The Confidence Cost of Pairwise Distribution Learning

Pahan Dewasurendra
Johns Hopkins University

Abstract

Pairwise conditional access to an unknown distribution returns a draw from the distribution restricted to any requested pair. At constant confidence, this weak oracle is as sample-efficient for labeled total-variation learning as ordinary sampling, up to constants. We show that this equivalence fails at high confidence.

For a distribution on n labels, accuracy ε, and failure probability δ, we prove the exact minimax query complexity

$$\Theta\left(\frac{n\log(1/\delta)}{\varepsilon^2}\right).$$

The corresponding ordinary-sampling complexity is $\Theta((n+\log(1/\delta))/\varepsilon^2)$. Thus confidence amplification costs up to a factor n under pairwise access.

The upper bound reuses one trajectory of a pairwise-winner Markov chain whose stationary law is the target. Although its worst-case relaxation time is linear in n, we prove that its empirical distribution has the i.i.d.-scale stationary risk $O(\sqrt{n/T})$. The proof compares its conductances with a minimum-mass graph. After sorting the target probabilities, this graph has an explicit Helmert eigensystem, and a capped-mass Green-function bound telescopes without any minimum-mass or dynamic-range assumption. A geometric median of independent trajectories gives the high-confidence result. The lower bound uses a heavy label and uniform light labels. Every adaptive pair query then carries only $O(\varepsilon^2/n)$ KL information, which makes the multiplicative confidence cost unavoidable.

1 Introduction

Suppose an unknown distribution p on $[n]$ cannot be sampled directly. Instead, a learner chooses two labels i, j and observes a draw from p conditioned on $\{i, j\}$. This pairwise conditional, or PCOND, oracle is the sampling form of the Bradley–Terry–Luce choice model (Bradley and Terry, 1952; Luce, 1959). It appears when providers expose only local menus, when comparisons are easier to elicit than calibrated scores, and when global samples cannot be observed.

Pairwise access identifies every labeled distribution, including targets with arbitrarily large dynamic range. The quantitative picture was recently clarified by Kleinberg et al. (2026). Their structural theory of restricted providers gives

$$\Omega\left(\frac{n+\log(1/\delta)}{\varepsilon^2}\right) \le q_{\text{pair}} \le O\left(\frac{n\log^2 n\log(n/\delta)}{\varepsilon^2}\right), \tag{1}$$

and asks whether the remaining polylogarithmic gap can be closed. Existing comparison estimators do not settle this question. Their useful uniform bounds generally impose a bounded parameter range, a fixed passive design, or a different loss on log qualities (Shah et al., 2015; Agarwal et al., 2018; Hendrickx et al., 2019, 2020).

We close the gap, but the answer differs from the ordinary-sampling benchmark. Pairwise learning has a multiplicative, not additive, dependence on confidence.

Theorem 1 (Minimax pairwise learning)**.** *There are universal constants $c, C > 0$ such that, for $n \ge 4$, $0 < \varepsilon \le 1/16$, and $0 < \delta \le 1/8$,*

$$c\frac{n\log(1/\delta)}{\varepsilon^2} \le q_{\text{pair}}(n,\varepsilon,\delta) \le C\frac{n\log(1/\delta)}{\varepsilon^2}. \tag{2}$$

The upper bound is attained by a polynomial-time adaptive learner and allows arbitrary zero probabilities.

For comparison, ordinary i.i.d. sampling has complexity $\Theta((n+\log(1/\delta))/\varepsilon^2)$ (Canonne, 2020). At constant confidence the two models both require $\Theta(n/\varepsilon^2)$ obser-

Table 1: Minimax labeled TV learning rates.

Observation model	query or sample complexity
Ordinary samples	$\Theta((n+\log(1/\delta))/\varepsilon^2)$
Pairwise conditional	$\Theta(n\log(1/\delta)/\varepsilon^2)$
Prior pairwise upper	$O(n\log^2 n\log(n/\delta)/\varepsilon^2)$

vations. If $\log(1/\delta) \gg n$, pairwise access requires a factor $\Theta(n)$ more. Table 1 summarizes the separation.

Why mixing time is misleading. The upper bound starts from a natural chain. From state i, propose a uniformly random $j \neq i$, query $\{i, j\}$, and retain the returned winner. Its stationary law is p. Restarting after each $O(n)$-step burn-in gives a quadratic learner. A single trajectory has worst-case relaxation time $\Theta(n)$, so a black-box spectral-gap bound still suggests quadratic cost. The slow mode, however, occurs only when its stationary mass is small or concentrated in a few coordinates. Total variation automatically weights these modes by their statistical relevance.

Our main technical result makes this compensation exact. We dominate the chain's Green function by that of the complete graph with edge conductance $\min(p_i, p_j)/n$. Once $p_1 \geq \cdots \geq p_m > 0$, this graph is diagonalized by the Helmert contrasts. Its eigenvalues are capped masses

$$c_k/n, \qquad c_k = kp_k + \sum_{j>k} p_j = \sum_j \min\{p_j, p_k\}.$$

The coordinate Green functions then split into left, local, and right terms. Each aggregate term is bounded by one telescoping inequality. This yields $O(\sqrt{n/T})$ TV risk uniformly over the whole simplex.

Why confidence is different. Take one label of mass near $1/2$ and spread the remaining mass uniformly over $n-1$ labels. Changing the heavy mass by $\Theta(\varepsilon)$ changes the outcome probability of every informative pair by only $\Theta(\varepsilon/n)$, while that outcome itself has probability $\Theta(1/n)$. Each query therefore contributes only $O(\varepsilon^2/n)$ KL divergence. This obstruction applies transcript by transcript to adaptive learners and forces the product $n\log(1/\delta)/\varepsilon^2$.

Contributions. Our results are entirely theoretical.

1. We determine the minimax PCOND complexity at every confidence level, closing the dimension and accuracy gap in (1).
2. We prove a distribution-free occupation-risk theorem for the winner chain. The result has no lower mass, bounded ratio, or warm-start assumption.
3. We identify a confidence separation between pairwise conditional and ordinary sampling, and prove that adaptive pair selection cannot remove it.

