/A
LWGP	0.4266 ±0.2064	0.5930 ±0.0465	0.7300 ±0.0524	0.1845 ±0.2296	0.5551 ±0.0266	0.5478 ±0.0414	-0.0060 ±0.0000	0.2119 ±0.0037	0.0465 ±0.0502
SpFC	0.2804 ±0.0022	0.6576 ±0.0442	0.8250 ±0.0775	0.0452 ±0.0000	0.5126 ±0.0006	0.5357 ±0.0401	-0.0034 ±0.0000	-0.0013 ±0.0004	N/A
USENC	0.7409 ±0.0160	0.6233 ±0.0361	0.6836 ±0.0776	0.0144 ±0.0296	0.5329 ±0.0347	0.5021 ±0.1038	-0.0060 ±0.0000	0.1914 ±0.0405	0.3434 ±0.1223
F3KM	0.7041 ±0.0000	0.5633 ±0.0354	0.5806 ±0.0583	-0.0010 ±0.0000	0.4625 ±0.0478	0.4727 ±0.0507	-0.0060 ±0.0000	0.0633 ±0.0002	0.0038 ±0.0000
FCMI	0.4557 ±0.2698	0.5102 ±0.0449	0.5108 ±0.0655	0.0406 ±0.0455	0.2734 ±0.0715	0.3696 ±0.1173	0.0226 ±0.0249	0.0209 ±0.0275	N/A
CFC	0.7643 ±0.0016	0.5104 ±0.0013	0.7104 ±0.0020	0.5499 ±0.0015	0.4981 ±0.0017	0.6262 ±0.0018	0.0589 ±0.0005	0.0408 ±0.0003	N/A
BFKM	0.5949 ±0.0062	0.0064 ±0.0017	0.6384 ±0.0588	-0.0010 ±0.0000	0.4971 ±0.0604	0.5280 ±0.0533	-0.0060 ±0.0000	0.0634 ±0.0000	0.0045 ±0.0000
FKKM	0.5907 ±0.0000	0.5539 ±0.0337	0.6360 ±0.0322	0.1564 ±0.0000	0.3463 ±0.0008	0.6066 ±0.0716	0.1414 ±0.0000	0.1931 ±0.0000	N/A
FMSC	0.5559 ±0.2689	0.5270 ±0.0557	0.6204 ±0.0803	0.0174 ±0.0165	0.4751 ±0.1384	0.3279 ±0.0682	0.0095 ±0.0102	0.0006 ±0.0007	N/A
FCE	0.0711 ±0.0590	0.2725 ±0.0291	0.0726 ±0.0084	-0.0001 ±0.0004	0.0013 ±0.0009	0.0106 ±0.0053	0.0045 ±0.0034	0.0034 ±0.0046	N/A
GPEC	0.5770 ±0.2076	0.6551 ±0.0296	0.7497 ±0.0543	0.4940 ±0.2292	0.5998 ±0.0820	N/A	N/A	N/A	N/A
DREAM	0.6065 ±0.2094	0.6638 ±0.0323	0.7064 ±0.0631	0.3450 ±0.2682	0.4693 ±0.0476	0.4811 ±0.1039	0.0363 ±0.0349	0.0005 ±0.0023	0.0233 ±0.0348
RANGE	-0.0020 ±0.0000	0.6769 ±0.0161	0.7362 ±0.0414	-0.0001 ±0.0000	0.5772 ±0.0588	0.5614 ±0.0478	0.0000 ±0.0000	0.0000 ±0.0000	0.0000 ±0.0000
FKCC	**0.7750** ±0.0230	**0.7044** ±0.0140	**0.8341** ±0.0327	**0.8224** ±0.0525	**0.7716** ±0.0406	**0.6358** ±0.0112	**0.1585** ±0.0046	**0.2119** ±0.0616	**0.5192** ±0.0043

B Additional Experimental Results

B.1 Detailed Matched KCC–FKCC Results

Outcome changes. We conduct a matched KCC–FKCC comparison to examine the empirical effect of allocation-dependent group–cluster weighting. Here KCC denotes the unweighted squared-Euclidean reference in the same association-space consensus setting. For each paired trial, the two methods use identical base partitions and the same input representation, and FKCC is initialized by the corresponding KCC solution. At the objective level, the resulting difference is the replacement of unit coefficients by the proposed p_{tq}^2 coefficients.

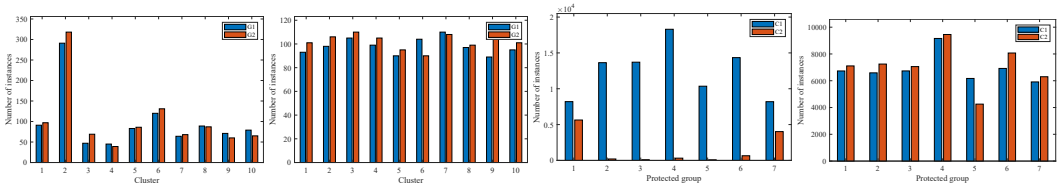
Table 9 reports both the component utility and fairness metrics and their joint score. FKCC improves every macro-average, increasing ACC from 0.7514 to 0.8364, MNCE from 0.9525 to 0.9809, and F_{fair} from 0.8324 to 0.8998. The dataset-level mean also increases for each of the five metrics on all nine datasets. At the run level, MNCE increases in $86/90$ paired trials and ACC in $79/90$, with one ACC tie.

Group-allocation measures. Using the same paired outputs as the matched KCC–FKCC comparison, we examine two quantities directly related to the fairness definition in Section 3.1. Table 10 reports the group-allocation discrepancy Δ_{prof} and the mass-weighted cluster-composition total variation (AvgTV). Both compare the group distributions across output clusters with the marginal cluster distribution. FKCC lowers both dataset-level means on all nine datasets. Across the 90 paired runs, Δ_{prof} decreases in 81 and AvgTV in 80.

Group distributions across clusters. We further inspect how the evaluation groups are distributed across the output clusters. Figure 3 compares KCC and FKCC on D3 and D9 by reporting the

Table 7: MNCE results on the nine datasets (mean ± standard deviation over ten runs). Higher is better. Bold and underline indicate the best and second-best available means, respectively.

Method	D1	D2	D3	D4	D5	D6	D7	D8	D9
PTASL	0.9590 ±0.0318	0.1790 ±0.3579	0.9835 ±0.0065	0.9496 ±0.0819	0.0000 ±0.0000	0.0000 ±0.0000	0.8893 ±0.2968	0.1899 ±0.3805	N/A
LWGP	0.9391 ±0.0081	0.8088 ±0.2744	0.9660 ±0.0559	0.5640 ±0.4696	0.9967 ±0.0012	0.6315 ±0.2532	0.9983 ±0.0001	0.9121 ±0.0024	0.5649 ±0.2376
SpFC	0.9259 ±0.0140	0.8756 ±0.0622	0.9897 ±0.0074	0.9787 ±0.0000	0.9991 ±0.0001	0.2107 ±0.0416	<u>0.9992</u> ±0.0000	0.8571 ±0.0476	N/A
USENC	0.9279 ±0.0026	0.9073 ±0.0257	0.9722 ±0.0405	0.5700 ±0.4705	0.9959 ±0.0010	0.8037 ±0.2024	0.9983 ±0.0000	0.9090 ±0.0147	<u>0.9297</u> ±0.0828
F3KM	0.9327 ±0.0000	0.9640 ±0.0036	0.9872 ±0.0061	0.9364 ±0.0000	0.9967 ±0.0011	0.8293 ±0.0700	0.9982 ±0.0000	0.8477 ±0.0006	0.7077 ±0.0000
FCMI	0.9734 ±0.0142	0.7155 ±0.3826	0.9877 ±0.0083	0.9855 ±0.0092	0.9983 ±0.0014	0.9093 ±0.0571	0.9979 ±0.0022	0.9610 ±0.0557	N/A
CFC	0.9349 ±0.0000	0.0000 ±0.0000	0.0000 ±0.0000	0.9747 ±0.0000	0.9967 ±0.0000	0.9367 ±0.0000	0.9985 ±0.0000	<u>0.9987</u> ±0.0000	N/A
BFKM	0.9704 ±0.0008	<u>0.9931</u> ±0.0059	0.9870 ±0.0030	0.9134 ±0.0000	0.9968 ±0.0012	0.9832 ±0.0064	0.9982 ±0.0000	0.8474 ±0.0000	0.7311 ±0.0001
FKKM	<u>0.9983</u> ±0.0000	0.9888 ±0.0009	<u>0.9995</u> ±0.0004	<u>0.9921</u> ±0.0000	<u>0.9998</u> ±0.0000	0.9929 ±0.0016	0.9977 ±0.0000	0.9607 ±0.0000	N/A
FMSC	0.9436 ±0.0201	0.9306 ±0.0423	0.9884 ±0.0055	0.9695 ±0.0176	0.9969 ±0.0024	0.8388 ±0.0686	0.9990 ±0.0013	0.9927 ±0.0053	N/A
FCE	**0.9999** ±0.0001	**0.9978** ±0.0016	**0.9997** ±0.0001	**0.9984** ±0.0002	**1.0000** ±0.0000	**0.9999** ±0.0000	**1.0000** ±0.0000	**1.0000** ±0.0000	N/A
GPEC	0.9344 ±0.0049	0.9207 ±0.0181	0.9906 ±0.0026	0.9782 ±0.0087	0.9982 ±0.0010	N/A	N/A	N/A	N/A
DREAM	0.9261 ±0.0161	0.9189 ±0.0103	0.9836 ±0.0217	0.9881 ±0.0085	0.9966 ±0.0019	0.8860 ±0.0419	0.9981 ±0.0031	0.9654 ±0.0498	0.4895 ±0.2414
RANGE	0.0000 ±0.0000	0.9080 ±0.0499	0.9887 ±0.0027	0.0000 ±0.0000	0.9972 ±0.0010	0.8804 ±0.0406	0.0000 ±0.0000	0.0000 ±0.0000	0.0000 ±0.0000
FKCC	0.9458 ±0.0041	0.9332 ±0.0349	0.9946 ±0.0010	0.9889 ±0.0069	0.9989 ±0.0001	<u>0.9929</u> ±0.0013	0.9990 ±0.0002	0.9796 ±0.0059	**0.9952** ±0.0003



(a) D3-KCC (b) D3-FKCC (c) D9-KCC (d) D9-FKCC

Figure 3: Group distributions across clusters on D3 and D9.

number of instances from different groups in each cluster. This visualization provides a more direct view of the group-representation behavior behind the MNCE and F_{fair} results.

Relative to KCC, FKCC yields cluster compositions closer to the dataset-level group proportions on D3 and D9, matching the lower Δ_{prof} and AvgTV values in Table 10. The figure provides a cluster-level view of the same matched outputs summarized quantitatively above.

