Mallows Model with Learned Distance Metrics: Sampling and Maximum Likelihood Estimation

Yeganeh Alimohammadi
University of Southern California
yalimoha@usc.edu

Kiana Asgari
Stanford University
asgkiana@stanford.edu

Abstract

Mallows model is a widely-used probabilistic framework for learning from ranking data, with applications ranging from recommendation systems and voting to aligning language models with human preferences [9, 22, 36]. Under this model, observed rankings are noisy perturbations of a central ranking σ, with likelihood decaying exponentially in distance from σ, i.e, $\mathbb{P}(\pi) \propto \exp\left(-\beta \cdot d(\pi, \sigma)\right)$, where $\beta > 0$ controls dispersion and d is a distance function.

Existing methods mainly focus on fixed distances (such as Kendall's τ distance), with no principled approach to learning the distance metric directly from data. In practice, however, rankings naturally vary by context; for instance, in some sports we regularly see long-range swaps (a low-rank team beating a high-rank one), while in others such events are rare. Motivated by this, we propose a generalization of Mallows model that learns the distance metric directly from data. Specifically, we focus on L_α distances: $d_\alpha(\pi, \sigma) := \sum_{i=1} |\pi(i) - \sigma(i)|^\alpha$.

For any $\alpha \geq 1$ and $\beta > 0$, we develop a Fully Polynomial-Time Approximation Scheme (FPTAS) to efficiently generate samples that are ϵ- close (in total variation distance) to the true distribution. Even in the special cases of L_1 and L_2, this generalizes prior results that required vanishing dispersion ($\beta \to 0$). Using this sampling algorithm, we propose an efficient Maximum Likelihood Estimation (MLE) algorithm that jointly estimates the central ranking, the dispersion parameter, and the optimal distance metric. We prove strong consistency results for our estimators, and we validate our approach empirically using datasets from sports rankings.

1 Introduction

Ranking tasks arise in many applications, from online recommendation systems, sports competitions, and voting, to more recently, the alignment of Large Language Models (LLMs) with human preferences [9, 22, 36]. A fundamental challenge is common to all these scenarios: given multiple observed rankings, how do we aggregate them into a single central ranking, *learn* how rankings vary around this center, and *generate* new rankings reflecting observed patterns?

Probabilistic models offer natural tools to address these questions. Among them, the *Mallows model*[29] stands out due to its conceptual simplicity and interpretability. Under this model, observed rankings are viewed as noisy perturbations of a central permutation[1] σ, with their likelihood decaying exponentially according to a chosen distance function d:

$$\mathbb{P}(\pi) := \frac{\exp\left(-\beta \, d(\pi, \sigma)\right)}{Z},$$

---

[1]Throughout this paper, we use the terms *ranking* and *permutation* interchangeably.

where $\beta > 0$ controls the dispersion around σ, and Z is the normalizing constant (often referred to as the partition function). Due to its interpretability and desirable structural properties, Mallows model has found widespread adoption in recommendation systems, voting predictions, retail assortment optimization, and, more recently, post-training of large language models [9, 10, 11, 23].

Initially, Mallows [29] proposed this model specifically with Kendall's tau distance, and it was subsequently generalized by Diaconis [12] to any right-invariant distance metric. In this paper, we focus on the L_α *Mallows model*, defined by the family of L_α distances: $d_\alpha(\pi, \sigma) := \sum_{i=1}^{n} |\pi(i) - \sigma(i)|^\alpha$. Intuitively, the parameter α controls the penalty for *long-range* swaps (e.g., changes from rank 1 to rank 100) versus *local* swaps, with larger values of α penalizing long-range swaps more strongly.

Our contributions focus on two complementary aspects:
Parameter Estimation. We present an efficient Maximum Likelihood Estimation (MLE) algorithm that jointly estimates the central ranking σ, the dispersion parameter β, and the distance parameter α. Unlike previous studies on the Mallows model –which assume a fixed distance metric– our estimator learns the distance metric jointly with other model parameters, marking the first such result in the literature. Further, we prove strong consistency of our estimators under general conditions $\alpha > 0, \beta > 0$ (see Theorem 2.4).
Efficient Sampling. We develop a Fully Polynomial-Time Approximation Scheme (FPTAS) for sampling permutations within ϵ-total variation distance of the true L_α-Mallows distribution, valid for any $\alpha \geq 1$ and $\beta > 0$ (Theorem 3.1). Our result generalizes previous sampling methods limited to special cases (L_1, L_2) that required vanishing dispersion ($\beta \to 0$) [50, 32].

We validate our algorithms empirically using sports ranking datasets from American college football and basketball (Section 4, codebase [2]). Our results reveal substantial differences between these two sports: basketball rankings yield a larger $\alpha \approx 1.09$, reflecting a rapid decay in the probability of large inversions (e.g., low-ranked teams beating top-ranked teams). In contrast, football rankings yield a smaller $\alpha \approx 0.44$, indicating a relatively higher likelihood of large inversions. Such domain-specific variations highlight the importance of learning distance metrics from data rather than fixing them a priori.

Related Work Prior works on sampling and learning Mallows model assume a fixed distance function, with Kendall's tau distance (counting inversions between permutations) [21] being a particularly common choice due to its closed-form expression [13, 16, 18, 26, 40]. However, it implicitly assumes equal probabilities for all inversions. For example, in a web search ranking, swapping the top two results has far more impact on user experience than swapping the 91st and 92nd, yet under Kendall's tau distance the inversions appear with equal probabilities. This uniform penalty is misaligned with real-world tasks where the location of ranking errors matters—such as in recommendation systems, sports ranking, and preference learning tasks.