Algorithm 1 High-confidence pairwise learner

Set $C_0 = 2 + 2\sqrt{2}$, $B = \lceil (n-1)\log(48/\varepsilon)\rceil$, $T = \lceil (48^2 C_0^2 n/\varepsilon^2\rceil$, and $R = \lceil 18\log(1/\delta)\rceil$.
for $r = 1, \ldots, R$ **do**
  Start at an arbitrary $X_0 \in [n]$.
  for $t = 0, \ldots, B+T-1$ **do**
    Draw J_t uniformly from $[n] \setminus \{X_t\}$.
    Query $\{X_t, J_t\}$ and set X_{t+1} to its winner.
  end for
  Let $\widehat{p}_r$ be the empirical law of $X_{B+1}, \ldots, X_{B+T}$.
end for
return an L_1 geometric median of $\widehat{p}_1, \ldots, \widehat{p}_R$.

2 Model and Algorithm

Let Δ_n be the simplex on $[n]$. For $p \in \Delta_n$, a query to an unordered pair $\{i, j\}$ returns i with probability $p_i/(p_i+p_j)$ and j otherwise. If both masses vanish, the response may follow any fixed convention. A learner may choose each pair using its past queries, responses, and private randomness.

We define $q_{\mathrm{pair}}(n, \varepsilon, \delta)$ as the least fixed query budget for which some learner returns $\widehat{p}$ satisfying

$$\sup_{p\in\Delta_n} \mathbb{P}_p\{d_{\mathrm{TV}}(\widehat{p}, p) > \varepsilon\} \leq \delta,$$
$$d_{\mathrm{TV}}(p, q) = \tfrac{1}{2}\|p - q\|_1.$$

One base estimator follows a single winner-chain trajectory. Repeating it and aggregating by a metric median gives Algorithm 1. An L_1 geometric median is any minimizer over Δ_n of the sum of distances to the candidates. It is computable by a linear program.

The stated constants simplify the proof and are not optimized. Since $\log(48/\varepsilon) = O(1/\varepsilon^2)$, the total query count has the order in Theorem 1.

2.1 What the algorithm changes

The global chain itself is not new. It appears in the generic complete-family learner of Kleinberg et al. (2026). The difference is which states are kept and how they are analyzed. That learner runs the chain for $B = \Theta(n\log(1/\varepsilon))$ steps, keeps only the terminal state, and restarts independently $M = \Theta((n+\log(1/\delta))/\varepsilon^2)$ times. Even in the pairwise case, this costs

$$\Theta\left(\frac{n(n+\log(1/\delta))\log(1/\varepsilon)}{\varepsilon^2}\right)$$

queries. Their faster hierarchical learner keeps trajectories inside many tree cells, then reconstructs p from binary split estimates. It pays two tree-depth logarithms and a simultaneous-confidence logarithm.

Algorithm 1 instead burns in one global chain and retains every later state. This removes repeated mixing, but standard MCMC analysis does not justify the improvement. The uniform minorization in (17) gives gap only $1/(n-1)$. Applying that gap to each coordinate and summing gives the unhelpful bound $O(n/\sqrt{T})$. The next section replaces this worst-mode argument by a joint coordinate calculation.

3 One Trajectory Has i.i.d.-Scale Risk

We first analyze a stationary trajectory. On the positive support of p, the winner chain has transition kernel

$$P(i,j) = \frac{1}{n-1}\frac{p_j}{p_i+p_j}, \quad i \neq j, \tag{3}$$

with the diagonal chosen to make rows sum to one. Detailed balance holds because

$$p_i P(i,j) = \frac{p_i p_j}{(n-1)(p_i+p_j)} = p_j P(j,i). \tag{4}$$

Theorem 2 (Stationary occupation risk)**.** *Let $(X_t)_{t=1}^T$ be a stationary trajectory of* (3)*, and let $\widehat{p}_T$ be its empirical law. For every $p \in \Delta_n$,*

$$\mathbb{E}_p d_{\mathrm{TV}}(\widehat{p}_T, p) \le (2+2\sqrt{2})\sqrt{\frac{n}{T}}. \tag{5}$$

The result is not a consequence of the worst-case gap, which can be $\Theta(1/n)$. We prove it through coordinate-specific Green functions. Let $f_i(x) = \mathbf{1}\{x = i\} - p_i$ and

$$H_i = \langle f_i, (I-P)^{-1} f_i \rangle_p, \tag{6}$$

where the inverse acts on mean-zero functions in $L_2(p)$.

There is a useful electrical interpretation. Let $D = \mathrm{diag}(p)$ and let $L = D(I-P)$ be the weighted graph Laplacian associated with (4). Since

$$Df_i = p_i(e_i - p) =: b_i,$$

the coordinate Green function is

$$H_i = b_i^\top L^\dagger b_i. \tag{7}$$

Thus H_i is the minimum energy needed to send source $p_i(1-p_i)$ from i and absorb $p_i p_j$ at every $j \neq i$. A small global gap can coexist with small coordinate energy because each source is scaled by its target mass.

Lemma 3 (Positive spectrum and variance)**.** *The kernel P is positive semidefinite on $L_2(p)$. Moreover,*

$$\mathrm{Var}_p\left(\frac{1}{T}\sum_{t=1}^T f_i(X_t)\right) \le \frac{2H_i}{T}. \tag{8}$$

For positive semidefiniteness, use

$$\frac{p_i p_j}{p_i+p_j}(f_i - f_j)^2 \le p_i f_i^2 + p_j f_j^2$$

and sum over pairs. Reversibility then gives eigenvalues in $[0,1]$. Expanding an additive functional in the eigenbasis and summing its geometric covariances proves (8).

It remains to control all H_i jointly. Let $m = |\mathrm{supp}(p)|$ and relabel the positive masses so that $p_1 \ge \cdots \ge p_m > 0$. Define

$$c_k = kp_k + \sum_{j>k} p_j. \tag{9}$$

Lemma 4 (Capped-mass Green bound)**.** *For $i \in [m]$, define*

$$A_i = \sum_{k=2}^{i-1} \frac{(1-c_k)^2}{k(k-1)c_k} + \mathbf{1}\{i \ge 2\}\frac{(i-c_i)^2}{i(i-1)c_i} + \sum_{k=i+1}^{m} \frac{c_k}{k(k-1)}. \tag{10}$$

Then

$$H_i \le 2np_i^2 A_i, \qquad \sum_{i=1}^m p_i\sqrt{A_i} \le 2+2\sqrt{2}. \tag{11}$$

Proof idea. The Dirichlet conductance in (4) obeys

$$\frac{p_i p_j}{(n-1)(p_i+p_j)} \ge \frac{\min\{p_i,p_j\}}{2n}. \tag{12}$$

Let L_0 be the graph Laplacian with conductance $\min\{p_i,p_j\}/n$. The Helmert vectors

$$u_k = \frac{(\underbrace{1,\ldots,1}_{k-1}, -(k-1), 0, \ldots, 0)}{\sqrt{k(k-1)}} \tag{13}$$

diagonalize L_0, with $L_0 u_k = (c_k/n)u_k$. Projecting the coordinate source $p_i(e_i - p)$ onto this basis gives (10) exactly. The factor two in (11) follows from (12).