B.2 Robustness to Heterogeneous Base-Partition Sources

The main experiments use candidate partitions produced by coordinate-descent k-means. We additionally construct 20-partition ensembles from four generators: coordinate-descent k-means and Lloyd-type k-means as centroid-based procedures, spectral clustering as a graph-based procedure, and Ward agglomerative clustering as a hierarchical procedure. In the generator order above, the three source mixtures are *Equal mixture* $(5/5/5/5)$, *Coordinate-descent-heavy* $(10/4/3/3)$, and *Without coordinate descent* $(0/7/7/6)$. The Lloyd, spectral, and Ward pools use matched cluster-count schedules and random feature subspaces. We evaluate these mixtures on D1–D5, the common subset on which the full heterogeneous construction protocol is applied. Each paired KCC–FKCC

Table 8: F_{fair} results on the nine datasets.

Method	D1	D2	D3	D4	D5	D6	D7	D8	D9
PTASL	0.8215 ±0.0622	0.1419 ±0.2842	0.7802 ±0.0569	0.7119 ±0.0328	0.0000 ±0.0000	0.0000 ±0.0000	0.6463 ±0.2161	0.1487 ±0.2983	N/A
LWGP	0.8750 ±0.0367	0.6814 ±0.2312	0.8610 ±0.0576	0.4987 ±0.4174	0.8192 ±0.0283	0.5807 ±0.2050	0.6702 ±0.0002	0.8111 ±0.0021	0.5486 ±0.1936
SpFC	0.8419 ±0.0058	0.7717 ±0.0517	0.9210 ±0.0449	0.7492 ±0.0000	0.8198 ±0.0014	0.3085 ±0.0440	0.7014 ±0.0000	0.6475 ±0.0142	N/A
USENC	0.9293 ±0.0030	0.7600 ±0.0333	0.8490 ±0.0573	0.4177 ±0.3439	0.8161 ±0.0331	0.6548 ±0.1084	0.6702 ±0.0002	0.8017 ±0.0226	0.8486 ±0.0846
F3KM	0.9262 ±0.0000	0.7490 ±0.0345	0.8019 ±0.0540	0.6586 ±0.0000	0.8104 ±0.0431	0.6745 ±0.0102	0.6702 ±0.0000	0.7210 ±0.0001	0.6081 ±0.0000
FCMI	0.8893 ±0.0591	0.5655 ±0.2986	0.7790 ±0.0465	0.7350 ±0.0416	0.7093 ±0.0578	0.7009 ±0.0887	0.7287 ±0.0327	0.7075 ±0.0389	N/A
CFC	0.9361 ±0.0010	0.0000 ±0.0000	0.0000 ±0.0000	0.9198 ±0.0012	0.7779 ±0.0008	0.8490 ±0.0005	0.7696 ±0.0004	0.7520 ±0.0011	N/A
BFKM	0.9263 ±0.0013	0.1927 ±0.0063	0.8432 ±0.0506	0.6525 ±0.0000	0.8334 ±0.0487	0.7669 ±0.0455	0.6702 ±0.0000	0.7210 ±0.0000	0.6182 ±0.0000
FKKM	0.9381 ±0.0000	0.7768 ±0.0274	0.8404 ±0.0351	0.8195 ±0.0000	0.8316 ±0.0003	0.8438 ±0.0576	0.8149 ±0.0000	0.8230 ±0.0000	N/A
FMSC	0.8915 ±0.0679	0.7370 ±0.0442	0.8313 ±0.0461	0.7071 ±0.0308	0.7982 ±0.0781	0.6337 ±0.0389	0.7025 ±0.0209	0.6735 ±0.0069	N/A
FCE	0.7669 ±0.0412	0.5815 ±0.0329	0.3967 ±0.0235	0.6728 ±0.0047	0.4312 ±0.0062	0.3599 ±0.0156	0.6930 ±0.0128	0.6867 ±0.0155	N/A
GPEC	0.9013 ±0.0392	0.7931 ±0.0231	0.9123 ±0.0369	0.9007 ±0.0539	0.8698 ±0.0475	N/A	N/A	N/A	N/A
DREAM	0.9035 ±0.0428	0.8067 ±0.0219	0.8832 ±0.0432	0.8652 ±0.0626	0.8120 ±0.0377	0.7768 ±0.0502	0.7442 ±0.0351	0.6798 ±0.0100	0.4925 ±0.1720
RANGE	0.0000 ±0.0000	0.8032 ±0.0219	0.8994 ±0.0327	0.0000 ±0.0000	0.8794 ±0.0239	0.7658 ±0.0292	0.0000 ±0.0000	0.0000 ±0.0000	0.0000 ±0.0000
FKCC	**0.9431** ±0.0052	**0.8281** ±0.0212	**0.9545** ±0.0143	**0.9706** ±0.0080	**0.9485** ±0.0133	**0.8757** ±0.0055	**0.8225** ±0.0020	**0.8323** ±0.0325	**0.9228** ±0.0008

Table 9: Matched KCC–FKCC comparison on the nine datasets. Cells report mean ± population standard deviation over ten paired runs; Avg. is the macro-average of the dataset means. Higher is better.

Method	Metric	D1	D2	D3	D4	D5	D6	D7	D8	D9	Avg.
KCC	ACC	0.9341 ±0.0056	0.6990 ±0.0329	0.7656 ±0.0671	0.8711 ±0.0968	0.7969 ±0.1574	0.6549 ±0.1035	0.6133 ±0.0917	0.6012 ±0.0648	0.8261 ±0.0956	0.7514
	NMI	0.6931 ±0.0167	0.8057 ±0.0148	0.7828 ±0.0249	0.5700 ±0.2207	0.6678 ±0.1084	0.6709 ±0.0470	0.0622 ±0.0531	0.0443 ±0.0438	0.3963 ±0.1313	0.5215
	ARI	0.7530 ±0.0194	0.6514 ±0.0288	0.7010 ±0.0464	0.5881 ±0.2554	0.6509 ±0.1627	0.5578 ±0.0707	0.0813 ±0.0793	0.0573 ±0.0592	0.4618 ±0.1522	0.5003
	MNCE	0.9305 ±0.0033	0.8974 ±0.0404	0.9908 ±0.0036	0.9755 ±0.0073	0.9976 ±0.0016	0.8709 ±0.0959	0.9977 ±0.0005	0.9591 ±0.0277	0.9528 ±0.1158	0.9525
	F_{fair}	0.9323 ±0.0040	0.7850 ±0.0257	0.8621 ±0.0420	0.9172 ±0.0565	0.8761 ±0.1153	0.7427 ±0.0880	0.7555 ±0.0715	0.7359 ±0.0408	0.8849 ±0.1048	0.8324
FKCC	ACC	**0.9404** ±0.0065	**0.7449** ±0.0211	**0.9179** ±0.0256	**0.9532** ±0.0144	**0.9032** ±0.0233	**0.7832** ±0.0085	**0.6991** ±0.0029	**0.7252** ±0.0476	**0.8603** ±0.0015	**0.8364**
	NMI	**0.7127** ±0.0208	**0.8259** ±0.0094	**0.8499** ±0.0199	**0.7747** ±0.0519	**0.7379** ±0.0314	**0.7109** ±0.0095	**0.1209** ±0.0039	**0.1588** ±0.0463	**0.4212** ±0.0043	**0.5903**
	ARI	**0.7750** ±0.0230	**0.7044** ±0.0140	**0.8341** ±0.0327	**0.8224** ±0.0525	**0.7716** ±0.0406	**0.6358** ±0.0112	**0.1585** ±0.0046	**0.2119** ±0.0616	**0.5192** ±0.0043	**0.6037**
	MNCE	**0.9458** ±0.0041	**0.9332** ±0.0349	**0.9946** ±0.0010	**0.9889** ±0.0069	**0.9989** ±0.0001	**0.9929** ±0.0013	**0.9990** ±0.0002	**0.9796** ±0.0059	**0.9952** ±0.0003	**0.9809**
	F_{fair}	**0.9431** ±0.0052	**0.8281** ±0.0212	**0.9545** ±0.0143	**0.9706** ±0.0080	**0.9485** ±0.0133	**0.8757** ±0.0055	**0.8225** ±0.0020	**0.8323** ±0.0325	**0.9228** ±0.0008	**0.8998**

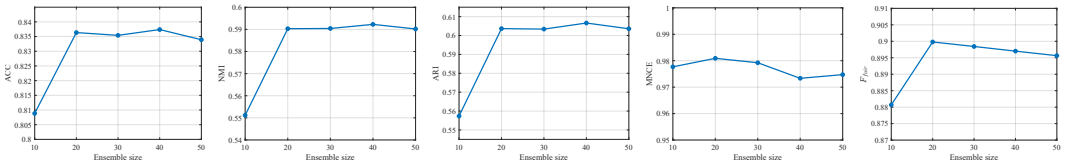
trial uses the same 20 base partitions, input representation, and initial labels, with ten matched runs for every dataset–mixture setting.

Table 10: Group-allocation measures for KCC and FKCC over ten matched runs. Entries are dataset means for KCC→FKCC; lower is better.

Dataset	Δ_{prof}	AvgTV
D1	$0.04240 \to 0.03290$	$0.14550 \to 0.12806$
D2	$0.000874 \to 0.000710$	$0.05011 \to 0.04615$
D3	$0.000268 \to 0.000159$	$0.01929 \to 0.01343$
D4	$0.000149 \to 0.000034$	$0.00552 \to 0.00232$
D5	$0.000181 \to 0.000085$	$0.01215 \to 0.00773$
D6	$0.02976 \to 0.00460$	$0.15243 \to 0.06491$
D7	$0.000356 \to 0.000182$	$0.01226 \to 0.00947$
D8	$0.000198 \to 0.000048$	$0.00503 \to 0.00278$
D9	$0.03645 \to 0.00247$	$0.08011 \to 0.02083$
Mean	$0.012293 \to 0.004576$	$0.053599 \to 0.032855$

Table 11: Robustness to heterogeneous base-partition sources on D1–D5. Each entry is averaged over ten matched runs per dataset and then macro-averaged across the five datasets. Higher is better.

Source mixture	Method	ACC	NMI	ARI	MNCE	F_{fair}
Equal mixture	KCC	0.7565	0.6066	0.5726	0.9640	0.8395
	FKCC	**0.7953**	**0.6123**	**0.5963**	**0.9737**	**0.8674**
Coordinate-descent-heavy	KCC	0.7356	0.5657	0.5217	0.9622	0.8250
	FKCC	**0.7845**	**0.5815**	**0.5677**	**0.9802**	**0.8645**
Without coordinate descent	KCC	0.7691	0.6052	0.5743	0.9630	0.8466
	FKCC	**0.8017**	**0.6066**	**0.5931**	**0.9722**	**0.8706**



(a) ACC (b) NMI (c) ARI (d) MNCE (e) F_{fair}

Figure 4: Sensitivity to the number of base clusterings, averaged over the nine datasets.

Across all three source mixtures, FKCC improves the cross-dataset averages of ACC, NMI, ARI, MNCE, and F_{fair}. The improvement persists without any coordinate-descent partition: relative to KCC, FKCC increases ACC from 0.7691 to 0.8017, MNCE from 0.9630 to 0.9722, and F_{fair} from 0.8466 to 0.8706. The controlled comparison therefore indicates that the empirical effect of the allocation-dependent consensus objective is not tied to a single base-partition generator.