Alternative metrics like Spearman's footrule (L_1) and rank correlation (L_2) or more generally the L_α model used in this paper, aim to mitigate the limitations of Kendall's tau by assigning smaller penalties to long-range inversions. Exact sampling under these metrics is known to be NP-hard [4], leading to the development of approximate methods such as Markov chain Monte Carlo (MCMC) algorithms [50, 47]. While these methods represent significant progress, prior to this work, their computational efficiency was established only under the assumption of vanishing dispersion ($\beta \to 0$).

To address the limitations of Kendall's tau in another direction, generalizations of the Mallows model have been proposed [7, 24, 31], introducing inversion penalties through graphical models or mixtures of multiple Mallows distributions. While these methods increase expressiveness, they often do so at the cost of substantial model complexity and reduced interpretability. In contrast, the L_α-Mallows framework we study offers a principled and tunable way to learn the shape of the distribution through a single metric parameter. Although such extensions are beyond the scope of this work, the L_α-Mallows model could similarly be adapted to mixture or hierarchical frameworks to further increase modeling flexibility.

From the perspective of estimation, extensive research has examined consistency and convergence of Mallows model parameters [47, 3, 15, 30, 8]. In cases where the partition function is intractable (e.g., for L_1 and L_2), and the maximum likelihood estimators cannot be computed in closed form, estimation methods often rely on Bayesian approaches [47] or Monte Carlo Maximum Likelihood Estimation (MCMLE) [17], which approximates the likelihood using sampling. Among more recent

works, Tang [45] analyzed MLE under fixed metrics like Kendall's tau and L_2, identifying the central ranking and potential upward biased estimator for the dispersion parameter. Mukherjee and Tagami [32, 33] provided consistency results for generalized Mallows-type exponential families, though assuming a known central ranking. Extending this line of research, our work jointly estimates all model parameters, and establishes strong consistency guarantees for these estimators.

Alternative probabilistic frameworks include pairwise-comparison models such as Bradley–Terry [6] and its generalization Plackett–Luce [27, 35], which assign latent scores to items and generate rankings sequentially. Their computational simplicity makes them popular in practice [37, 48], but unlike the Mallows model, the number of parameters grows linearly with the number of items. We include the Plackett–Luce model as a high-parameter benchmark in our empirical evaluation and consistently observe improved performance from our approach (see Section 4).

2 Parameter Estimation via Maximum Likelihood

Our first goal is to estimate the parameters of the L_α-Mallow's model, including the central ranking σ_0, the dispersion parameter β_0, and the distance function α_0. Suppose we have observed m rankings $\{\pi^{(1)}, \ldots, \pi^{(m)}\}$ i each independently drawn from a Mallows distribution with unknown parameters $(\alpha_0, \beta_0, \sigma_0)$. To estimate these parameters, we use Maximum Likelihood Estimation (MLE), choosing values that maximize the average log-likelihood of the observed data:

$$\mathcal{L}_m(\alpha, \beta, \sigma) = \frac{1}{m} \sum_{l=1}^m \log \mathbb{P}_{\alpha,\beta,\sigma}(\pi^{(l)}) = -\frac{\beta}{m} \sum_{l=1}^m d_\alpha(\pi^{(l)}, \sigma) - \log Z_n(\beta, \alpha).$$

where the probability of observing a ranking π is given by

$$\mathbb{P}_{\alpha,\beta,\sigma}(\pi) = \frac{e^{-\beta d_\alpha(\pi,\sigma)}}{Z_n(\beta, \alpha)}, \quad \text{with} \quad Z_n(\beta, \alpha) = \sum_{\pi \in S_n} e^{-\beta d_\alpha(\pi,\sigma)},$$

and S_n is the set of all permutations over n items.
With these definitions in place, we now describe a two-step estimation of the MLE (Algorithm 1).

2.1 Step 1: Estimating the central ranking

First, observe that the partition function $Z_n(\beta, \alpha)$ is invariant to the choice of the central ranking σ (Proposition A.1 in Appendix A). Thus, minimizing the negative log-likelihood over σ effectively reduces to solving a simpler optimization problem: $\hat{\sigma}_m \in \arg\min_{\sigma \in S_n} \sum_{l=1}^m d_\alpha(\pi^{(l)}, \sigma)$.

Since the value of α_0 is unknown, we instead solve for L_1 distance:

$$\hat{\sigma}_m \in \arg\min_\sigma \sum_{l=1}^m d_1(\pi^{(l)}, \sigma). \tag{1}$$

This simplification is theoretically justified by the following lemma, which guarantees consistency of the estimated central ranking regardless of the choice of the distance function.

Lemma 2.1. *Let $\{\pi^{(1)}, \ldots, \pi^{(m)}\}$ be m observed rankings, which are assumed to be sampled i.i.d from Mallow's model with parameters $(\alpha_0, \beta_0, \sigma_0)$. Fix any $\tilde{\alpha} > 0$, and let $\hat{\sigma}_m \in \arg\min_{\sigma \in S_n} \sum_{l=1}^m d_{\tilde{\alpha}}(\pi^{(l)}, \sigma)$. Then as $m \to \infty$ the central ranking estimator converges almost surely $\hat{\sigma}_m \xrightarrow{\mathbb{P}} \sigma_0$. Further, σ_0 is the unique minimizer of the expected distance, i.e.,*

$$\{\sigma_0\} = \arg\min_{\tilde{\sigma} \in S_n} \left(\mathbb{E}[d_{\tilde{\alpha}}(\Pi, \tilde{\sigma})] \right),$$

where Π is a random permutation sampled from $\mathbb{P}_{\alpha_0,\beta_0,\sigma_0}$.

To see why Lemma 2.1 holds, consider any ranking that differs from σ_0. Such a ranking contains at least one inversion relative to σ_0, and we show that correcting this inversion strictly reduces the expected distance. Repeatedly applying this procedure yields a strictly decreasing sequence of expected distances, ensuring that the true central ranking σ_0 uniquely emerges as the global minimizer. A complete proof appears in Appendix B.