For completeness, the projections are explicit. One has

$$u_k^\top p = \frac{1-c_k}{\sqrt{k(k-1)}},$$

and hence

$$\frac{u_k^\top b_i}{p_i} = \begin{cases} -(1-c_k)/\sqrt{k(k-1)}, & k < i, \\ -(i-c_i)/\sqrt{i(i-1)}, & k = i, \\ c_k/\sqrt{k(k-1)}, & k > i. \end{cases} \tag{14}$$

Dividing their squares by c_k/n produces the three terms of (10). Since the exact conductances dominate

half of those of L_0, the inverse-Laplacian variational principle yields $L^\dagger \preceq 2L_0^\dagger$ on the zero-sum subspace. This proves the first part of Lemma 4.

For the aggregate bound, split A_i into the three sums in (10). Jensen's inequality and a change in summation order bound the first and third contributions by one each. For example,

$$\sum_i p_i \sqrt{\sum_{k<i} \frac{(1-c_k)^2}{k(k-1)c_k}} \le \sqrt{\sum_{k\ge 2} \frac{1}{k(k-1)}} \le 1,$$

where $\sum_{i>k} p_i \le c_k$. For the middle term, let $S_i = \sum_{j\ge i} p_j$. Since $c_i \ge S_i$,

$$\sum_{i\ge 2} p_i \frac{i-c_i}{\sqrt{i(i-1)c_i}} \le \sqrt{2} \sum_i \frac{p_i}{\sqrt{S_i}} \le 2\sqrt{2}.$$

The last inequality telescopes using $(S_i - S_{i+1})/\sqrt{S_i} \le 2(\sqrt{S_i} - \sqrt{S_{i+1}})$. The supplement gives every projection and boundary case.

Corollary 5 (Profile-aware risk)**.** *With A_i as in* (10)*, every stationary trajectory satisfies*

$$\mathbb{E} d_{\mathrm{TV}}(\widehat{p}_T, p) \le \sqrt{\frac{n}{T}}\, \Psi(p), \qquad \Psi(p) := \sum_{i \in \mathrm{supp}(p)} p_i \sqrt{A_i} \le 2 + 2\sqrt{2}. \tag{15}$$

Proof of Theorem 2. By Lemmas 3 and 4,

$$\mathbb{E} d_{\mathrm{TV}}(\widehat{p}_T, p) \le \frac{1}{2} \sum_i \sqrt{\mathrm{Var}(\widehat{p}_T(i))} \le \sqrt{\frac{n}{T}} \sum_i p_i \sqrt{A_i} \le (2+2\sqrt{2}) \sqrt{\frac{n}{T}}.$$

3.1 Examples and the slow mode

The profile functional in (15) reveals why a global gap bound is loose. If p is uniform on all n labels, then $c_k = 1$ and direct substitution gives

$$A_i = 1 - \frac{1}{n}, \qquad \Psi(p) = \sqrt{1 - \frac{1}{n}}.$$

The winner chain is a lazy complete-graph walk. Its TV occupation risk is therefore bounded by $\sqrt{(n-1)/T}$, at the ordinary empirical-distribution scale.

At the other extreme, take

$$p_1 = \frac{1}{2}, \qquad p_j = \frac{1}{2(n-1)}, \quad j \ge 2. \tag{16}$$

The indicator $Y_t = \mathbf{1}\{X_t = 1\}$ is itself a two-state Markov chain. From either aggregate state it flips with probability exactly $1/n$. Its nonconstant eigenvalue is $1 - 2/n$, so estimating the heavy mass from one trajectory has variance of order n/T. This is a genuine slow mode. Nevertheless, there is only one such aggregate coordinate. The remaining $n-1$ light coordinates leave their individual states at constant rate. The sum of all coordinate standard deviations remains $O(\sqrt{n/T})$, exactly as Lemma 4 certifies.

For a nearly point-mass target, the slow aggregate mode becomes still slower, but its stationary variance shrinks in the same proportion. In capped-mass language, the local denominator c_k decreases only where the corresponding tail mass also decreases. The left and middle telescopes in the proof are the finite-dimensional expression of this compensation.

These examples also separate estimation from independent sampling. The uniform target behaves nearly as if the retained states were independent. The half-heavy target has only about T/n effective observations of its heavy-versus-light split. That split is precisely the source of the high-confidence lower bound in Section 5.

4 Burn-in and Confidence Amplification

The winner chain admits the pointwise minorization

$$P(i, \cdot) \ge \frac{1}{n-1} p(\cdot). \tag{17}$$

This also holds from a zero-mass state. Hence, from every initial law μ,

$$d_{\mathrm{TV}}(\mu P^B, p) \le e^{-B/(n-1)}. \tag{18}$$

Applying the same Markov kernel to two initial states cannot increase TV, so (18) also bounds the distance between the entire post-burn path and a stationary path.

With the B, T in Algorithm 1, Theorem 2 and (18) give

$$\mathbb{E} d_{\mathrm{TV}}(\widehat{p}_r, p) \le \varepsilon/24.$$

Markov's inequality shows that each candidate is within $\varepsilon/4$ with probability at least $5/6$. Hoeffding's inequality implies that at least $2R/3$ candidates are good except with probability $e^{-R/18} \le \delta$.

If G of R points in any metric space lie within radius r of p, where $G > R/2$, every geometric median q satisfies

$$d(q, p) \le \frac{2G}{2G - R} r. \tag{19}$$

Indeed, compare its sum of distances with the sum at p and apply the triangle inequality separately to the

good and bad points. With $G \geq 2R/3$ and $r = \varepsilon/4$, (19) proves the upper bound in Theorem 1.

Computation. The chain uses one oracle call and $O(1)$ additional work per transition if a uniform challenger can be sampled in constant time. The candidate histograms require $O(Rn)$ storage in a direct implementation. The geometric median in (19) is the linear program

$$\min_{q \in \Delta_n, z \geq 0} \sum_{r,i} z_{ri} \quad \text{subject to} \quad -z_{ri} \leq q_i - \widehat{p}_r(i) \leq z_{ri}.$$

Thus the statistical upper bound is also polynomial-time. Approximate minimization to $o(\varepsilon R)$ preserves the guarantee up to a constant change in accuracy.