B.3 Additional Sensitivity Analyses

The allocation-weight exponent sensitivity is reported in Section 5.5 and Table 2. Here we report additional sensitivity analyses not included in the main paper.

Ensemble size. We study the sensitivity of FKCC to the ensemble size, i.e., the number of base clusterings. We vary the ensemble size in $\{10, 20, 30, 40, 50\}$ and keep the other settings unchanged. For each ensemble size, FKCC is run on all nine datasets, and the average ACC, NMI, ARI, MNCE, and F_{fair} over the nine datasets are reported. As shown in Fig. 4, the largest change occurs between ensemble sizes 10 and 20. Beyond 20, all five macro-averaged metrics remain within relatively narrow ranges. We therefore use $m = 20$ as the fixed ensemble size in the main experiments, balancing empirical performance and computational cost without dataset-specific selection.

B.4 Numerical Checks of the Analysis

We numerically examine three consequences of the analysis using saved experimental states: global stability, zero group–cluster cells, and the closed-form allocation under fixed cell-average distortions.

Global stability. For Theorem 2, 239 statistic pairs cover initial-to-final KCC–FKCC changes, consecutive FKCC sweeps, and boundary constructions. Every pair satisfies Eq. (17), with a largest observed risk-change-to-bound ratio of 0.0821.

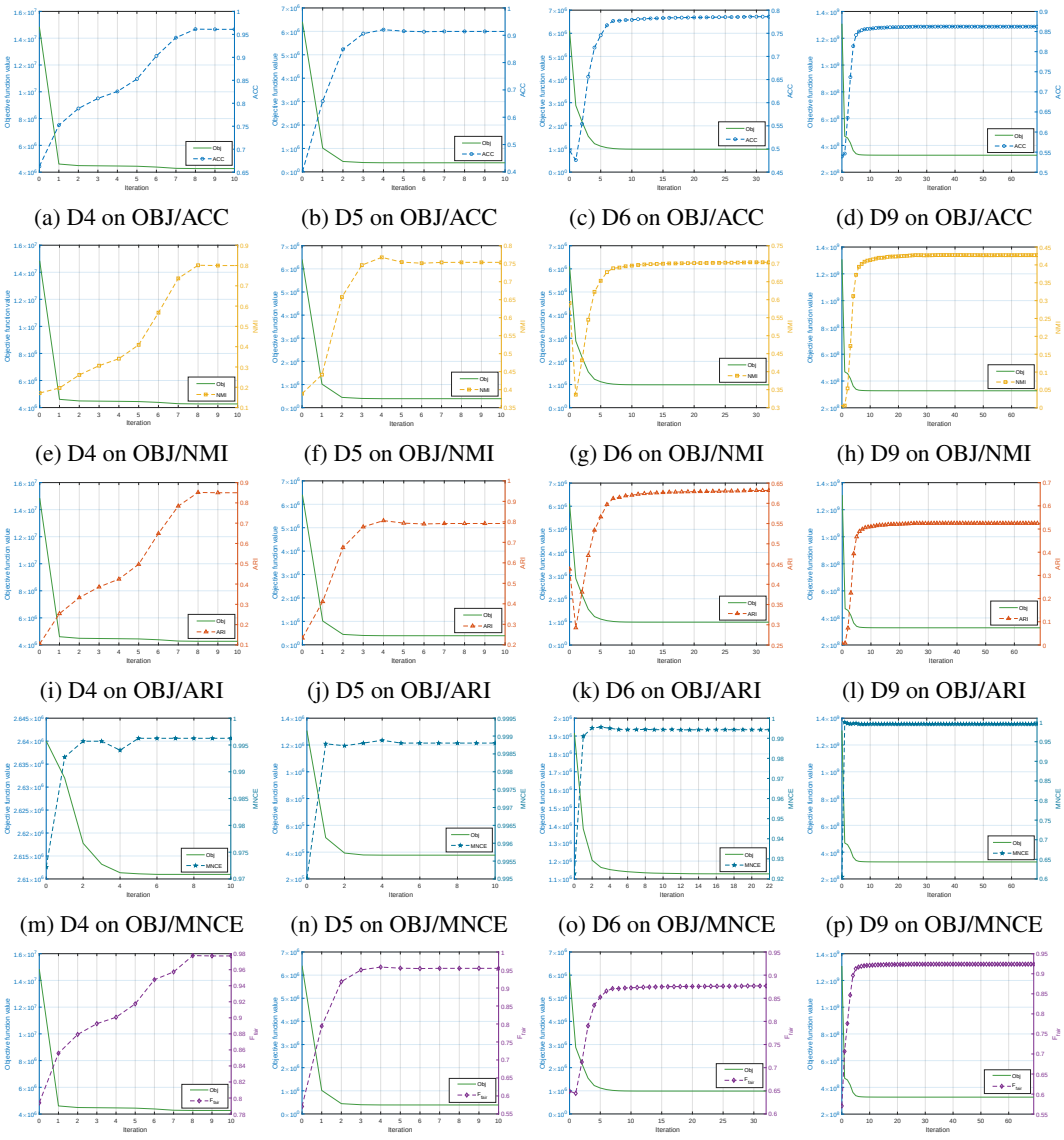
Zero cell counts. On each dataset, we empty one occupied group–cluster cell while retaining a nonempty output cluster. The center-eliminated objective remains finite in all nine cases, confirming the numerical handling of zero group–cluster counts.

Optimal allocation under fixed distortions. Holding the observed cell-average distortions fixed, we numerically solve the corresponding simplex problems. All 51 group-level instances satisfy $\bar{\ell}_{tq} > 0$ for every candidate cluster, so Proposition 2 applies directly. The numerical solutions agree with the closed-form allocation in Eq. (7), with maximum ℓ_1 discrepancy 4.29×10^{-7}.

B.5 Optimization Behavior and Runtime

Optimization trajectories. We further examine the optimization behavior of FKCC on D4, D5, D6, and D9. Figure 5 reports representative FKCC optimization trajectories together with ACC, NMI, ARI, MNCE, and F_{fair}. The objective decreases rapidly during the early sweeps and then stabilizes, consistent with the descent property of the sufficient-statistics coordinate updates. Across the representative trajectories, the external evaluation metrics generally improve rapidly and subsequently remain stable, with only mild fluctuations. These results show both the descent of the optimized objective and the corresponding empirical evolution of clustering utility and group-representation performance.

Runtime. Figure 6 compares FKCC with 14 baselines on D1–D9 using a shared logarithmic scale. On D5, FKCC takes 0.24 seconds, compared with 0.87, 2.03, and 551.06 seconds for DREAM, RANGE, and GPEC, respectively. On D6, FKCC takes 3.01 seconds, compared with 4.31 for DREAM, 14.42 for RANGE, and at least 300 for SpFC, F3KM, FCMI, and CFC. On D9, FKCC takes 171.59 seconds; DREAM takes 71.48 seconds and RANGE takes 462.61 seconds. USENC and LWGP remain faster on several datasets, while FKCC is substantially faster than multiple fairness-aware baselines.



(a) D4 on OBJ/ACC (b) D5 on OBJ/ACC (c) D6 on OBJ/ACC (d) D9 on OBJ/ACC

(e) D4 on OBJ/NMI (f) D5 on OBJ/NMI (g) D6 on OBJ/NMI (h) D9 on OBJ/NMI

(i) D4 on OBJ/ARI (j) D5 on OBJ/ARI (k) D6 on OBJ/ARI (l) D9 on OBJ/ARI

(m) D4 on OBJ/MNCE (n) D5 on OBJ/MNCE (o) D6 on OBJ/MNCE (p) D9 on OBJ/MNCE

(q) D4 on OBJ/F_{fair} (r) D5 on OBJ/F_{fair} (s) D6 on OBJ/F_{fair} (t) D9 on OBJ/F_{fair}

Figure 5: Representative FKCC optimization trajectories on D4, D5, D6, and D9 with ACC, NMI, ARI, MNCE, and F_{fair}. Each panel overlays the optimized objective with one evaluation metric.

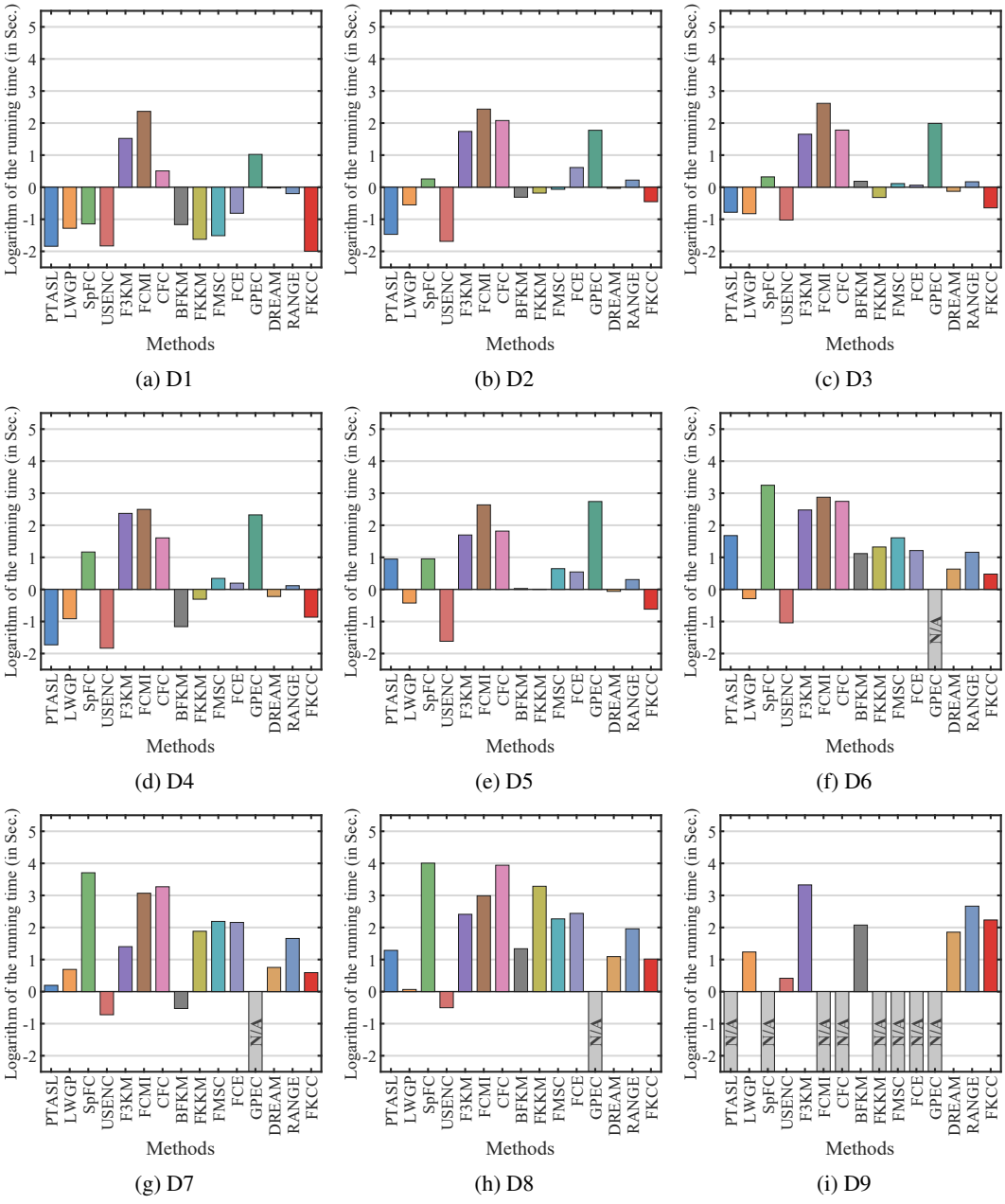


Figure 6: Runtime comparison on D1–D9. Bars show the base-10 logarithm of the mean running time in seconds over ten runs; all panels use the same vertical scale. Gray N/A entries indicate unavailable ten-run means.