3

Crucially, the simplified optimization in (1) is computationally efficient. It corresponds exactly to a *minimum-weight perfect matching* problem on a bipartite graph where the weight of an edge between node i on the first side and node j on the second side is given by: $\sum_{l=1}^{m} |\pi^{(l)}(i) - j|$. This matching problem can be solved efficiently in near-linear time [14], giving the central ranking estimator:

Proposition 2.2. *The minimizer $\hat{\sigma}_m$ can be obtained by solving a minimum-weight perfect matching problem in a bipartite graph, with each side of size n, where the edge weight between node i on one side and node j on the other side is:* $\sum_{l=1}^{m} |\pi^{(l)}(i) - j|$.

2.2 Step 2: Estimating the distance and dispersion parameters

With the central ranking $\hat{\sigma}_m$ fixed, we now estimate the continuous parameters (α, β). We formulate this as a two-dimensional root-finding problem, where we seek parameter values that align empirical and model-implied distances. Specifically,

$$\Psi_m(\alpha, \beta; \hat{\sigma}_m) := \left(-\frac{1}{m} \sum_{i=1}^{m} d_\alpha(\pi^{(i)}, \hat{\sigma}_m) + \mathbb{E}[d_\alpha(\Pi, \hat{\sigma}_m)], \quad -\frac{1}{m} \sum_{i=1}^{m} \dot{d}_\alpha(\pi^{(i)}, \hat{\sigma}_m) + \mathbb{E}[\dot{d}_\alpha(\Pi, \hat{\sigma}_m)] \right),$$

where Π is a random permutation with distribution $\mathbb{P}_{\alpha, \beta, \hat{\sigma}_m}$, and $\dot{d}_\alpha(\pi, \sigma) := \sum_{i=1}^{n} |\pi(i) - \sigma(i)|^\alpha \log |\pi(i) - \sigma(i)|$ denotes the partial derivative of d_α with respect to α.

The function Ψ_m corresponds to the gradient of the log-likelihood, with terms rescaled by β. We estimate (α_0, β_0) by solving:

$$(\hat{\alpha}_m, \hat{\beta}_m) \in \{(\alpha > 0, \beta > 0) : \ \Psi_m(\alpha, \beta; \hat{\sigma}_m) = (0, 0)\}.$$

In practice, a simple zero-order optimization method, such as the differential evolution algorithm [44], can be used to solve this system efficiently. Note that each iteration of this method requires evaluating the partition function and its derivatives, since we have

$$\mathbb{E}[d_\alpha(\Pi, \hat{\sigma}_m)] = -\frac{\partial \log Z_n(\alpha, \beta)}{\partial \beta}, \quad \beta \cdot \mathbb{E}[\dot{d}_\alpha(\Pi, \hat{\sigma}_m)] = -\frac{\partial \log Z_n(\alpha, \beta)}{\partial \alpha}.$$

For large n, we can use a sampling algorithm to approximate these expectations. In Section 3, we develop a sampling algorithm which yields provable error guarantees for all $\alpha \geq 1$.

Algorithm 1: MLE for $\{\alpha, \beta, \sigma\}$ in the Mallows Model

Input : Samples $\{\pi^{(1)}, \ldots, \pi^{(m)}\}$;

Step 1: Solve for $\hat{\sigma}_m$ using minimum matching problem. $\hat{\sigma}_m := \arg\min_\sigma \sum_{l=1}^{m} d_1(\pi^{(l)}, \sigma)$.

Step 2: Estimate parameters $(\hat{\alpha}_m, \hat{\beta}_m)$ by solving the system:

$$(\hat{\alpha}_m, \hat{\beta}_m) \leftarrow \text{Solve} \left\{ (\alpha, \beta) : \Psi_m(\alpha, \beta; \hat{\sigma}_m) = (0, 0) \right\}.$$

return $\hat{\alpha}_m, \hat{\beta}_m, \hat{\sigma}_m$.

Remark 2.3. *While our estimation procedure works for any $\alpha > 0$, the sampling algorithm developed in Section 3.3 requires $\alpha \geq 1$ to guarantee provable efficiency and approximation accuracy when evaluating the function Ψ_m. This constraint is purely computational and does not impact the statistical properties of the estimator or its empirical performance. The sampler simply enables efficient evaluation of Ψ_m, making root-finding for parameter estimation computationally tractable.*

2.3 Statistical Guarantees

Under standard assumptions of i.i.d. samples from the Mallows distribution, the maximum likelihood estimator proposed in Algorithm 1 jointly estimates the dispersion parameters and central ranking. The result below establishes its statistical guarantees when the score function Ψ_m is computed exactly: the estimator is consistent, and the continuous parameters converge at the optimal $\sqrt{m}$ rate with an asymptotically normal distribution.

Theorem 2.4 (MLE Consistency and Asymptotic Normality). *Suppose the rankings $\pi^{(1)}, \ldots, \pi^{(m)}$ are drawn i.i.d. from the Mallows distribution $\mathbb{P}_{\alpha_0, \beta_0, \sigma_0}$ with true parameters $\alpha_0, \beta_0, \sigma_0$. Let $(\hat{\alpha}_m, \hat{\beta}_m, \hat{\sigma}_m)$ be the estimators returned by Algorithm 1. Then, as $m \to \infty$, the following hold:*

1. *Consistency of $\hat{\sigma}_m$: The estimated central ranking $\hat{\sigma}_m$ converges almost surely to the true central ranking σ_0. Moreover, the probability of error admits the finite-sample bound*

$$\mathbb{P}\big(\hat{\sigma}_m \neq \sigma_0\big) \;\leq\; n! \cdot \exp\left(-\frac{m}{3n^4}\right).$$