Why ordinary confidence is additive. For T i.i.d. samples, the empirical TV risk is $O(\sqrt{n/T})$. Changing one sample changes TV loss by at most $1/T$, so bounded differences adds only $O(\sqrt{\log(1/\delta)/T})$ to this expectation (Canonne, 2020). This gives the sum $n + \log(1/\delta)$. Winner-chain states do not have this product concentration. The half-heavy example in (16) retains one transition for $\Theta(n)$ steps on average. The lower bound below shows that no different adaptive estimator can bypass the resulting confidence cost.

5 The Confidence Cost Is Necessary

We now show that no adaptive pair selection can attain the additive confidence dependence of ordinary sampling. For $a \in (0,1)$, define

$$p_a(1) = a, \qquad p_a(j) = \frac{1-a}{n-1}, \quad j \geq 2. \tag{20}$$

Take $a_- = 1/2 - 2\varepsilon$ and $a_+ = 1/2 + 2\varepsilon$. Then $d_{\mathrm{TV}}(p_{a_-}, p_{a_+}) = 4\varepsilon$.

A query not containing label 1 returns a fair draw under both hypotheses. For $\{1, j\}$, let the response be one when the light label wins. Its Bernoulli parameter is

$$r(a) = \frac{1-a}{1+(n-2)a}. \tag{21}$$

On $a \in [1/4, 3/4]$, one has $r(a) = \Theta(1/n)$ and

$$|r'(a)| = \frac{n-1}{(1+(n-2)a)^2} = O(1/n). \tag{22}$$

The elementary Bernoulli bound $D_{\mathrm{KL}}(x\|y) \leq (x-y)^2/[y(1-y)]$ therefore gives

$$D_{\mathrm{KL}}(\mathrm{Bern}(r(a_-))\|\,\mathrm{Bern}(r(a_+))) \leq C_1 \frac{\varepsilon^2}{n}. \tag{23}$$

For an adaptive learner, condition on each realized transcript before the next query. Every possible pair obeys (23), so the KL chain rule bounds the divergence between length-Q transcript laws by $C_1 Q\varepsilon^2/n$.

If an estimator is ε accurate with probability $1-\delta$ under both hypotheses, choosing the closer of p_{a_-} and p_{a_+} produces a test whose two error probabilities are at most δ. The Bretagnolle–Huber inequality then requires transcript KL at least $\log(1/(4\delta))$. Consequently,

$$Q \geq c_1 \frac{n \log(1/\delta)}{\varepsilon^2},$$

which proves the lower bound in Theorem 1.

6 Bounded Menus Interpolate the Obstruction

The confidence phenomenon is not unique to pairs. It weakens continuously when a query may expose more light labels at once. Let $q_{\leq k}(n, \varepsilon, \delta)$ denote minimax labeled TV complexity when the learner may condition on any nonempty menu of size at most k.

Proposition 6 (Bounded-menu confidence lower bound)**.** *For $2 \leq k \leq n$ and the parameter ranges of Theorem 1,*

$$q_{\leq k}(n, \varepsilon, \delta) = \Omega\left(\frac{n + (n/k)\log(1/\delta)}{\varepsilon^2}\right). \tag{24}$$

Proof. The $\Omega(n/\varepsilon^2)$ term already holds with unrestricted conditional access, since learning all labels is no easier than the standard multinomial packing used for ordinary samples (Canonne, 2020; Kleinberg et al., 2026).

For confidence, use the family (20). A menu excluding label 1 is uniform under both hypotheses. Suppose a menu contains label 1 and $s \leq k-1$ light labels. Conditional on a light outcome, its identity is uniform under both hypotheses, so all information is in the Bernoulli light probability

$$r_s(a) = \frac{s(1-a)}{(n-1)a + s(1-a)}. \tag{25}$$

Uniformly for $a \in [1/4, 3/4]$, one has $r_s(a) = \Theta(s/n)$ and

$$|r_s'(a)| = \frac{s(n-1)}{((n-1)a + s(1-a))^2} = O(s/n).$$

The Bernoulli KL is therefore $O(s\varepsilon^2/n)$, and hence at most $O(k\varepsilon^2/n)$, for every conditional menu and every transcript. The KL chain rule and Bretagnolle–Huber argument from Section 5 require $Q = \Omega((n/k)\log(1/\delta)/\varepsilon^2)$. Taking the maximum of the two lower bounds proves (24). □

For $k = 2$, Theorem 1 matches (24). For $k = n$, the lower bound becomes the ordinary additive rate. This identifies a candidate interpolation law, but we do not claim its upper bound for intermediate k. A direct k-winner chain has stationary law p and a stronger minorization, yet its coordinate Green function is no longer the minimum-mass graph in Lemma 4. Determining whether its occupation profile attains (24) is a separate open problem.

7 Related Work and Scope

Restricted conditional learning. Conditional sampling was introduced for distribution testing (Canonne et al., 2015; Chakraborty et al., 2016). The PCOND restriction is powerful for many label-invariant testing problems, but our target is the entire labeled distribution. Kleinberg et al. (2026) characterize which fixed menu families permit pointwise and uniform labeled learning. Their global and local winner chains motivate ours. Their global learner restarts after every terminal sample, while their hierarchical learner keeps local trajectories to estimate tree splits. Our result instead controls the full empirical law of one global pairwise trajectory through its complete mass profile. It both removes the dimension logarithms and supplies the matching confidence lower bound.

Comparison estimation and sampling. Classical BTL work estimates rankings, log qualities, or normalized weights under passive comparison graphs. Sharp bounds depend on graph Laplacians, effective resistance, and often bounded dynamic range (Shah et al., 2015; Negahban et al., 2017; Agarwal et al., 2018; Hendrickx et al., 2019, 2020). Our active estimator permits all simplex boundary points and is analyzed in labeled TV. Fotakis et al. (2022) construct exact samples from exogenously drawn pair comparisons using coupling from the past. Their distribution-free method first learns every weight to relative accuracy and solves a different exact-sampling objective.

Biased sampling and Markov concentration. Conditional menus are a finite selection-bias model. Classical work proves identifiability and asymptotic efficiency for fixed biased-sampling designs (Vardi, 1985; Gill et al., 1988). Generic reversible-chain concentration scales with a worst-case spectral gap (Jiang et al., 2018). That route loses a factor n here. The capped mass calculation is coordinate-specific and nonasymptotic.