C Derivations for the Allocation Analysis

This appendix collects the derivations underlying Section 4.1. We first give the complete finite separation, then analyze the fixed-distortion power family, and finally prove and calibrate the bound for the complete objective. Per-sample distortion coefficients are fixed and positive only in the fixed-distortion subsections unless a boundary case is stated explicitly.

C.1 Finite Separation in Both Directions

We give two exact constructions in the raw concatenated one-hot association representation. They show that allocation-dependent weighting changes the complete finite preference, but does not impose a universal direction on Δ_{prof}.

Example 1: FKCC has the lower-discrepancy optima. Consider four objects and three base partitions

$$\Pi^{(1)} = \{\{1,2,3\},\{4\}\}, \qquad \Pi^{(2)} = \{\{1,2,4\},\{3\}\}, \qquad \Pi^{(3)} = \{\{1,3,4\},\{2\}\}.$$

Using the raw association representation gives

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}.$$

Let $\mathcal{G}_1 = \{1,2\}$, $\mathcal{G}_2 = \{3,4\}$, and $k = 2$. For the unweighted squared-Euclidean KCC reference define

$$F_0(\mathbf{Y}) = \min_{\mathbf{U}} \sum_{q=1}^{2} \sum_{i=1}^{4} y_{iq} \left\| \mathbf{h}_i - \mathbf{u}_q \right\|_2^2.$$

For FKCC, $F(\mathbf{Y})$ is the center-eliminated objective in Eq. (5). Up to output-cluster label permutation there are seven nontrivial bipartitions:

Partition	$F_0(\mathbf{Y})$	$F(\mathbf{Y})$	Δ_{prof}
$123 \mid 4$	$8/3$	$14/9$	$1/8$
$124 \mid 3$	$8/3$	$14/9$	$1/8$
$12 \mid 34$	3	3	$1/2$
$134 \mid 2$	$8/3$	$20/9$	$1/8$
$13 \mid 24$	3	$3/4$	0
$14 \mid 23$	3	$3/4$	0
$1 \mid 234$	4	$8/3$	$1/8$

Thus the reference value is minimized at $8/3$ by $123 \mid 4$, $124 \mid 3$, and $134 \mid 2$, all with $\Delta_{\mathrm{prof}} = 1/8$, while FKCC is minimized at $3/4$ by $13 \mid 24$ and $14 \mid 23$, both with $\Delta_{\mathrm{prof}} = 0$. The cross-objective preferences are strict, and the two FKCC optima have $\lambda = 1/2 > 0$.

Example 2: the reference has the lower-discrepancy optimum. Consider six objects with

$$\Pi^{(1)} = \{\{1,3,6\},\{2,4,5\}\}, \quad \Pi^{(2)} = \{\{1,2,6\},\{3,4,5\}\},$$
$$\Pi^{(3)} = \{\{1,2,3\},\{4,5,6\}\}.$$

protected groups $\mathcal{G}_1 = \{1,2,4\}$ and $\mathcal{G}_2 = \{3,5,6\}$, and $k = 2$. The raw association matrix is

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

Fixing object 1 in the first output cluster removes label duplication and leaves exactly $2^5 - 1 = 31$ nontrivial bipartitions. Exact evaluation gives:

Partition	F_0	F	Partition	F_0	F
$1 \mid 23456$	$36/5$	$1756/315$	$1234 \mid 56$	$15/2$	$643/126$
$12 \mid 3456$	6	$289/63$	$1235 \mid 46$	$15/2$	$8/3$
$13 \mid 2456$	6	$7/3$	$1236 \mid 45$	$9/2$	2
$14 \mid 2356$	9	$118/21$	$1245 \mid 36$	$15/2$	$643/126$
$15 \mid 2346$	9	3	$1246 \mid 35$	$15/2$	$643/126$
$16 \mid 2345$	6	$7/3$	$1256 \mid 34$	$15/2$	$8/3$
$123 \mid 456$	$16/3$	$136/81$	$1345 \mid 26$	$15/2$	$8/3$
$124 \mid 356$	8	8	$1346 \mid 25$	$15/2$	$8/3$
$125 \mid 346$	8	$56/27$	$1356 \mid 24$	$15/2$	$643/126$
$126 \mid 345$	$16/3$	$136/81$	$1456 \mid 23$	$15/2$	$8/3$
$134 \mid 256$	8	$8/3$	$12345 \mid 6$	$36/5$	$1756/315$
$135 \mid 246$	8	$64/27$	$12346 \mid 5$	$36/5$	$1756/315$
$136 \mid 245$	$16/3$	$160/81$	$12356 \mid 4$	$36/5$	$1796/315$
$145 \mid 236$	8	$8/3$	$12456 \mid 3$	$36/5$	$1756/315$
$146 \mid 235$	8	$8/3$	$13456 \mid 2$	$36/5$	$572/105$
$156 \mid 234$	8	$64/27$			

The unweighted reference has the unique optimum (up to label permutation) $1236 \mid 45$, with $F_0 = 9/2$ and $\Delta_{\mathrm{prof}} = 0$; both groups have allocation profile $(2/3, 1/3)$. FKCC instead has exactly two optima, $123 \mid 456$ and $126 \mid 345$, with $F = 136/81$ and $\Delta_{\mathrm{prof}} = 1/18$. The reference optimum has FKCC value $2 > 136/81$, whereas each FKCC optimum has reference value $16/3 > 9/2$. At both FKCC optima, $\lambda = 43/81 > 0$. This proves Proposition 1 and shows that neither direction relies on a zero-distortion occupied cell.

C.2 Mass-Only Minimum-Face Characterization

Fix a group t and abbreviate $\bar{\ell}_q = \bar{\ell}_{tq}$. Let $\ell_{\min} = \min_q \bar{\ell}_q$ and $\mathcal{I}_{\min} = \{q : \bar{\ell}_q = \ell_{\min}\}$. For any $\mathbf{p} \in \Delta_k$,

$$\sum_{q=1}^{k} p_q \bar{\ell}_q = \ell_{\min} + \sum_{q=1}^{k} p_q (\bar{\ell}_q - \ell_{\min}).$$

Hence the minimum equals $\ell_{\min}$ and is attained exactly by distributions supported on $\mathcal{I}_{\min}$. Equivalently, the solution set is the face $\mathrm{conv}\{\mathbf{e}_q : q \in \mathcal{I}_{\min}\}$ of the simplex. When the minimum-distortion coordinate is unique, the solution is the corresponding vertex.

C.3 Optimal Allocation for Positive Powers and the Zero-Distortion Case

For $\alpha > 0$, consider

$$\min_{\mathbf{p} \in \Delta_k} \sum_{q=1}^{k} \bar{\ell}_q p_q^{\alpha+1}.$$

If every $\bar{\ell}_q > 0$, the objective is strictly convex. For a Lagrange multiplier ν, the KKT conditions give

$$(\alpha + 1)\bar{\ell}_q p_q^{\alpha} = \nu,$$

so the unique minimizer is

$$p_q^{\star} = \frac{\bar{\ell}_q^{-1/\alpha}}{\sum_{j=1}^{k} \bar{\ell}_j^{-1/\alpha}}.$$

Therefore

$$\frac{p_q^{\star}}{p_r^{\star}} = \left(\frac{\bar{\ell}_r}{\bar{\ell}_q}\right)^{1/\alpha}.$$

The quadratic FKCC coefficient corresponds to $\alpha = 2$, which yields Eq. (7). The mass-only case is the separate linear case $\alpha = 0$ and is characterized by the minimum face above.

If the zero-distortion set $\mathcal{Z} = \{q : \bar{\ell}_q = 0\}$ is nonempty, the minimum objective value is zero. The minimizers are exactly the distributions supported on $\mathcal{Z}$. Thus a single zero-distortion coordinate gives a unique vertex solution, whereas multiple zero-distortion coordinates give a nonunique simplex face.

C.4 Quadratic Decomposition

Using $\bar{\mathbf{p}} = \sum_t \rho_t \mathbf{p}_t$ and $\sum_t \rho_t = 1$,

$$\sum_t \rho_t \|\mathbf{p}_t - \bar{\mathbf{p}}\|_2^2 = \sum_t \rho_t \|\mathbf{p}_t\|_2^2 - \|\bar{\mathbf{p}}\|_2^2.$$

Moreover, since $\mathbf{1}_k^\top \bar{\mathbf{p}} = 1$,

$$\left\| \bar{\mathbf{p}} - \frac{1}{k}\mathbf{1}_k \right\|_2^2 = \|\bar{\mathbf{p}}\|_2^2 - \frac{1}{k}.$$

Combining the two identities proves Proposition 3.

C.5 Euclidean Compatibility within the Power Family

We prove Theorem 1. Assume that Eq. (9) holds. Choose any $\boldsymbol{\rho} \in \Delta_T^\circ$ with $\rho_1 = \rho_2 = a > 0$; for $T = 2$ this means $a = 1/2$, while for $T > 2$ the remaining positive shares may sum to $1 - 2a$. Fix an interior marginal profile $\mathbf{r} \in \Delta_k^\circ$ and a nonzero tangent direction $\mathbf{v}$ satisfying $\mathbf{1}_k^\top \mathbf{v} = 0$. For sufficiently small ε, let

$$\mathbf{p}_1 = \mathbf{r} + \varepsilon \mathbf{v}, \qquad \mathbf{p}_2 = \mathbf{r} - \varepsilon \mathbf{v},$$

and set $\mathbf{p}_t = \mathbf{r}$ for $t \geq 3$. Then $\bar{\mathbf{p}} = \mathbf{r}$ and $\Delta_{\mathrm{prof}} = 2a\varepsilon^2 \|\mathbf{v}\|_2^2$. Differentiating Eq. (9) twice with respect to ε at zero and dividing by $2a$ gives

$$\alpha(\alpha - 1) \sum_{q=1}^{k} r_q^{\alpha-2} v_q^2 = 2c \|\mathbf{v}\|_2^2 . \tag{27}$$

If $0 < \alpha < 1$, the left-hand side is negative while the right-hand side is positive; if $\alpha = 1$, the left-hand side is zero. Hence $\alpha > 1$. Taking $\mathbf{v} = \mathbf{e}_i - \mathbf{e}_j$ in Eq. (27) gives

$$\alpha(\alpha - 1) \left(r_i^{\alpha-2} + r_j^{\alpha-2} \right) = 4c.$$

Because $k \geq 3$, we may choose interior profiles with $r_i = r_j = x$ while varying x over a nondegenerate interval and keeping the remaining coordinates positive. Thus

$$2\alpha(\alpha - 1) x^{\alpha-2} = 4c$$

must be independent of x, which forces $\alpha = 2$. Substitution gives $c = 1$. Finally, setting all group profiles equal to an arbitrary interior $\mathbf{r}$ yields $M(\mathbf{r}) = \sum_q r_q^2 = \|\mathbf{r}\|_2^2$. The converse follows directly from Proposition 3.