2. *Consistency and asymptotic normality of $(\hat{\alpha}_m, \hat{\beta}_m)$: The estimators $(\hat{\alpha}_m, \hat{\beta}_m)$ converge in probability to (α_0, β_0), and further is an efficient estimator; namely,*

$$\sqrt{m} \begin{pmatrix} \hat{\alpha}_m - \alpha_0 \\ \hat{\beta}_m - \beta_0 \end{pmatrix} \xrightarrow{d} \mathcal{N}\left(\mathbf{0},\, \mathcal{I}_{\alpha_0, \beta_0}^{-1}\right),$$

where $\mathcal{I}_{\alpha_0, \beta_0}$ is the Fisher information matrix, given by

$$\begin{pmatrix} \beta^2 \operatorname{Var}[\dot{d}_\alpha(\Pi, id)] & \beta \operatorname{Cov}[\dot{d}_\alpha(\Pi, id),\, d_\alpha(\Pi, id)] \\ \beta \operatorname{Cov}[\dot{d}_\alpha(\Pi, id),\, d_\alpha(\Pi, id)] & \operatorname{Var}[d_\alpha(\Pi, id)] \end{pmatrix},$$

where and id is the identity permutation and Π is a random permutation with distribution $\mathbb{P}_{\alpha_0, \beta_0, id}$.

The main challenge for establishing consistency results on the parameters α and β, is that the log-likelihood is not jointly convex in both parameters. However, by carefully analyzing the structure of Mallow's distribution, we show that the log-likelihood gradient admits a *unique*, stable minimizer. For this purpose, we show that the function $\Psi(\alpha, \beta)$ is identifiable, by showing it is locally invertible (Lemma B.2), and proper (Lemma B.3). These conditions, along with uniform convergence of Ψ_m to Ψ, allow us to apply Theorem 5.9 of [46] to conclude consistency and asymptotic normality. Full technical details appear in Appendix B.

Remark 2.5. *The convergence rate in part* (i) *of Theorem 2.4 depends on the choice of the distance exponent $\tilde{\alpha}$ used in the first step of the algorithm. A direct optimization of the error bound shows that setting $\tilde{\alpha} = n/3$ minimizes the upper bound on the probability of misidentifying the true ranking, yielding the improved rate*

$$\mathbb{P}\left(\hat{\sigma}_m \neq \sigma_0\right) \leq n! \cdot \exp\left(-\frac{m}{54n^2}\right).$$

However, for ease of implementation and interpretation, we set $\tilde{\alpha} = 1$ in all theoretical and empirical results.

Approximate computation of Ψ_m. When exact evaluation of the gradient Ψ_m is not possible –typically due to the intractability of the partition function $\ln Z(\alpha, \beta)$ and its derivatives– one can work with the approximations of Ψ_m. We show that the same statistical guarantees continue to hold, provided the approximation error vanishes faster than $m^{-1/2}$. The next result formalizes this statement. Proof appear in Appendix B.

Theorem 2.6 (Asymptotic Normality under Approximate Score). *Suppose the rankings $\pi^{(1)}, \ldots, \pi^{(m)}$ are drawn i.i.d. from the Mallows distribution $\mathbb{P}_{\alpha_0, \beta_0, \sigma_0}$ with true parameters $\alpha_0, \beta_0, \sigma_0$, where (α_0, β_0) lies in a compact subset of $\Theta \in (0, \infty)^2$. Let $(\tilde{\alpha}_m, \tilde{\beta}_m)$ be approximate estimators that solve the estimating equation*

$$\widetilde{\Psi}_m(\tilde{\alpha}_m, \tilde{\beta}_m) := \Psi_m(\tilde{\alpha}_m, \tilde{\beta}_m) + \Delta(\tilde{\alpha}_m, \tilde{\beta}_m) = 0,$$

where the approximation error $\Delta(\alpha, \beta)$ satisfies $\sup_{(\alpha, \beta) \in \Theta} \|\Delta(\alpha, \beta)\| = o(m^{-1/2})$. Then, the conclusions of Theorem 2.4 continue to hold for $(\tilde{\alpha}_m, \tilde{\beta}_m)$, i.e.,

$$\sqrt{m} \begin{pmatrix} \tilde{\alpha}_m - \alpha_0 \\ \tilde{\beta}_m - \beta_0 \end{pmatrix} \xrightarrow{d} \mathcal{N}\left(\mathbf{0},\, \mathcal{I}_{\alpha_0, \beta_0}^{-1}\right).$$

5

3 Efficient Sampling

Sampling rankings play a central role in many applications. In sports analytics, it enables the simulation of tournament outcomes; in recommendation systems, it supports generating user suggestions. Further, sampling is needed for computing the normalizing constant (also known as the partition function) required for maximum likelihood estimation of the Mallows model.

Exact computation of this partition function becomes infeasible as n grows—indeed, it is equivalent to Kemeny's consensus ranking problem, which is known to be NP-hard [43]. Although Markov chain Monte Carlo (MCMC) methods have been developed to approximate sampling [19, 5] (often phrased in terms of approximating the permanent of a matrix), the MCMC methods rarely scale well in practice [34, 42].

Instead of relying on these general-purpose algorithms, we exploit the structure of the L_α-Mallow's model to efficiently sample rankings. Our main result on sampling is summarized next:

Theorem 3.1. *For any $\alpha \geq 1$ and $\beta > 0$, and any desired accuracy $\epsilon > 0$, there exists a fully polynomial-time approximation scheme (FPTAS) that, for the Mallow's distribution, (i) estimates the partition function $Z_n(\beta, \alpha)$ within a multiplicative factor of $1 - \epsilon$, and (ii) generates samples ϵ-close (in total variation distance) to the true distribution.*

Proof Outline: Throughout the proof, we assume without loss of generality that the central ranking is identity (i.e., $\sigma(i) = i$). Our starting point is the construction of a matrix $\mathbf{A}_n$ whose permanent equals $Z_n(\beta, \alpha)$. Specifically, let $\mathbf{A}_n[i, j] = e^{-\beta|i-j|^\alpha}$, for $i, j \in [n]$. Note that sampling a permutation according to the probabilities defined by $\mathbf{A}_n$ corresponds to sampling from the L_α-Mallow's with dispersion β. For instance, assigning items along the diagonal corresponds to the central ranking.