Limits and open directions. Theorem 1 concerns the family of all pairs. It does not close the polylogarithmic gaps for every hierarchically comparable menu family. Proposition 6 gives a broader necessary interpolation, but whether it is sufficient for every menu cap remains open. The present algorithm also assumes that every pair is available. Graph-restricted active designs may require additional connectivity and dynamic-range parameters.

8 Conclusion

Pairwise conditional samples retain enough aggregate information to learn a labeled distribution at the ordinary n/ε^2 rate at constant confidence. They do not retain the ordinary model's cheap confidence amplification. Reusing a single winner-chain trajectory attains the optimal constant-confidence rate, while independent trajectories and a metric median match the unavoidable multiplicative confidence cost. The proof explains why a linear relaxation time does not imply a linear loss in estimation risk: the slow modes and the TV sensitivity are coupled through the target's capped mass profile.

A Pairwise Winner Chain

Fix a distribution $p \in \Delta_n$. From state i, the chain chooses $j \neq i$ uniformly, queries the pair $\{i, j\}$, and makes the returned label the next state. If $p_i + p_j > 0$, then

$$P(i,j) = \frac{p_j}{(n-1)(p_i+p_j)}, \qquad i \neq j. \tag{26}$$

If $p_i = p_j = 0$, the pair response may follow any fixed convention. The diagonal is the remaining row mass.

Lemma 7 (Stationarity, reversibility, and minorization)**.** *The distribution p is stationary. On its positive support the chain is reversible, and for every state i,*

$$P(i,\cdot) \geq \frac{1}{n-1} p(\cdot) \tag{27}$$

coordinatewise.

Proof. For distinct positive-mass states i, j,

$$p_i P(i,j) = \frac{p_i p_j}{(n-1)(p_i+p_j)} = p_j P(j,i).$$

Transitions from a positive-mass state to a zero-mass state have probability zero. Detailed balance on the support proves reversibility and stationarity.

For the minorization, first take $j \neq i$. If $p_i + p_j > 0$, then

$$P(i,j) = \frac{1}{n-1}\frac{p_j}{p_i+p_j} \geq \frac{p_j}{n-1},$$

since $p_i + p_j \leq 1$. If both masses vanish, the right side is zero. If $p_i = 0$, the diagonal inequality is immediate. If $p_i > 0$, then against every challenger j, the probability of remaining at i is at least $p_i/(p_i+p_j) \geq p_i$. Averaging over challengers gives $P(i,i) \geq p_i \geq p_i/(n-1)$. □

For $n > 2$, the minorization gives a Markov kernel R such that

$$P = \frac{1}{n-1}\Pi_p + \left(1 - \frac{1}{n-1}\right) R, \tag{28}$$

where every row of Π_p is p. Thus, for any initial law μ,

$$d_{\mathrm{TV}}(\mu P^B, p) \leq \left(1 - \frac{1}{n-1}\right)^B \leq e^{-B/(n-1)}. \tag{29}$$

For $n = 2$, one query returns an exact p draw from either state.

B Positive Spectrum and Coordinate Variance

We work on the positive support of p. Zero-mass coordinates contribute nothing to TV under stationarity. Let $m = |\operatorname{supp}(p)|$. For a function $f \in L_2(p)$, reversibility gives the Dirichlet form

$$\langle f, (I-P)f\rangle_p = \sum_{1 \leq i < j \leq m} \frac{p_i p_j}{(n-1)(p_i+p_j)}(f_i - f_j)^2. \tag{30}$$

Lemma 8 (Positive semidefinite kernel)**.** *Every eigenvalue of P on $L_2(p)$ lies in $[0,1]$.*

Proof. The Markov and reversibility properties put the spectrum in $[-1,1]$. For each positive pair,

$$\frac{p_i p_j}{p_i+p_j}(f_i - f_j)^2 \leq p_i f_i^2 + p_j f_j^2. \tag{31}$$

Indeed, the right side minus the left side is $(p_i f_i + p_j f_j)^2/(p_i+p_j)$. Each coordinate appears in at most $m-1$ positive pairs. Summing (31) and dividing by $n-1$ gives

$$\langle f, (I-P)f\rangle_p \leq \langle f, f\rangle_p.$$

Hence $\langle f, Pf\rangle_p \geq 0$ for every f. □

For coordinate i, put $f_i(x) = \mathbf{1}\{x = i\} - p_i$ and

$$H_i = \langle f_i, (I-P)^{-1} f_i\rangle_p, \tag{32}$$

where the inverse acts on the mean-zero subspace.

Lemma 9 (Finite-horizon coordinate variance)**.** *For a stationary length-T trajectory,*

$$\operatorname{Var}_p\left(\frac{1}{T}\sum_{t=1}^{T} f_i(X_t)\right) \leq \frac{2H_i}{T}. \tag{33}$$

Proof. Expand f_i in an orthonormal eigenbasis of the mean-zero subspace. If a^2 is the squared coefficient associated with eigenvalue $\rho \in [0,1)$, its contribution to the variance is

$$\frac{a^2}{T^2}\left(T + 2\sum_{s=1}^{T-1}(T-s)\rho^s\right) \leq \frac{2a^2}{T(1-\rho)}.$$

Summing the contributions gives (33) by the spectral definition of H_i. □

C An Explicit Comparison Green Function

Relabel the positive probabilities so that $p_1 \geq \cdots \geq p_m > 0$. Define the graph Laplacian L_0 on $[m]$ whose

edge conductances are

$$w_{ij}^0 = \frac{\min\{p_i, p_j\}}{n}. \tag{34}$$

For $k \in [m]$, define

$$c_k = kp_k + \sum_{j>k} p_j = \sum_{j=1}^{m} \min\{p_j, p_k\}. \tag{35}$$

Lemma 10 (Helmert diagonalization)**.** *For $k = 2, \dots, m$, let*

$$u_k = \frac{(\underbrace{1, \dots, 1}_{k-1}, -(k-1), 0, \dots, 0)}{\sqrt{k(k-1)}}. \tag{36}$$

The vectors $(u_k)_{k=2}^m$ are an orthonormal basis of the Euclidean zero-sum subspace and

$$L_0 u_k = \frac{c_k}{n} u_k. \tag{37}$$

Proof. Orthonormality is the standard Helmert calculation. Fix k. Every coordinate larger than k sees equal conductance p_j/n to all first k coordinates, whose entries in u_k sum to zero. Its image is therefore zero. A coordinate $i < k$ differs only from coordinate k among the first k, contributing kp_k/n. Every coordinate $j > k$ contributes p_j/n. Thus the ith image is $(c_k/n)u_k(i)$. The same sum at coordinate k is multiplied by $-(k-1)$, which proves (37). □