The restriction $k \geq 3$ is essential. On the binary simplex, $\sum_q p_q^3 = \frac{3}{2}\|\mathbf{p}\|_2^2 - \frac{1}{2}$, so the cubic coefficient aggregate also admits a decomposition into a positive multiple of Δ_{prof} plus a marginal-profile-only term. The theorem therefore characterizes the quadratic exponent for cluster numbers $k \geq 3$, or equivalently among powers required to satisfy the decomposition uniformly across arbitrary cluster numbers.

C.6 Allocation under Shared and Nearly Shared Relative Distortions

Corollary 1 (Allocation under shared or nearly shared relative distortions)**.** *Let* $\bar{\mathbf{L}} = [\bar{\ell}_{tq}] \in \mathbb{R}_{++}^{T \times k}$*, and define*

$$\varepsilon_\star(\bar{\mathbf{L}}) = \min_{\substack{\mathbf{a} \in \mathbb{R}^T,\, \mathbf{b} \in \mathbb{R}^k \\ a_1 = 0}} \max_{t,q} \left| \log \bar{\ell}_{tq} - a_t - b_q \right| . \tag{28}$$

Then the optimal quadratic allocation profiles satisfy

$$\Delta_{\mathrm{prof}}^\star \leq \min \left\{ \left(e^{\varepsilon_\star(\bar{\mathbf{L}})} - 1 \right)^2 , 1 - \frac{1}{k} \right\} .$$

In particular, $\varepsilon_\star(\bar{\mathbf{L}}) = 0$ *if and only if*

$$\bar{\ell}_{tq} = c_t d_q$$

for some $c_t, d_q > 0$. In this case,

$$p^\star_{tq} = \frac{d_q^{-1/2}}{\sum_{j=1}^k d_j^{-1/2}},$$

independently of t, and hence

$$\mathbf{p}^\star_1 = \cdots = \mathbf{p}^\star_T, \qquad \Delta^\star_{\mathrm{prof}} = 0.$$

Proof. Let $(\mathbf{a}^\star, \mathbf{b}^\star)$ be an optimizer of Eq. 28, and set

$$c_t = e^{a^\star_t}, \qquad d_q = e^{b^\star_q}, \qquad \xi_{tq} = \log \bar{\ell}_{tq} - a^\star_t - b^\star_q.$$

Then

$$\bar{\ell}_{tq} = c_t d_q e^{\xi_{tq}}, \qquad |\xi_{tq}| \le \varepsilon_\star(\bar{\mathbf{L}}).$$

Define

$$r_q = \frac{d_q^{-1/2}}{\sum_{j=1}^k d_j^{-1/2}}, \qquad z_{tq} = e^{-\xi_{tq}/2}.$$

By Eq. (7),

$$p^\star_{tq} = \frac{r_q z_{tq}}{\sum_{j=1}^k r_j z_{tj}}.$$

For brevity, write $\varepsilon_\star = \varepsilon_\star(\bar{\mathbf{L}})$. Since

$$z_{tq} \in \left[e^{-\varepsilon_\star/2}, e^{\varepsilon_\star/2}\right],$$

and $\sum_j r_j = 1$, the denominator $\sum_j r_j z_{tj}$ lies in the same interval. Hence

$$e^{-\varepsilon_\star} \le \frac{p^\star_{tq}}{r_q} \le e^{\varepsilon_\star}.$$

Therefore

$$\|\mathbf{p}^\star_t - \mathbf{r}\|_1 = \sum_q r_q \left|\frac{p^\star_{tq}}{r_q} - 1\right| \le e^{\varepsilon_\star} - 1,$$

and consequently

$$\|\mathbf{p}^\star_t - \mathbf{r}\|_2 \le e^{\varepsilon_\star} - 1.$$

Because the weighted mean

$$\bar{\mathbf{p}}^\star = \sum_t \rho_t \mathbf{p}^\star_t$$

minimizes

$$\sum_t \rho_t \|\mathbf{p}^\star_t - \mathbf{v}\|_2^2$$

over $\mathbf{v}$, we obtain

$$\Delta^\star_{\mathrm{prof}} = \sum_t \rho_t \|\mathbf{p}^\star_t - \bar{\mathbf{p}}^\star\|_2^2 \le \sum_t \rho_t \|\mathbf{p}^\star_t - \mathbf{r}\|_2^2 \le \left(e^{\varepsilon_\star} - 1\right)^2.$$

Moreover,

$$\Delta_{\mathrm{prof}} = \sum_t \rho_t \|\mathbf{p}_t\|_2^2 - \|\bar{\mathbf{p}}\|_2^2 \le 1 - \frac{1}{k},$$

because $\|\mathbf{p}_t\|_2^2 \le 1$ for every $\mathbf{p}_t \in \Delta_k$ and $\|\bar{\mathbf{p}}\|_2^2 \ge 1/k$. Thus

$$\Delta^\star_{\mathrm{prof}} \le \min\left\{\left(e^{\varepsilon_\star} - 1\right)^2, 1 - \frac{1}{k}\right\}.$$

Finally, $\varepsilon_\star(\bar{\mathbf{L}}) = 0$ if and only if

$$\log \bar{\ell}_{tq} = a_t + b_q$$

for all t, q, equivalently

$$\bar{\ell}_{tq} = c_t d_q$$

for some $c_t, d_q > 0$. In this case, substituting the factorization into Eq. (7) cancels the group-specific factor $c_t^{-1/2}$ and yields

$$p^\star_{tq} = \frac{d_q^{-1/2}}{\sum_{j=1}^k d_j^{-1/2}},$$

independently of t. □

C.7 Bound for the Complete Objective

For any finite state with $\lambda = \lambda(\mathbf{Y}, \mathbf{U}) > 0$, the occupied-cell identity Eq. (3) gives

$$\frac{J(\mathbf{Y}, \mathbf{U})}{n} = \sum_t \rho_t \sum_{q:\, n_{tq}>0} \bar{\ell}_{tq} p_{tq}^3 \tag{29}$$

$$\geq \lambda \sum_t \rho_t \sum_{q:\, n_{tq}>0} p_{tq}^3 = \lambda \sum_t \rho_t \sum_q p_{tq}^3. \tag{30}$$

For each group,

$$\left(\sum_q p_{tq}^2\right)^2 = \left(\sum_q p_{tq}^{3/2} p_{tq}^{1/2}\right)^2 \leq \left(\sum_q p_{tq}^3\right)\left(\sum_q p_{tq}\right) = \sum_q p_{tq}^3. \tag{31}$$

Jensen's inequality therefore yields

$$\sum_t \rho_t \sum_q p_{tq}^3 \geq \sum_t \rho_t \|\boldsymbol{p}_t\|_2^4 \geq \left(\sum_t \rho_t \|\boldsymbol{p}_t\|_2^2\right)^2. \tag{32}$$

Writing $S = \sum_t \rho_t \|\boldsymbol{p}_t\|_2^2$, Eq. (8) gives $S = 1/k + \Delta_{\mathrm{prof}} + \Delta_{\mathrm{marg}}$. Combining Eq. (30) through Eq. (32) yields

$$\Delta_{\mathrm{prof}} + \Delta_{\mathrm{marg}} = S - \frac{1}{k} \leq \sqrt{\frac{J(\mathbf{Y}, \mathbf{U})}{n\lambda}} - \frac{1}{k}.$$

Since $\Delta_{\mathrm{marg}} \geq 0$, the same right-hand side also bounds Δ_{prof}, proving Proposition 4. At the minimizing centers, $J(\mathbf{Y}, \mathbf{U}^\star) = F(\mathbf{Y})$.

When positive global rescalings are allowed, some distortion-scale normalization is necessary, and the bound is sharp within the association-derived representation setting. Fix any $k \geq 2$ and let $T = k$, with two objects per group and equal group shares. Take two identical base partitions, each placing the $n = 2k$ objects in singleton clusters, so the raw association matrix is $\mathbf{B} = [\mathbf{I}_n, \mathbf{I}_n]$. For any $s > 0$, use the positively rescaled representation $\mathbf{H} = s\mathbf{B}$ and assign the two objects of group t to output cluster t. Then $\boldsymbol{p}_t = \mathbf{e}_t$, $\bar{\boldsymbol{p}} = k^{-1}\mathbf{1}_k$, $\Delta_{\mathrm{marg}} = 0$, and

$$\Delta_{\mathrm{prof}} = 1 - \frac{1}{k}.$$

At the minimizing center of each output cluster, both assigned objects have squared distortion s^2. Hence every occupied-cell average distortion equals $\lambda = s^2$, while

$$F(\mathbf{Y}) = ns^2 = n\lambda.$$

Consequently

$$\sqrt{\frac{F(\mathbf{Y})}{n\lambda}} - \frac{1}{k} = 1 - \frac{1}{k} = \Delta_{\mathrm{prof}} + \Delta_{\mathrm{marg}},$$

so equality holds in Proposition 4 for every $k \geq 2$. Along the same strictly positive-distortion family, $F(\mathbf{Y})/n = s^2 \to 0$ as $s \downarrow 0$, while $\Delta_{\mathrm{prof}} = 1 - 1/k$ remains fixed. Thus no universal bound depending only on the per-sample objective and vanishing without a distortion-scale normalization can control Δ_{prof} over this representation class. If zero occupied-cell distortion is allowed outright, the limiting case also gives $F = 0$ with nonzero profile discrepancy.

C.8 Non-Additivity of Allocation-Dependent Distortion

Let

$$J_0(\mathbf{Y}, \mathbf{U}) = \sum_{q=1}^{k} \sum_{t=1}^{T} \sum_{i=1}^{n} g_{it} y_{iq} \|\mathbf{h}_i - \mathbf{u}_q\|_2^2$$

denote the ordinary unweighted consensus distortion.