A crucial step in our proof (Lemma 3.2) shows that the probability of assigning any item to a position far from its location in the central ranking decays exponentially. Using this property, we approximate $\mathbf{A}_n$ by truncation – setting entries sufficiently far from the diagonal to zero. This truncation approach, explained in detail in Sections 3.2, yields a probabilistic ranking model within ϵ-total variation distance to the true distribution. The resulting truncated structure then enables the use of an efficient dynamic programming (DP) algorithm for sequential sampling, which we present in Section 3.3.

3.1 Marginal Decay

We begin by showing that marginal probabilities in the L_α Mallow's distribution exhibits exponential decay away from the diagonal.

Lemma 3.2. *Let $\alpha \geq 1$, $\beta > 0$. There exist constants $k > 0$ and $c(\alpha, \beta) < 1$, independent of n, such that for all $i, j \in [n]$ with $|j - i| \geq k$,*

$$\mathbb{P}_{\alpha,\beta}\left(\pi(i) = j\right) \leq c(\alpha, \beta)^{|j-i|}.$$

Furthermore, $c(\alpha, \beta)$ is monotone decreasing in α and β, with $c(1, \beta) \leq \frac{2e^{-2\beta}}{1+e^{-2\beta}}$.

This is the key lemma used to design the sampling algorithm. The proof carefully analyzes marginal probability ratios $\mathbb{P}_{\alpha,\beta}(\pi(i) = j + 1)/\mathbb{P}_{\alpha,\beta}(\pi(i) = j)$ by comparing permutations assigning i to neighboring positions j vs. $j + 1$. Specifically, we define two types of mappings: The first involves simple transpositions that often directly increase permutation probabilities by resolving inversions. When such direct mappings fail to increase probability, we use a second, more complicated mapping that ensures the resulting permutation moves closer to the identity and thus achieves higher probability. Combining these two mapping types establishes geometric decay in probabilities with respect to $|j - i|$. Full proof appears in Appendix C.1.

3.2 Truncated Distribution

We approximate $\mathbb{P}_{\alpha,\beta}$ by truncating its distribution. Define the *truncated matrix* $\mathbf{A}_n^{(k)}$ by zeroing out entries of $\mathbf{A}_n$ that lie far from the main diagonal: $\mathbf{A}_n^{(k)}[i, j] = e^{-\beta|i-j|^\alpha} \mathbf{1}_{\{|i-j| \leq k\}}$. Let $\mathbb{P}_{\alpha,\beta}^{(k)}$ denote the probability distribution induced by this truncated matrix. We show that for $k = O(\log n)$ the truncated distribution $\mathbb{P}_{\alpha,\beta}^{(k)}$ closely approximates the original Mallows distribution.
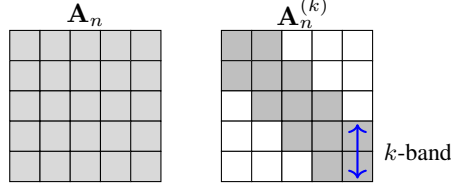
Lemma 3.3. *Given $\epsilon > 0$, $\alpha \geq 1$, and $\beta > 0$, let $k = \log(n/\epsilon)\Big(\log\big(\frac{1+e^{-2\beta}}{2e^{-2\beta}}\big)\Big)^{-1}$, then*

1. $\|\mathbb{P}_{\alpha,\beta} - \mathbb{P}_{\alpha,\beta}^{(k)}\|_{TV} \leq \epsilon$, where TV denotes the total variation distance.

2. $per(\mathbf{A}_n^{(k)})$ is a multiplicative approximation of $\mathbf{Z}_n(\alpha,\beta)$, i.e., $\frac{|per(\mathbf{A}_n) - per(\mathbf{A}_n^{(k)})|}{per(\mathbf{A}_n)} \leq \epsilon$.

To prove this result, we use Lemma 3.2 which shows that the probability of a permutation containing at least one element displaced by more than distance k decays exponentially with k (Lemma 3.2). Therefore, if we choose $k = O(\log n)$, a union bound over all indices yields that the total variation distance between $\mathbb{P}_{\alpha,\beta}^{(k)}$ and $\mathbb{P}_{\alpha,\beta}$ is sufficiently small. Full proof appears in Appendix C.2.

3.3 Sampling from Truncated Distribution

The final step is to efficiently sample a permutation π from the (truncated) distribution $\mathbb{P}_{\alpha,\beta}^{(k)}(\pi)$. We achieve this via dynamic programming (DP), structured in layers that assign elements to positions. Specifically, a DP state $DP[i][s]$ tracks the cumulative weight of all partial permutations that have assigned the first $i - 1$ elements, where the state s indicates available columns. Recall that each element i sampled from $\mathbb{P}_{\alpha,\beta}^{(k)}(\pi)$ can be only matched to those $\pm k$ around i. Thus, the state s is a binary vector of length $2k$, marking columns as available (0) or assigned (1). Transitions between layers assign the current element i to an available column, updating the state to mark this column as assigned, and weighting the transition by the corresponding entry from $\mathbf{A}_n^{(k)}$ (see Figure 1). The final DP state at layer captures the total weight of all perfect matchings, equal to the permanent of $\mathbf{A}_n^{(k)}$.

Once the forward DP pass is complete, sampling can be done by reversing through DP states, choosing transitions with probability proportional to their DP-computed weights (marginal probabilities). Detailed algorithms and rigorous analysis for DP updates, graph construction, and sampling appear in Appendices C.3 and C.4 (Algorithms 2 and 3).