Let $D = \mathrm{diag}(p)$. The graph Laplacian associated with (30) is $L = D(I - P)$. For the centered indicator f_i,

$$Df_i = p_i(e_i - p) =: b_i. \tag{38}$$

Therefore

$$H_i = b_i^\top L^\dagger b_i. \tag{39}$$

The exact conductance of L satisfies

$$\frac{p_i p_j}{(n-1)(p_i + p_j)} \geq \frac{\min\{p_i, p_j\}}{2(n-1)} \geq \frac{1}{2} w_{ij}^0. \tag{40}$$

It follows by the variational characterization of inverse Laplacians that

$$H_i \leq 2 b_i^\top L_0^\dagger b_i. \tag{41}$$

Lemma 11 (Exact diagonal formula)**.** *Define*

$$A_i = \sum_{k=2}^{i-1} \frac{(1-c_k)^2}{k(k-1)c_k} + \mathbf{1}\{i \geq 2\} \frac{(i-c_i)^2}{i(i-1)c_i} + \sum_{k=i+1}^{m} \frac{c_k}{k(k-1)}. \tag{42}$$

Then

$$b_i^\top L_0^\dagger b_i = n p_i^2 A_i, \qquad H_i \leq 2n p_i^2 A_i. \tag{43}$$

Proof. First observe from (35) and (36) that

$$u_k^\top p = \frac{1 - c_k}{\sqrt{k(k-1)}}. \tag{44}$$

Since $b_i = p_i(e_i - p)$, its projection onto u_k is p_i times

$$u_k(i) - u_k^\top p = \begin{cases} -(1-c_k)/\sqrt{k(k-1)}, & k < i, \\ -(i-c_i)/\sqrt{i(i-1)}, & k = i, \\ c_k/\sqrt{k(k-1)}, & k > i. \end{cases} \tag{45}$$

Divide the squared projections by the eigenvalues c_k/n from Lemma 10 and sum over k. This gives the equality in (43). The inequality follows from (41). □

D The Profile Sum Telescopes

Write $A_i = L_i + M_i + R_i$ for the three terms in (42). We prove the aggregate inequality used in the main paper.

Lemma 12 (Uniform profile bound)**.** *For every non-increasing positive probability vector,*

$$\sum_{i=1}^{m} p_i \sqrt{A_i} \leq 2 + 2\sqrt{2}. \tag{46}$$

Proof. By Jensen's inequality and a change in summation order,

$$\sum_i p_i \sqrt{L_i} \leq \sqrt{\sum_i p_i L_i} \tag{47}$$

$$= \sqrt{\sum_{k=2}^{m-1} \frac{(1-c_k)^2}{k(k-1)c_k} \sum_{i>k} p_i} \tag{48}$$

$$\leq \sqrt{\sum_{k=2}^{m-1} \frac{1}{k(k-1)}} \leq 1. \tag{49}$$

The last line uses $\sum_{i>k} p_i \leq c_k$ and $1 - c_k \leq 1$. Similarly,

$$\sum_i p_i \sqrt{R_i} \leq \sqrt{\sum_i p_i R_i} \tag{50}$$

$$= \sqrt{\sum_{k=2}^{m} \frac{c_k}{k(k-1)} \sum_{i<k} p_i} \tag{51}$$

$$\leq \sqrt{\sum_{k=2}^{m} \frac{1}{k(k-1)}} \leq 1. \tag{52}$$

For the middle term, define $S_i = \sum_{j \geq i} p_j$. Since $c_i = ip_i + \sum_{j>i} p_j \geq S_i$ and $i/(i-1) \leq 2$,

$$\sum_i p_i \sqrt{M_i} \leq \sqrt{2} \sum_i \frac{p_i}{\sqrt{c_i}} \leq \sqrt{2} \sum_i \frac{S_i - S_{i+1}}{\sqrt{S_i}}. \tag{53}$$

For $0 \le b \le a$,

$$\frac{a-b}{\sqrt{a}} \le 2(\sqrt{a}-\sqrt{b}). \tag{54}$$

Apply this with $a = S_i, b = S_{i+1}$ and telescope to obtain

$$\sum_i p_i \sqrt{M_i} \le 2\sqrt{2}. \tag{55}$$

Finally, $\sqrt{L_i + M_i + R_i} \le \sqrt{L_i} + \sqrt{M_i} + \sqrt{R_i}$. Combining (49), (52), and (55) proves the claim. □

E Stationary Occupation Risk

Theorem 13 (Stationary risk)**.** *For a stationary length-T winner-chain trajectory with empirical law $\widehat{p}_T$,*

$$\mathbb{E} d_{\mathrm{TV}}(\widehat{p}_T, p) \le (2 + 2\sqrt{2})\sqrt{\frac{n}{T}}. \tag{56}$$

Proof. If the support has size one, the claim is immediate. Otherwise, Lemmas 9 and 11 give

$$\mathrm{Var}(\widehat{p}_T(i)) \le \frac{4np_i^2 A_i}{T}.$$

Jensen's inequality and Lemma 12 yield

$$\begin{aligned}\mathbb{E} d_{\mathrm{TV}}(\widehat{p}_T, p) &\le \frac{1}{2}\sum_i \mathbb{E}|\widehat{p}_T(i) - p_i| \\ &\le \frac{1}{2}\sum_i \sqrt{\mathrm{Var}(\widehat{p}_T(i))} \\ &\le \sqrt{\frac{n}{T}}\sum_i p_i\sqrt{A_i} \le (2+2\sqrt{2})\sqrt{\frac{n}{T}}.\end{aligned}$$

Zero-mass coordinates are never visited under stationarity and add zero. □

F Arbitrary Starts and High Confidence

Set $C_0 = 2 + 2\sqrt{2}$ and

$$B = \left\lceil (n-1)\log\frac{48}{\varepsilon} \right\rceil, \qquad T = \left\lceil 48^2 C_0^2 \frac{n}{\varepsilon^2} \right\rceil. \tag{57}$$

Start the winner chain from an arbitrary state and form its empirical law $\widehat{p}$ over the T states after burn-in.