Proposition 6 (Non-additivity of allocation-dependent distortion)**.** *Fix a feasible assignment* $\mathbf{Y}$ *containing two occupied group–cluster cells* (t, q) *and* (t', q') *with* $p_{tq}^2 \neq p_{t'q'}^2$*. Over the Euclidean representation class* $\mathbf{H} \in \mathbb{R}^{n \times d}$*, there are no constant* c *and function* A *depending only on the group–cluster counts* $\{n_{tq}\}$ *such that*

$$J(\mathbf{Y}, \mathbf{U}) = cJ_0(\mathbf{Y}, \mathbf{U}) + A(\{n_{tq}\})$$

for every representation $\mathbf{H}$ *and every choice of centers* $\mathbf{U}$*.*

Proof. Set all representations and centers to zero. Both distortions vanish, so the putative identity gives $A(\{n_{tq}\}) = 0$ for the fixed assignment. Next, change only the representation of one object in occupied cell (t, q) from zero to an arbitrary vector $\mathbf{v}$, while retaining zero centers. The two sides increase by $p_{tq}^2 \|\mathbf{v}\|_2^2$ and $c \|\mathbf{v}\|_2^2$, respectively; hence $c = p_{tq}^2$. Repeating the same construction in cell (t', q') gives $c = p_{t'q'}^2$, contradicting the assumed identity. Thus no decomposition of the stated form exists. □

The proposition is deliberately stated over the unrestricted Euclidean representation class; it does not assert the same impossibility for every restricted association-derived subclass.

D Finite Optimization and Objective Closure

D.1 Center Elimination for the Finite Objective

Write $w_{tq} = p_{tq}^2$. Fix a feasible assignment matrix $\mathbf{Y}$ and a final cluster q. Conditional on this assignment, the weights w_{tq} are fixed when optimizing the center $\mathbf{u}_q$. The part of the finite objective in Eq. (2) involving $\mathbf{u}_q$ is

$$J_q(\mathbf{u}_q) = \sum_t w_{tq} \sum_i g_{it} y_{iq} \|\mathbf{h}_i - \mathbf{u}_q\|_2^2 . \tag{33}$$

Expanding the squared norm and using $a_{tq} = \sum_i g_{it} y_{iq} \|\mathbf{h}_i\|_2^2$ and $\mathbf{r}_{tq} = \sum_i g_{it} y_{iq} \mathbf{h}_i$, we obtain

$$J_q(\mathbf{u}_q) = \sum_t w_{tq} a_{tq} - 2\Big\langle \mathbf{u}_q, \sum_t w_{tq} \mathbf{r}_{tq} \Big\rangle + \|\mathbf{u}_q\|_2^2 \sum_t w_{tq} n_{tq}. \tag{34}$$

Let $D_q = \sum_t w_{tq} n_{tq}$ and $\mathbf{R}_q = \sum_t w_{tq} \mathbf{r}_{tq}$. If $D_q > 0$, completing the square gives

$$J_q(\mathbf{u}_q) = \sum_t w_{tq} a_{tq} - \frac{\|\mathbf{R}_q\|_2^2}{D_q} + D_q \|\mathbf{u}_q - \mathbf{R}_q / D_q\|_2^2 . \tag{35}$$

Hence the minimizing center is $\mathbf{u}_q^\star = \mathbf{R}_q / D_q$, and the minimized cluster contribution is $\sum_t w_{tq} a_{tq} - \|\mathbf{R}_q\|_2^2 / D_q$. For every feasible $\mathbf{Y} \in \mathcal{Y}_k$, the cluster is nonempty, so some $n_{tq} > 0$; then $w_{tq} = (n_{tq}/n_t)^2 > 0$ and therefore $D_q > 0$. Thus the finite solver never encounters a zero denominator. For algebraic extensions that allow an empty sample-index cluster, we assign that cluster zero contribution, consistent with the population zero-mass convention. Summing over q yields $F(\mathbf{Y})$.

D.2 Cached Objective and Exact One-Object Updates

We optimize the center-eliminated objective $F(\mathbf{Y})$ by coordinate descent, updating one row of $\mathbf{Y}$ at a time. For the finite optimization procedure, we use the standard fixed-k assignment set

$$\mathcal{Y}_k = \left\{\mathbf{Y} \in \{0,1\}^{n\times k} : \mathbf{Y}\mathbf{1}_k = \mathbf{1}_n,\ \mathbf{1}_n^\top \mathbf{y}_q \geq 1 \text{ for all } q \in \{1,\dots,k\}\right\}. \tag{36}$$

Thus every sample has one cluster assignment and every output cluster remains nonempty during the finite optimization.

Lemma 1 (Global-scale invariance)**.** *Let $F_{\mathbf{H}}$ denote the center-eliminated objective induced by $\mathbf{H}$. For every $c > 0$, $F_{c\mathbf{H}}(\mathbf{Y}) = c^2 F_{\mathbf{H}}(\mathbf{Y})$. Consequently, a common positive rescaling of the representation leaves the minimizing assignments and the signs of exact reassignment changes unchanged.*

Proof. Under $\mathbf{h}_i \mapsto c\mathbf{h}_i$ and $\mathbf{u}_q \mapsto c\mathbf{u}_q$, every squared distance in Eq. (2) is multiplied by c^2; minimizing over the centers preserves this common factor. □

For cluster q, define its current sample count by

$$n_q = \mathbf{1}_n^\top \mathbf{y}_q = \sum_{t=1}^{T} n_{tq}. \tag{37}$$

To evaluate candidate moves efficiently, we maintain the group–cluster statistics $\{n_{tq}, a_{tq}, \mathbf{r}_{tq}, w_{tq}\}$, where $w_{tq} = p_{tq}^2 = (n_{tq}/n_t)^2$ and the remaining quantities are defined in Sections 3.1 and 3.3, together with the following cluster-level cached quantities:

$$S_q = \sum_{t=1}^{T} w_{tq} a_{tq}, \qquad \mathbf{R}_q = \sum_{t=1}^{T} w_{tq}\mathbf{r}_{tq}, \qquad D_q = \sum_{t=1}^{T} w_{tq} n_{tq}, \tag{38}$$

and the cluster contribution

$$\psi_q = \begin{cases} S_q - \|\mathbf{R}_q\|_2^2/D_q, & D_q > 0, \\ 0, & D_q = 0. \end{cases} \tag{39}$$

On the feasible set $\mathcal{Y}_k$, only the first branch is used; the zero branch is retained solely as a continuous algebraic extension. Then $F(\mathbf{Y}) = \sum_{q=1}^{k} \psi_q$.

To update the i-th row of $\mathbf{Y}$, let $\mathbf{e}_q$ denote the q-th standard basis vector in $\mathbb{R}^k$. Define $\mathbf{Y}^{(q)}$ as the assignment matrix obtained by replacing the i-th row of $\mathbf{Y}$ with $\mathbf{e}_q$, and let $\mathbf{Y}^{(0)}$ be the reference matrix obtained by setting all entries in the i-th row to zero. Suppose the current assignment of sample i is the $\hat{q}$-th cluster. Updating the i-th row amounts to solving

$$\min_{\mathbf{Y}^{(q)}} \quad F(\mathbf{Y}^{(q)}) - F(\mathbf{Y}^{(0)}), \tag{40}$$

$$\text{s.t.} \quad \mathbf{Y}^{(q)} \in \{\mathbf{Y}^{(1)},\dots,\mathbf{Y}^{(k)}\} \cap \mathcal{Y}_k. \tag{41}$$

Since $\mathbf{Y}^{(0)}$ removes sample i from its current cluster $\hat{q}$, while $\mathbf{Y}^{(q)}$ further assigns it to candidate cluster q, we have

$$F(\mathbf{Y}) - F(\mathbf{Y}^{(0)}) = \psi_{\hat{q}} - \psi_{\hat{q}}^{(0)}. \tag{42}$$

For $q \neq \hat{q}$, $\mathbf{Y}^{(0)}$ leaves cluster q unchanged, so $\psi_q^{(0)} = \psi_q$, and

$$F(\mathbf{Y}^{(q)}) - F(\mathbf{Y}^{(0)}) = \psi_q^{(q)} - \psi_q. \tag{43}$$

Therefore, for $q \neq \hat{q}$, the objective change caused by moving sample i from cluster $\hat{q}$ to cluster q is

$$F(\mathbf{Y}^{(q)}) - F(\mathbf{Y}) = \psi_{\hat{q}}^{(0)} + \psi_q^{(q)} - \psi_{\hat{q}} - \psi_q. \tag{44}$$

For $q = \hat{q}$, we have $\mathbf{Y}^{(\hat{q})} = \mathbf{Y}$, so $F(\mathbf{Y}^{(\hat{q})}) - F(\mathbf{Y}) = 0$.

Suppose that sample i belongs to the t-th protected group. Let $\mathbf{x} = \mathbf{h}_i$ and $c_i = \|\mathbf{h}_i\|_2^2$. For a candidate target $q \neq \hat{q}$, the remove-and-add operation affects only the source cell $(t, \hat{q})$ and the target cell (t, q). We compute the source removal once and then evaluate each possible target addition.

Source removal. The reference assignment $\mathbf{Y}^{(0)}$ removes sample i from cluster $\hat{q}$. The updated source-cell statistics are

$$n_{t\hat{q}}^{(0)} = n_{t\hat{q}} - 1, \quad a_{t\hat{q}}^{(0)} = a_{t\hat{q}} - c_i, \quad \mathbf{r}_{t\hat{q}}^{(0)} = \mathbf{r}_{t\hat{q}} - \mathbf{x}, \quad w_{t\hat{q}}^{(0)} = \big((n_{t\hat{q}} - 1)/n_t\big)^2. \tag{45}$$

The corresponding source-cluster caches are

$$S_{\hat{q}}^{(0)} = S_{\hat{q}} - w_{t\hat{q}} a_{t\hat{q}} + w_{t\hat{q}}^{(0)} a_{t\hat{q}}^{(0)},$$
$$\mathbf{R}_{\hat{q}}^{(0)} = \mathbf{R}_{\hat{q}} - w_{t\hat{q}} \mathbf{r}_{t\hat{q}} + w_{t\hat{q}}^{(0)} \mathbf{r}_{t\hat{q}}^{(0)}, \tag{46}$$
$$D_{\hat{q}}^{(0)} = D_{\hat{q}} - w_{t\hat{q}} n_{t\hat{q}} + w_{t\hat{q}}^{(0)} n_{t\hat{q}}^{(0)}. \tag{47}$$

Then $\psi_{\hat{q}}^{(0)}$ is computed from Eq. (39) using $S_{\hat{q}}^{(0)}, \mathbf{R}_{\hat{q}}^{(0)}, D_{\hat{q}}^{(0)}$.

Target addition. For a candidate target cluster $q \neq \hat{q}$, the assignment $\mathbf{Y}^{(q)}$ adds sample i to cluster q. The updated target-cell statistics are

$$n_{tq}^{(q)} = n_{tq} + 1, \quad a_{tq}^{(q)} = a_{tq} + c_i, \quad \mathbf{r}_{tq}^{(q)} = \mathbf{r}_{tq} + \mathbf{x}, \quad w_{tq}^{(q)} = \big((n_{tq} + 1)/n_t\big)^2. \tag{48}$$

The corresponding target-cluster caches are

$$S_q^{(q)} = S_q - w_{tq} a_{tq} + w_{tq}^{(q)} a_{tq}^{(q)},$$
$$\mathbf{R}_q^{(q)} = \mathbf{R}_q - w_{tq} \mathbf{r}_{tq} + w_{tq}^{(q)} \mathbf{r}_{tq}^{(q)}, \tag{49}$$
$$D_q^{(q)} = D_q - w_{tq} n_{tq} + w_{tq}^{(q)} n_{tq}^{(q)}. \tag{50}$$

Then $\psi_q^{(q)}$ is computed from Eq. (39) using $S_q^{(q)}, \mathbf{R}_q^{(q)}, D_q^{(q)}$.