Lemma 3.4. *For any given n and k: Algorithm 2 constructs the DP table and $per(\mathbf{A}_n^{(k)})$ in time $O\left(n\binom{2k}{k}\right)$. Then given the precomputed DP table, Algorithm 3 samples each permutation from $\hat{\mu}_n^{(k)}$ in time $O(nk)$.*

Combining Lemma 3.4 with the truncation guarantee in Section 3.2, we obtain the PTAS described in Theorem 3.1. Specifically, for $k = O(\log(n/\epsilon))$ the total variation distance between the true Mallow's distribution and the truncated model is at most ϵ. Under this setting, the preprocessing step takes time $O(\epsilon^{-C} n^{1+C})$, where $C = 2\log((1 + e^{-2\beta})/2e^{-2\beta})$ (constant derived from Lemma 3.3). After this setup, each sample can be drawn in $O(n \log(n/\epsilon))$ time. Together, these results yield a PTAS for both partition function estimation and sampling.

4 Experiments

We empirically evaluate our methods with two primary goals: (1) assessing the predictive performance of our model compared to classical ranking models on real-world and synthetic datasets; and (2) validating our approximation scheme for the partition function $Z_n(\beta, \alpha)$ and the sampling algorithm.

Reproducibility and Computational Resources. Our implementation, which also includes tools for sampling and learning from L_α Mallows models, is available at [2]. The MLE evaluation experiments were conducted on a server with dual AMD EPYC 7763 64-Core processors and 1.5 TB of RAM. Sampling experiments were run on an Intel Core i7 CPU with 16 GB RAM.
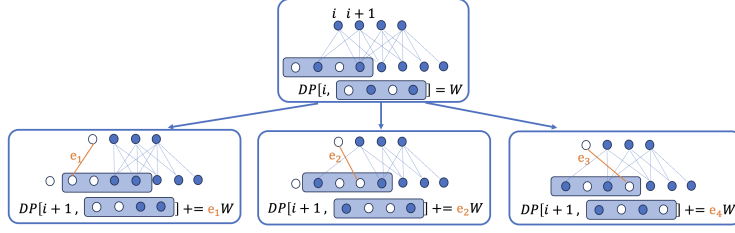
Figure 1: DP state transitions showing available (white) and assigned (blue) columns within bandwidth ($k = 2$). Matched edges for i are shown in orange; DP values represent weights of partial matchings.

4.1 Predictive Performance of the L_α-Mallow's Model

Datasets. We evaluate our methods on the following datasets:

American college football rankings: We use weekly power rankings of college football teams from the 2019 to 2021 seasons, sourced from [38], providing over 4,100 distinct observed rankings across various expert sources. These rankings—known as *power rankings*—are ordinal lists produced by independent media outlets and analysts. Each ranking reflects an expert's assessment of team strength during a particular week, typically incorporating factors such as win–loss records, strength of schedule, point differentials, and recent performance. Unlike standings, which are derived purely from match outcomes, power rankings aim to capture an underlying strength signal. We focus on learning the top 10 teams identified in [39]: Georgia, Alabama, Michigan, Cincinnati, Baylor, Ohio State, Oklahoma State, Notre Dame, Michigan State, and Oklahoma. To evaluate scalability, we also conduct experiments on a larger set of the top 100 most frequently ranked teams.

College basketball rankings: We similarly use weekly power rankings from the 2020 and 2021 college basketball seasons, also sourced from [38], resulting in over 1,700 observed rankings. We further, select the top 10 teams as reported in [1]: Gonzaga, Illinois, Baylor, Michigan, Alabama, Houston, Ohio State, Iowa, Texas, and Arkansas. For large-scale evaluation, we additionally sample a set of 100 teams at random.

SUSHI preference rankings: We also include the SUSHI preference dataset from [20], which contains 5,000 human-annotated rankings over 10 different types of SUSHI.

Synthetic data: We generate rankings using our sampler (Algorithm 3). The parameters are set to $\alpha = 1.5$, $\beta = 0.5$, with $n = 15$ items, and the truncation of $k = 9$ is used for sample generation.

Data Splitting Strategy. To avoid temporal leakage in the sports datasets, we train on earlier rankings and hold out the final sixth of the data for evaluation. We repeat this sampling procedure independently over 50 trials to derive confidence intervals. In each of 50 trials, we randomly sample 700 training and 150 testing examples for basketball, and 800 training and 250 testing examples for football. For the SUSHI dataset, we conduct 25 independent trials, each with a random 80%- 20% training-testing split. We report mean and standard deviation across trials.

Baselines. To highlight the performance of our model, we compare against two classical models:

Plackett–Luce (PL) model [35]: The PL model is a classical ranking model that assumes rankings are generated sequentially. At each stage, the probability of choosing item i for the next available position is proportional to a parameter θ_i. Once selected, the item is removed from the pool, and the process continues with the remaining items. Formally,

$$\mathbb{P}_\theta^{\text{PL}}(\pi) = \frac{\prod_{i=1}^{n} \theta_{\pi(i)}}{\prod_{i=1}^{n} \left(\sum_{j=i}^{n} \theta_{\pi(j)} \right)}.$$

Mallow's τ (Kendall's tau distance): Mallow's τ assumes the underlying distance metric is the Kendall's tau distance $d_\tau(\pi, \sigma)$ [21], which counts the number of pairwise disagreements (inversions) between π and σ:

$$d_\tau(\pi, \sigma) = |\{(i,j) : i < j, \ (\pi(i) < \pi(j)) \neq (\sigma(i) < \sigma(j))\}|.$$

We defer the details of implementation to the Appendix D.