Lemma 14 (One base estimate)**.** *The estimate satisfies*

$$\mathbb{P}\{d_{\mathrm{TV}}(\widehat{p}, p) > \varepsilon/4\} \le 1/6. \tag{58}$$

Proof. By (29), the state after burn-in is within $\varepsilon/48$ TV of stationarity. Generate the entire subsequent path by applying one common path kernel to either initial law. Data processing shows that the arbitrary-start and stationary path laws remain within $\varepsilon/48$ TV. Since TV loss lies in $[0,1]$, their expected losses differ by at most $\varepsilon/48$. Theorem 13 and the choice of T give

$$\mathbb{E} d_{\mathrm{TV}}(\widehat{p}, p) \le \varepsilon/48 + \varepsilon/48 = \varepsilon/24.$$

Markov's inequality proves (58). □

Repeat this construction independently R times and let $\widehat{p}_r$ be the candidates. Define an L_1 geometric median

$$\widetilde{p} \in \arg\min_{q \in \Delta_n} \sum_{r=1}^{R} d_{\mathrm{TV}}(q, \widehat{p}_r). \tag{59}$$

Existence follows from compactness. The optimization is a linear program after introducing one absolute-value slack variable per candidate and coordinate.

Lemma 15 (Metric median)**.** *In any metric space, suppose $G > R/2$ of the points $z_1, \dots, z_R$ lie within distance r of p. Every minimizer z of the sum of distances to the points satisfies*

$$d(z,p) \le \frac{2G}{2G-R} r. \tag{60}$$

Proof. Let $\mathcal{G}$ denote the good indices and $d = d(z,p)$. For a good point,

$$d(z, z_r) - d(p, z_r) \ge d - 2r.$$

For a bad point, the triangle inequality gives $d(z,z_r) - d(p,z_r) \ge -d$. Since the sum at z is no larger than the sum at p,

$$0 \ge G(d-2r) - (R-G)d = (2G-R)d - 2Gr.$$

Rearranging proves the result. □

Theorem 16 (High-confidence upper bound)**.** *If $R = \lceil 18\log(1/\delta)\rceil$, then*

$$\mathbb{P}\{d_{\mathrm{TV}}(\widetilde{p}, p) > \varepsilon\} \le \delta. \tag{61}$$

The total number of pair queries is $O(n\log(1/\delta)/\varepsilon^2)$.

Proof. Every candidate is good with probability at least $5/6$ by Lemma 14. Hoeffding's inequality gives

$$\mathbb{P}\{G < 2R/3\} \le e^{-R/18} \le \delta.$$

On the complementary event, apply Lemma 15 with $r = \varepsilon/4$. The multiplicative factor in (60) is at most four, so the output error is at most ε.

Each repetition uses $B + T$ queries. For $0 < \varepsilon \le 1/16$, $\log(48/\varepsilon) = O(1/\varepsilon^2)$, hence $B + T = O(n/\varepsilon^2)$. Multiplying by R proves the query bound. □

G Adaptive Confidence Lower Bound

Let $n \geq 4$, $0 < \varepsilon \leq 1/16$, and define

$$p_a(1) = a, \qquad p_a(j) = \frac{1-a}{n-1}, \quad j \geq 2. \tag{62}$$

Take $a_- = 1/2 - 2\varepsilon$ and $a_+ = 1/2 + 2\varepsilon$. Both lie in $[1/4, 3/4]$, and

$$d_{\mathrm{TV}}(p_{a_-}, p_{a_+}) = |a_- - a_+| = 4\varepsilon. \tag{63}$$

If a queried pair excludes label 1, its response is uniform under both hypotheses. If the pair is $\{1, j\}$, write one for a light-label response. The Bernoulli parameter is

$$r(a) = \frac{1-a}{1+(n-2)a}. \tag{64}$$

For $a \in [1/4, 3/4]$ and $n \geq 4$,

$$\frac{1}{3n} \leq r(a) \leq \frac{3}{n}, \qquad |r'(a)| = \frac{n-1}{(1+(n-2)a)^2} \leq \frac{16}{n}. \tag{65}$$

Also $1 - r(a) \geq 1/4$. By the mean value theorem and the Bernoulli inequality

$$D_{\mathrm{KL}}(\mathrm{Bern}(x) \| \mathrm{Bern}(y)) \leq \frac{(x-y)^2}{y(1-y)}, \tag{66}$$

every possible pair query has conditional KL divergence at most

$$C_1 \frac{\varepsilon^2}{n} \tag{67}$$

between the two response laws, for a universal C_1.

Theorem 17 (Adaptive lower bound)**.** *Any possibly randomized and adaptive learner using a fixed budget Q and satisfying*

$$\sup_{p \in \Delta_n} \mathbb{P}_p\{d_{\mathrm{TV}}(\widehat{p}, p) > \varepsilon\} \leq \delta \tag{68}$$

must have

$$Q \geq c \frac{n \log(1/\delta)}{\varepsilon^2} \tag{69}$$

for $0 < \delta \leq 1/8$ and a universal $c > 0$.

Proof. Include the learner's private random seed in its transcript. Conditional on every past transcript, its next pair is fixed. The conditional response laws then obey (67), regardless of which pair was selected. The KL chain rule gives

$$D_{\mathrm{KL}}(\mathsf{P}_-^Q \| \mathsf{P}_+^Q) \leq C_1 Q \frac{\varepsilon^2}{n}, \tag{70}$$

where $\mathsf{P}_-^Q, \mathsf{P}_+^Q$ are the complete transcript laws.

Given the output $\widehat{p}$, test the two hypotheses by choosing the closer of p_{a_-} and p_{a_+} in TV. By (63), every estimate within ε of the truth is strictly closer to that truth. Thus the two testing errors are at most δ. The Bretagnolle–Huber inequality (Tsybakov, 2009) gives

$$2\delta \geq \frac{1}{2} \exp\{-D_{\mathrm{KL}}(\mathsf{P}_-^Q \| \mathsf{P}_+^Q)\}. \tag{71}$$

Combining (70) and (71) yields

$$Q \geq \frac{n}{C_1 \varepsilon^2} \log \frac{1}{4\delta}.$$

For $\delta \leq 1/8$, this is (69) after changing the universal constant. □

Theorem 16 and Theorem 17 together prove the minimax theorem in the main paper.

H A Bounded-Menu Confidence Lower Bound

The same family quantifies how menu size can reduce the confidence obstruction. Suppose every allowed conditional query has cardinality at most k, where $2 \leq k \leq n$.