Using the above updates, the best target for sample i is obtained by

$$q^\star = \arg \min_{q \in \{1,\dots,k\}} \big(F(\mathbf{Y}^{(q)}) - F(\mathbf{Y}^{(0)})\big). \tag{51}$$

Since $F(\mathbf{Y}) - F(\mathbf{Y}^{(0)})$ in Eq. (42) is independent of q, minimizing $F(\mathbf{Y}^{(q)}) - F(\mathbf{Y}^{(0)})$ is equivalent to minimizing $F(\mathbf{Y}^{(q)}) - F(\mathbf{Y})$ in Eq. (44). A candidate move is feasible whenever the resulting assignment remains in $\mathcal{Y}_k$. Hence, for $q \neq \hat{q}$, a move from the source cluster $\hat{q}$ is evaluated only when $n_{\hat{q}} > 1$. No positivity constraint is imposed on an individual group–cluster count n_{tq}; such a count may become zero, in which case the corresponding allocation weight and sufficient statistics are updated according to their definitions. Among the feasible candidates, a reassignment is accepted exactly when its objective change is negative.

D.3 Coordinate-Descent Algorithm

D.4 Descent, Finite Termination, and Complexity

Proposition 7 (Descent and finite termination)**.** *In exact arithmetic, every accepted move strictly decreases F. Since $\mathcal{Y}_k$ is finite, the coordinate descent procedure terminates after finitely many accepted moves. At termination, every feasible one-object reassignment in $\mathcal{Y}_k$ has a nonnegative exact objective change. The quantity in Eq. (44) gives this change for each candidate reassignment.*

Remark 1 (Objective-level connection). The sufficient statistics used by coordinate descent also define the empirical aggregation risk. For assignments induced by a group-aware rule $\boldsymbol{A} \in \mathbb{H}$, Proposition 5 gives $F(\mathbf{Y}_{\boldsymbol{A}}) = n\widehat{R}(\boldsymbol{A})$. Thus the statistical and optimization analyses share the same center-eliminated objective for assignments induced by the comparison class, while the descent result itself applies to the full finite assignment set $\mathcal{Y}_k$.

Precomputing the cell-level statistics $\{n_{tq}, a_{tq}, \mathbf{r}_{tq}\}$ requires one pass over the n samples, costing $\mathcal{O}(nd)$ for the vector aggregates $\mathbf{r}_{tq}$ and $\mathcal{O}(n)$ for the scalars n_{tq}, a_{tq}. Computing $w_{tq} = (n_{tq}/n_t)^2$ and the cluster caches $S_q, \mathbf{R}_q, D_q$ costs $\mathcal{O}(kTd)$.

In each sweep, updating a single sample i requires evaluating $k - 1$ candidate moves. For each candidate $q \neq \hat{q}$, the target-cluster cache update involves subtracting the old cell (t, q) contribution

Algorithm 1 Sufficient-Statistics Coordinate Descent for FKCC

Require: $\mathbf{H} \in \mathbb{R}^{n\times d}$, $\mathbf{G} \in \{0,1\}^{n\times T}$, number of clusters k
Ensure: $\mathbf{Y} \in \mathcal{Y}_k$
  Initialize $\mathbf{Y} \in \mathcal{Y}_k$ (e.g., from a nonempty k-means++ partition on $\mathbf{H}$).
  Compute $\{n_{tq}, a_{tq}, \mathbf{r}_{tq}\}$ from Sections 3.1 and 3.3, and set $w_{tq} = (n_{tq}/n_t)^2$, $n_q = \sum_{t=1}^{T} n_{tq}$ for all q, and $\{S_q, \mathbf{R}_q, D_q, \psi_q\}$ via Eq. (38) and Eq. (39).
  repeat
    for $i = 1$ **to** n **do**
      $(\hat{q}, t) \leftarrow$ (current cluster, group of sample i).
      if $n_{\hat{q}} \le 1$ **then**
        continue
      end if
      Compute $n_{t\hat{q}}^{(0)}, a_{t\hat{q}}^{(0)}, \mathbf{r}_{t\hat{q}}^{(0)}, w_{t\hat{q}}^{(0)}$ via Eq. (45), then $S_{\hat{q}}^{(0)}, \mathbf{R}_{\hat{q}}^{(0)}, D_{\hat{q}}^{(0)}$ via Eq. (47), and $\psi_{\hat{q}}^{(0)}$ via Eq. (39).
      $q^\star \leftarrow \hat{q}, \quad \Delta_{\min} \leftarrow 0.$
      for $q = 1$ **to** k, $q \neq \hat{q}$ **do**
        Compute $n_{tq}^{(q)}, a_{tq}^{(q)}, \mathbf{r}_{tq}^{(q)}, w_{tq}^{(q)}$ via Eq. (48), then $S_q^{(q)}, \mathbf{R}_q^{(q)}, D_q^{(q)}$ via Eq. (50), and $\psi_q^{(q)}$ via Eq. (39).
        $\Delta \leftarrow \psi_{\hat{q}}^{(0)} + \psi_q^{(q)} - \psi_{\hat{q}} - \psi_q$ Eq. (44).
        if $\Delta < \Delta_{\min}$ **then**
          $\Delta_{\min} \leftarrow \Delta, \quad q^\star \leftarrow q.$
        end if
      end for
      if $q^\star \neq \hat{q}$ **then**
        Set the i-th row of $\mathbf{Y}$ to $\mathbf{e}_{q^\star}$.
        $n_{\hat{q}} \leftarrow n_{\hat{q}} - 1, \qquad n_{q^\star} \leftarrow n_{q^\star} + 1.$
        Replace the source cell $(t, \hat{q})$ with the computed removal statistics and the target cell $(t, q^\star)$ with the stored addition statistics for $q^\star$.
        Replace $(S_{\hat{q}}, \mathbf{R}_{\hat{q}}, D_{\hat{q}}, \psi_{\hat{q}})$ with $(S_{\hat{q}}^{(0)}, \mathbf{R}_{\hat{q}}^{(0)}, D_{\hat{q}}^{(0)}, \psi_{\hat{q}}^{(0)})$.
        Replace $(S_{q^\star}, \mathbf{R}_{q^\star}, D_{q^\star}, \psi_{q^\star})$ with $(S_{q^\star}^{(q^\star)}, \mathbf{R}_{q^\star}^{(q^\star)}, D_{q^\star}^{(q^\star)}, \psi_{q^\star}^{(q^\star)})$.
      end if
    end for
  until no move accepted in a full sweep
  return $\mathbf{Y}$

and adding the new one, taking $\mathcal{O}(d)$ for the vector operations on $\mathbf{R}_q$ and $\mathcal{O}(1)$ for S_q, D_q, plus $\mathcal{O}(d)$ to evaluate $\psi_q^{(q)}$ via Eq. (39). The source contribution $\psi_{\hat{q}}^{(0)}$ is computed once per sample. Hence each sample update costs $\mathcal{O}(kd)$, and a full sweep over all n samples costs $\mathcal{O}(nkd)$.

The total complexity over $\widehat{\tau}$ sweeps is $\mathcal{O}(nkd\widehat{\tau})$. When the raw binary association representation is used, $d = d_B = \sum_{r=1}^{m} k_r$. Hence the complexity is $\mathcal{O}(nkd_B\widehat{\tau})$. It is linear in n when k, d_B, and $\widehat{\tau}$ are fixed; when the ensemble construction lets d_B vary with n, the corresponding dependence is retained explicitly through d_B.

D.5 Population Formulation and Exact Finite–Empirical Equality

Fix a group-aware rule $\boldsymbol{A} = (A^{(1)}, \ldots, A^{(T)})$ as in Section 4.2. For cluster q, the population term is

$$R_{0,q}(\mathbf{u}_q) = \sum_t \rho_t p_{tq}^2 \mathbb{E}_t[\mathbf{1}\{\mathbf{h} \in A_q^{(t)}\} \, \|\mathbf{h} - \mathbf{u}_q\|_2^2]. \tag{52}$$

Expanding the square and using Eq. (12) gives

$$R_{0,q}(\mathbf{u}_q) = \sum_t \rho_t p_{tq}^2 s_{tq} - 2\Big\langle \mathbf{u}_q, \sum_t \rho_t p_{tq}^2 \mathbf{m}_{tq} \Big\rangle + \|\mathbf{u}_q\|_2^2 \sum_t \rho_t p_{tq}^3 \tag{53}$$

$$= \Gamma_q - 2\langle \mathbf{u}_q, \boldsymbol{\mu}_q \rangle + \beta_q \, \|\mathbf{u}_q\|_2^2 \, . \tag{54}$$

If $\beta_q > 0$, completing the square gives

$$R_{0,q}(\mathbf{u}_q) = \Gamma_q - \frac{\|\boldsymbol{\mu}_q\|_2^2}{\beta_q} + \beta_q \|\mathbf{u}_q - \boldsymbol{\mu}_q/\beta_q\|_2^2 . \tag{55}$$

Thus the minimized risk is $R(\boldsymbol{A}) = \sum_q [\Gamma_q - Q_q]$, where $Q_q = \|\boldsymbol{\mu}_q\|_2^2 / \beta_q$ if $\beta_q > 0$ and $Q_q = 0$ otherwise.

For the induced assignment $\mathbf{Y}_{\boldsymbol{A}}$, the empirical statistics satisfy $\widehat{p}_{tq} = n_{tq}/n_t$, $\widehat{s}_{tq} = a_{tq}/n_t$, $\widehat{\mathbf{m}}_{tq} = \mathbf{r}_{tq}/n_t$, and $\rho_t = n_t/n$. Substitution gives

$$\widehat{\Gamma}_q = \frac{1}{n} \sum_t w_{tq} a_{tq}, \tag{56}$$

$$\widehat{\boldsymbol{\mu}}_q = \frac{1}{n} \sum_t w_{tq} \mathbf{r}_{tq}, \qquad \widehat{\beta}_q = \frac{1}{n} \sum_t w_{tq} n_{tq}. \tag{57}$$

Hence, under the common zero-mass convention,

$$n\widehat{R}(\boldsymbol{A}) = F(\mathbf{Y}_{\boldsymbol{A}}), \tag{58}$$

which proves Proposition 5.

For any feasible finite assignment matrix $\mathbf{Y}$, the same sample algebra remains valid after defining its empirical counts and moments directly. This yields a sample-index plug-in objective equal to $F(\mathbf{Y})/n$. The probability theorem, however, concerns assignments induced by the chosen group-aware comparison class $\mathbb{H}$.