8

Metrics. We measure predictive accuracy via the following metrics: (1) *Top-1/top-5 hit rates*: Probability the test ranking's top item appears in the top-1/top-5 predicted ranks (2) *Spearman's ρ*: rank correlation based on squared rank differences; (3) *Kendall's τ*: correlation based on pairwise agreements; (4) *Hamming distance*: fraction of differing positions; (5) *Pairwise accuracy*: agreement ratio of item pairs' relative ordering. Arrows next to each metric indicate whether higher ($\uparrow$) or lower ($\downarrow$) values correspond to better performance. Full description of these metrics appears in Appendix D.

MLE Evaluation on Real-World Datasets. Tables 1, 2, and 3 present results for football, basketball, and SUSHI datasets, respectively. The best results are bold-faced, showing that the L_α-Mallows model consistently outperforms the Plackett-Luce model and Mallow's τ model over several metrics on all datasets. Notably, the basketball dataset results (Table 2) highlight particularly poor performance of the Plackett–Luce model, indicated by negative correlation values, possibly due to overfitting. The better generalization of the L_α Mallows model is reflected in significantly higher predictive accuracy.

The parameters estimated from the L_α-Mallows model offer meaningful insights into ranking behavior. The dispersion parameter (β) captures stability—lower values imply frequent ranking shifts, while higher values indicate stability. The learned distance parameter quantifies ranking dynamics, highlighting deviations of the optimal model from traditional models like Kendall's τ and Spearman's ρ (e.g. $\alpha = 2$). The higher estimated $\alpha \approx 1.1$ indicates a lower likelihood of long-range swaps (e.g., a team ranked 10th defeating a top-ranked team) compared to football, which has a smaller $\alpha \approx 0.4$, allowing more frequent long-range swaps.

	L_α-**Mallow**	Mallow's τ	Plackett–Luce
Estimated α	0.442 ($\pm$ 0.061)	–	–
Estimated β	0.455 ($\pm$ 0.051)	–	–
$\uparrow$Spearman's ρ correlation	**0.094** ($\pm$ 0.010)	0.070 ($\pm$ 0.016)	0.093 ($\pm$ 0.009)
$\uparrow$Kendall's τ correlation	**0.070** ($\pm$ 0.007)	0.052 ($\pm$ 0.011)	0.065 ($\pm$ 0.007)
$\downarrow$Hamming distance	**0.888** ($\pm$ 0.005)	0.892 ($\pm$ 0.002)	0.920 ($\pm$ 0.001)
$\uparrow$Pairwise accuracy (%)	**53.5** ($\pm$ 0.4)	52.6 ($\pm$ 0.6)	53.3 ($\pm$ 0.3)
$\uparrow$Top-1 hit rate (%)	**8.0** ($\pm$ 0.1)	6.9 ($\pm$ 0.7)	3.5 ($\pm$ 0.1)
$\uparrow$Top-5 hit rate (%)	**41.9** ($\pm$ 0.5)	41.1 ($\pm$ 2.9)	30.5 ($\pm$ 0.7)

Table 1: College football dataset, model out-of-sample performance averaged over 50 independent trials (mean $\pm$ standard deviation).

	L_α-**Mallow**	Mallow's τ	Plackett–Luce
Estimated α	1.096 ($\pm$ 0.056)	–	–
Estimated β	0.178 ($\pm$ 0.019)	–	–
$\uparrow$Spearman's ρ correlation	**0.269** ($\pm$ 0.005)	0.174 ($\pm$ 0.011)	-0.020 ($\pm$ 0.009)
$\uparrow$Kendall's τ correlation	**0.199** ($\pm$ 0.004)	0.128 ($\pm$ 0.009)	-0.010 ($\pm$ 0.007)
$\downarrow$Hamming distance	**0.872** ($\pm$ 0.001)	0.880 ($\pm$ 0.001)	0.919 ($\pm$ 0.001)
$\uparrow$Pairwise accuracy (%)	**59.9** ($\pm$ 0.2)	56.4 ($\pm$ 0.4)	49.5 ($\pm$ 0.3)
$\uparrow$Top-1 hit rate (%)	**22.7** ($\pm$ 0.6)	17.5 ($\pm$ 0.3)	4.9 ($\pm$ 0.3)
$\uparrow$Top-5 hit rate (%)	**77.7** ($\pm$ 1.0)	68.4 ($\pm$ 0.5)	57.7 ($\pm$ 1.6)

Table 2: College basketball dataset, model out-of-sample performance averaged over 50 independent trials (mean $\pm$ standard deviation).

Large n Regime: MLE Evaluation on 100-Team College Sports Rankings. Our learning algorithm can efficiently scale to rankings over a large number of items (see the theoretical guarantees in 3.1). To evaluate the robustness and scalability of our model, we compared the performance of the L_α Mallow model against Kendall's τ-based model and Placket–Luce model on college sports datasets, this time considering 100 teams.

For the football dataset, teams were selected with the highest participation rates to ensure a sufficient number of full rankings remained. For the basketball dataset, 100 teams were selected at random. The results for both datasets are reported in Table 4 and Table 5, respectively. All experiments were conducted with truncation size $k = 7$.

9

	L_α-Mallow	Mallow's τ	Plackett-Luce
Estimated α	0.764 ($\pm$ 0.378)	–	–
Estimated β	0.159 ($\pm$ 0.172)	–	–
$\uparrow$ Spearman's ρ correlation	**0.043** ($\pm$ 0.006)	0.031 ($\pm$ 0.012)	-0.050 ($\pm$ 0.003)
$\uparrow$Kendall's τ correlation	**0.031** ($\pm$ 0.005)	0.023 ($\pm$ 0.009)	-0.037 ($\pm$ 0.002)
$\downarrow$Hamming distance	**0.893** ($\pm$ 0.001)	0.897 ($\pm$ 0.001)	0.916 ($\pm$ 0.000)
$\uparrow$Pairwise accuracy (%)	**51.6** ($\pm$ 0.2)	51.1 ($\pm$ 0.4)	48.2 ($\pm$ 0.1)
$\uparrow$Top-1 hit rate (%)	10.2 ($\pm$ 0.6)	**10.6** ($\pm$ 0.6)	9.3 ($\pm$ 0.1)
$\uparrow$Top-5 hit rate (%)	**52.0** ($\pm$ 1.0)	51.7 ($\pm$ 1.6)	49.5 ($\pm$ 0.3)

Table 3: SUSHI dataset, model out-of-sample performance averaged over 25 independent trials (mean $\pm$ standard deviation).