Proposition 18 (Bounded-menu lower bound)**.** *Under the parameter ranges of Theorem 17, every learner from menus of size at most k requires*

$$Q = \Omega\left(\frac{n}{k} \frac{\log(1/\delta)}{\varepsilon^2}\right). \tag{72}$$

Proof. Use the two distributions in (62). A menu excluding label 1 is uniform under both hypotheses and carries no information. If a menu contains label 1 and $s \leq k-1$ light labels, the identity of a light winner is uniform conditional on a light outcome. The only informative statistic is therefore Bernoulli with light-outcome probability

$$r_s(a) = \frac{s(1-a)}{(n-1)a + s(1-a)}. \tag{73}$$

Uniformly for $a \in [1/4, 3/4]$, $r_s(a) = \Theta(s/n)$ and

$$|r_s'(a)| = \frac{s(n-1)}{((n-1)a + s(1-a))^2} = O(s/n). \tag{74}$$

Equation (66) now bounds the one-query KL by $Ck\varepsilon^2/n$. The adaptive transcript and testing argument from Theorem 17 gives (72). □

Ordinary distribution-learning hardness also gives $\Omega(n/\varepsilon^2)$ even when the full domain is queryable (Canonne, 2020). Hence every bounded-menu learner has lower bound

$$\Omega\left(\frac{n + (n/k)\log(1/\delta)}{\varepsilon^2}\right).$$

We do not claim a matching upper bound when $k > 2$.

I Numerical Verification

The proofs are analytic. The script `verify_trajectory_theorem.py` performs independent floating-point checks designed to catch indexing, scaling, and inequality-direction errors. It does not supply evidence needed by any theorem.

Across 1,014 seeded mass profiles in dimensions 2 through 40, including uniform, one-heavy, geometric, and Dirichlet profiles, the script checked:

1. the winner kernel's symmetrized eigenvalues lie in $[0, 1]$,
2. the Helmert formula (43) agrees with a direct pseudoinverse of L_0,
3. the exact Green function is bounded by twice the comparison Green function coordinatewise,
4. each of the left, middle, and right profile sums obeys its analytic bound,
5. the exact finite-T stationary covariance for $T \in \{1, 2, 5, 20, 100\}$ obeys Lemma 9, and
6. the normalized one-query KL in the hard family stays below the loose analytic constant over $4 \leq n \leq 256$ and 80 logarithmically spaced accuracies.

The largest observed normalized KL, $nD_{\mathrm{KL}}/\varepsilon^2$, was 161.847400. The terminal output was `all trajectory-theorem checks passed`.

J Additional Related Work

The closest result is Kleinberg et al. (2026). Its generic complete-family learner runs a global winner chain, burns in, keeps one terminal state, and independently restarts. Its hierarchical learner keeps local trajectories but estimates one binary tree split at a time. Its pairwise corollary is $O(n \log^2 n \log(n/\delta)/\varepsilon^2)$, and its conclusion explicitly leaves the polylogarithmic gaps open. It contains neither the empirical-law risk theorem nor the multiplicative-confidence lower bound.

Accelerated spectral ranking (Agarwal et al., 2018), graph-resistance estimation (Hendrickx et al., 2019), and BTL minimax analyses (Shah et al., 2015; Hendrickx et al., 2020) study fixed comparison designs and parameter losses. Their uniform guarantees assume positive weights and bounded parameter ranges or enter a many-comparisons-per-edge regime. None analyzes the query-driven winner trajectory used here.

The exact-sampling method of Fotakis et al. (2022) uses coupling from the past for an exogenous pair design. Its distribution-free procedure first estimates all weights to relative accuracy under a learnability condition, then modifies the chain and performs exact correction. Its goal and complexity are distinct from minimax labeled total-variation learning.

Ordinary distribution-learning lower bounds (Canonne, 2020) give the additive confidence benchmark. Generic spectral-gap concentration (Jiang et al., 2018) incurs the relaxation-time factor avoided by the capped-mass profile. Classical biased-sampling results (Vardi, 1985; Gill et al., 1988) are asymptotic and fixed-design.

References

Agarwal, A., Patil, P., and Agarwal, S. (2018). Accelerated spectral ranking. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 70–79.

Bradley, R. A. and Terry, M. E. (1952). Rank analysis of incomplete block designs: I. the method of paired comparisons. *Biometrika*, 39(3/4):324–345.

Canonne, C. L. (2020). A short note on learning discrete distributions. *arXiv preprint arXiv:2002.11457*.

Canonne, C. L., Ron, D., and Servedio, R. A. (2015). Testing probability distributions using conditional samples. *SIAM Journal on Computing*, 44(3):540–616.

Chakraborty, S., Fischer, E., Goldhirsh, Y., and Matsliah, A. (2016). On the power of conditional samples in distribution testing. *SIAM Journal on Computing*, 45(4):1261–1296.

Fotakis, D., Kalavasis, A., and Tzamos, C. (2022). Perfect sampling from pairwise comparisons. In *Advances in Neural Information Processing Systems*, volume 35.

Gill, R. D., Vardi, Y., and Wellner, J. A. (1988). Large sample theory of empirical distributions in biased sampling models. *The Annals of Statistics*, 16(3):1069–1112.

Hendrickx, J., Olshevsky, A., and Saligrama, V. (2020). Minimax rate for learning from pairwise comparisons in the BTL model. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 4193–4202.

Hendrickx, J. M., Olshevsky, A., and Saligrama, V. (2019). Graph resistance and learning from pairwise comparisons. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 2702–2711.

Jiang, B., Sun, Q., and Fan, J. (2018). Bernstein's inequalities for general markov chains. *arXiv preprint arXiv:1805.10721*.

Kleinberg, J., Saberi, A., Tan, X., and Velegkas, G. (2026). Learning distributions from multiple data providers. *arXiv preprint arXiv:2607.24732*.

Luce, R. D. (1959). *Individual Choice Behavior: A Theoretical Analysis*. Wiley.

Negahban, S., Oh, S., and Shah, D. (2017). Rank centrality: Ranking from pairwise comparisons. *Operations Research*, 65(1):266–287.

Shah, N. B., Balakrishnan, S., Bradley, J., Parekh, A., Ramchandran, K., and Wainwright, M. J. (2015). Estimation from pairwise comparisons: Sharp minimax bounds with topology dependence. In *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics*, volume 38 of *Proceedings of Machine Learning Research*, pages 856–865.

Tsybakov, A. B. (2009). *Introduction to Nonparametric Estimation*. Springer.

Vardi, Y. (1985). Empirical distributions in selection bias models. *The Annals of Statistics*, 13(1):178–203.