E Statistical Theory and Consequences

E.1 Admissible Statistic Domain

For each group t, a group-aware partition $A^{(t)}$ induces $\boldsymbol{p}_t = (p_{t1}, \ldots, p_{tk})$, $\boldsymbol{s}_t = (s_{t1}, \ldots, s_{tk})$, and $\mathbf{M}_t = (\mathbf{m}_{t1}, \ldots, \mathbf{m}_{tk})$. We call such a collection admissible when $\boldsymbol{p}_t \in \Delta_k$, $0 \le s_{tq} \le B^2 p_{tq}$, and $\|\mathbf{m}_{tq}\|_2 \le B p_{tq}$ for every q. These conditions follow from

$$s_{tq} = \mathbb{E}_t[\|\mathbf{h}\|_2^2 \mathbf{1}\{\mathbf{h} \in A_q^{(t)}\}] \le B^2 p_{tq}$$

and

$$\|\mathbf{m}_{tq}\|_2 = \left\| \mathbb{E}_t[\mathbf{h}\mathbf{1}\{\mathbf{h} \in A_q^{(t)}\}] \right\|_2 \le B p_{tq}.$$

Empirical statistics satisfy the same inequalities. The admissible domain is convex: the simplex and scalar-moment constraints are preserved by interpolation, and

$$\left\|(1-\theta)\mathbf{m}_q + \theta \mathbf{m}'_q\right\|_2 \le (1-\theta) B p_q + \theta B p'_q = B p_q^{\theta}.$$

E.2 Zero-Mass Continuity

For any admissible statistic collection,

$$\|\boldsymbol{\mu}_q\|_2 = \left\| \sum_t \rho_t p_{tq}^2 \mathbf{m}_{tq} \right\|_2 \le B \sum_t \rho_t p_{tq}^3 = B\beta_q. \tag{59}$$

Consequently, $0 \le Q_q = \|\boldsymbol{\mu}_q\|_2^2 / \beta_q \le B^2 \beta_q$ whenever $\beta_q > 0$, and $Q_q \to 0$ as $\beta_q \downarrow 0$. The convention $Q_q = 0$ at $\beta_q = 0$ is therefore the continuous extension.

E.3 Proof of Global Stability

Let z and z' be two admissible statistic collections and let $z^\theta = (1-\theta)z + \theta z'$. For $\epsilon > 0$, define

$$Q_{q,\epsilon} = \|\boldsymbol{\mu}_q\|_2^2 / (\beta_q + \epsilon).$$

Along the segment,

$$\dot{\boldsymbol{\mu}}_q = \sum_t \rho_t \left(2p_{tq}\dot{p}_{tq}\mathbf{m}_{tq} + p_{tq}^2\dot{\mathbf{m}}_{tq}\right), \tag{60}$$

$$\dot{\beta}_q = \sum_t 3\rho_t p_{tq}^2 \dot{p}_{tq}. \tag{61}$$

Differentiating and using $\|\boldsymbol{\mu}_q\|_2 \le B\beta_q$, $\|\mathbf{m}_{tq}\|_2 \le Bp_{tq}$, and $\beta_q/(\beta_q+\epsilon) \le 1$ gives

$$\left|\dot{Q}_{q,\epsilon}\right| \le 7B^2 \sum_t \rho_t p_{tq}^2 |\dot{p}_{tq}| + 2B \sum_t \rho_t p_{tq}^2 \left\|\dot{\mathbf{m}}_{tq}\right\|_2. \tag{62}$$

Summing over q and using $\sum_q p_{tq}^4 \le 1$ yields

$$\sum_q |\dot{Q}_{q,\epsilon}| \le 7B^2 \sum_t \rho_t \left\|\dot{\boldsymbol{p}}_t\right\|_2 + 2B \sum_t \rho_t \left\|\dot{\mathbf{M}}_t\right\|_F. \tag{63}$$

The additive term $\sum_{t,q} \rho_t p_{tq}^2 s_{tq}$ satisfies

$$\left|\frac{d}{d\theta}\sum_{t,q} \rho_t p_{tq}^2 s_{tq}\right| \le 2B^2 \sum_t \rho_t \left\|\dot{\boldsymbol{p}}_t\right\|_2 + \sum_t \rho_t \left\|\dot{\boldsymbol{s}}_t\right\|_2. \tag{64}$$

Integrating the resulting derivative bound and taking $\epsilon \downarrow 0$ gives

$$|R(z) - R(z')| \le 9B^2 \sum_t \rho_t \left\|\boldsymbol{p}_t - \boldsymbol{p}_t'\right\|_2 + \sum_t \rho_t \left\|\boldsymbol{s}_t - \boldsymbol{s}_t'\right\|_2 + 2B \sum_t \rho_t \left\|\mathbf{M}_t - \mathbf{M}_t'\right\|_F, \tag{65}$$

which proves Theorem 2.

E.4 Uniform Concentration for Group-Aware Classes

We prove Theorem 3 for $\mathbb{H} = \mathcal{H}_1 \times \cdots \times \mathcal{H}_T$. Each $\mathcal{H}_t$ is the pointwise measurable partition class used for group t.

Generic Hilbert-valued bound. Let $X_1, \ldots, X_N$ be i.i.d. from P and let $\mathcal{G}$ be a class of Hilbert-valued functions with $\|g(X)\| \le U$ almost surely. We use the Hilbert-valued Rademacher complexity

$$\mathfrak{R}_N(\mathcal{G}) = \mathbb{E}_{X,\sigma} \sup_{g\in\mathcal{G}} \left\| \frac{1}{N}\sum_{i=1}^N \sigma_i g(X_i) \right\|,$$

where $\sigma_1, \ldots, \sigma_N$ are independent Rademacher signs. Standard symmetrization and bounded differences give, with probability at least $1-\eta$,

$$\sup_{g\in\mathcal{G}} \left\| \frac{1}{N}\sum_{i=1}^N (g(X_i) - Pg) \right\| \le 2\mathfrak{R}_N(\mathcal{G}) + U\sqrt{\frac{2\log(1/\eta)}{N}}. \tag{66}$$

If $\mathcal{G}$ induces at most L distinct value sequences on N fixed points, the Hilbert-space Hoeffding bound (Pinelis, 1994) and a union bound give

$$\mathfrak{R}_N(\mathcal{G}) \le CU\sqrt{\frac{\log(2L)}{N}} \tag{67}$$

for a universal constant C.

Application to each protected group. For $A^{(t)} = (A_1^{(t)}, \ldots, A_k^{(t)}) \in \mathcal{H}_t$, let $q_{A^{(t)}}(\mathbf{h})$ denote the unique label q for which $\mathbf{h} \in A_q^{(t)}$. Define

$$\phi_{p,t}^{A^{(t)}}(\mathbf{h}) = \mathbf{e}_{q_{A^{(t)}}(\mathbf{h})}, \tag{68}$$

$$\phi_{s,t}^{A^{(t)}}(\mathbf{h}) = \|\mathbf{h}\|_2^2\, \mathbf{e}_{q_{A^{(t)}}(\mathbf{h})}, \tag{69}$$

$$\phi_{m,t}^{A^{(t)}}(\mathbf{h}) = \mathbf{h} \otimes \mathbf{e}_{q_{A^{(t)}}(\mathbf{h})}. \tag{70}$$

Their envelopes are 1, B^2, and B. On n_t points, each value sequence is determined by a label sequence from $\mathcal{H}_t$, so there are at most $\Pi_{\mathcal{H}_t}(n_t)$ possibilities. Applying Eq. (66) and Eq. (67) and taking a union bound over the T groups and three statistic types yields, simultaneously for every t,

$$\sup_{A^{(t)} \in \mathcal{H}_t} \|\widehat{\boldsymbol{p}}_t - \boldsymbol{p}_t\|_2 \le C\alpha_t, \tag{71}$$

$$\sup_{A^{(t)} \in \mathcal{H}_t} \|\widehat{\boldsymbol{s}}_t - \boldsymbol{s}_t\|_2 \le CB^2\alpha_t, \tag{72}$$

$$\sup_{A^{(t)} \in \mathcal{H}_t} \left\|\widehat{\mathbf{M}}_t - \mathbf{M}_t\right\|_F \le CB\alpha_t, \tag{73}$$

where

$$\alpha_t = \sqrt{\frac{\log \Pi_{\mathcal{H}_t}(n_t) + \log(6T/\delta)}{n_t}}. \tag{74}$$

The global-stability bound then holds uniformly over every tuple $\boldsymbol{A} \in \mathbb{H}$, proving Theorem 3.

E.5 Growth-Function Specializations

Group-specific multiclass affine scores. Suppose each $\mathcal{H}_t$ is a class of partitions induced by

$$q_t(\mathbf{h}) = \arg\max_{j \in [k]} \langle \mathbf{a}_{tj}, (\mathbf{h}, 1) \rangle,$$

with a fixed deterministic tie rule. Viewing the $k(d+1)$ score parameters jointly, the $N\binom{k}{2}$ pairwise score comparisons define affine hyperplanes in parameter space. Classical hyperplane-arrangement counting (Cover, 1965) therefore gives universal constants $C_1, C_2 > 0$ such that

$$\log \Pi_{\mathcal{H}_t}(N) \le C_1 k(d+1) \log(C_2 eNk) \qquad \text{for every } t. \tag{75}$$

Substituting into Theorem 3 gives

$$\sup_{\boldsymbol{A} \in \mathbb{H}} |\widehat{R}(\boldsymbol{A}) - R(\boldsymbol{A})| \le CB^2 \sum_t \rho_t \sqrt{\frac{k(d+1)\log(C_2 e n_t k) + \log(6T/\delta)}{n_t}}. \tag{76}$$

A looser coordinatewise count gives

$$\log \Pi_{\mathcal{H}_t}(N) \le Ck^2(d+1)\log(CeNk), \tag{77}$$

which is sufficient as a conservative alternative.

These specializations make the dependence on the representation dimension explicit. In particular, the dimension need not remain fixed when the representation itself changes with the ensemble construction. For the raw binary association representation,

$$d = d_B = \sum_{r=1}^{m} k_r, \tag{78}$$

so the cluster-count schedule may induce a dimension $d = d(n)$ that grows with the sample size. In that case, the corresponding $d(n)$ should be substituted directly into the bounds above. Fixed-dimensional shorthand rates apply only when d is held constant.

E.6 Scope and Connection to Finite Optimization

Theorem 3 concerns the aggregation objective over the specified group-aware comparison class $\mathbb{H}$. For assignments induced by this class, Proposition 5 gives the exact identity

$$F(\mathbf{Y}_{\boldsymbol{A}}) = n\widehat{R}(\boldsymbol{A}). \tag{79}$$

Appendix D separately establishes descent and finite termination for coordinate descent over the full finite assignment set $\mathcal{Y}_k$. The two analyses therefore address different layers of the same center-eliminated aggregation objective: uniform estimation over a comparison class and finite local optimization.

The uniform-estimation result uses the fixed-representation sampling model of Section 4.2: the representation rule is fixed independently of the aggregation sample, after which the supplied representation vectors are sampled group-wise. This probability model does not describe the same-sample transductive construction in which the base partitions are learned from the data being aggregated. Proposition 5 remains an algebraic identity for any realized representation and induced assignment. For raw association vectors, boundedness follows from their binary construction; for other fixed representations, the condition $\|\mathbf{h}\|_2 \leq B$ specifies the analyzed sampling class.