Tables 4 and 5 report out-of-sample performance averaged over 50 independent trials. Across both datasets, the L_α-Mallows model significantly outperforms classical baselines (Mallows-τ, Plackett–Luce) across all metrics. The flexibility of choosing the distance metric is particularly evident as the number of teams increases in both datasets, the L_α Mallow model achieves more than fivefold improvement in top-1 hit rate compared to the baselines, significantly outperforming both Kendall's τ and Placket–Luce models.

	L_α-Mallow	Mallow's τ	Plackett–Luce
Estimated α	0.002 ($\pm$ 0.002)	–	–
Estimated β	0.504 ($\pm$ 0.023)	–	–
$\uparrow$Spearman's ρ correlation	**0.759** ($\pm$ 0.006)	0.478 ($\pm$ 0.006)	0.373 ($\pm$ 0.006)
$\uparrow$Kendall's τ correlation	**0.564** ($\pm$ 0.005)	0.328 ($\pm$ 0.005)	0.253 ($\pm$ 0.004)
$\downarrow$Hamming distance	**0.973** ($\pm$ 0.000)	0.984 ($\pm$ 0.000)	0.986 ($\pm$ 0.000)
$\uparrow$Pairwise accuracy (%)	**96.378** ($\pm$ 0.100)	91.754 ($\pm$ 0.152)	90.298 ($\pm$ 0.138)
$\uparrow$Top-1 hit rate (%)	**12.873** ($\pm$ 0.941)	2.212 ($\pm$ 0.807)	1.749 ($\pm$ 0.250)
$\uparrow$Top-5 hit rate (%)	**56.465** ($\pm$ 2.315)	10.857 ($\pm$ 1.401)	7.205 ($\pm$ 0.605)

Table 4: College basketball dataset for 100 teams, model out-of-sample performance averaged over 50 independent trials (mean $\pm$ standard deviation).

	L_α-Mallow	Mallow's τ	Plackett–Luce
Estimated α	0.003 ($\pm$ 0.002)	–	–
Estimated β	0.516 ($\pm$ 0.030)	–	–
$\uparrow$Spearman's ρ correlation	**0.454** ($\pm$ 0.006)	0.387 ($\pm$ 0.007)	0.138 ($\pm$ 0.005)
$\uparrow$Kendall's τ correlation	**0.318** ($\pm$ 0.004)	0.264 ($\pm$ 0.005)	0.093 ($\pm$ 0.004)
$\downarrow$Hamming distance	**0.981** ($\pm$ 0.000)	0.986 ($\pm$ 0.000)	0.989 ($\pm$ 0.000)
$\uparrow$Pairwise accuracy (%)	**91.163** ($\pm$ 0.165)	88.134 ($\pm$ 0.178)	86.944 ($\pm$ 0.133)
$\uparrow$Top-1 hit rate (%)	**2.057** ($\pm$ 0.535)	0.294 ($\pm$ 0.399)	1.524 ($\pm$ 0.899)
$\uparrow$Top-5 hit rate (%)	**23.590** ($\pm$ 3.979)	2.386 ($\pm$ 1.240)	7.097 ($\pm$ 1.154)

Table 5: College football dataset for 100 teams, model out-of-sample performance averaged over 50 independent trials (mean $\pm$ standard deviation).

Notably, the estimated distance parameter (α) remains close to zero in both datasets, suggesting that the model adaptively flattens the penalty for long-range swaps when necessary, capturing the heterogeneity in upset dynamics across sports. This reinforces the utility of learning α from data rather than fixing it a priori.

Effect of the Truncation size.
We demonstrate the exponential decay of the ℓ_2 estimation error on synthetic data. Specifically, we generate 50 training samples over 15 items using our sampling method with truncation order $k = 9$ and parameters $\alpha_0 = 1$ and $\beta_0 = 1$. We then train our model using varying truncation levels and report the mean and standard deviation of the estimation error in Figure 2. Notably, even though the training data was generated with a truncation of order 9, using a truncation as small as $k = 5$ yields stable and accurate estimates, highlighting the robustness of the model.
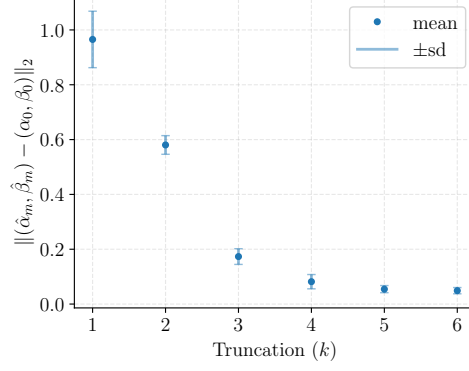
Figure 2: Estimation error versus the truncation size used for training L_α Mallow's model. Each point shows the ℓ_2-test error, averaged across 25 independent trials. As k increases, the model receives exponentially more information, leading to an exponential decay rate of the estimation error.

Synthetic Validation. We further establish the accuracy and robustness of MLE on the synthetic dataset. Specifically, we train our model using samples generated by the truncated sampling algorithm (Algorithm 3) with a smaller truncation parameter ($k = 6$) compared to the truncation parameter used to generate the original synthetic data ($k = 9$). Despite this deliberate mismatch, Figure 3 illustrates that our MLE procedure accurately recovers the true underlying parameters α and β, highlighting the robustness of our estimation method to the truncation choice.
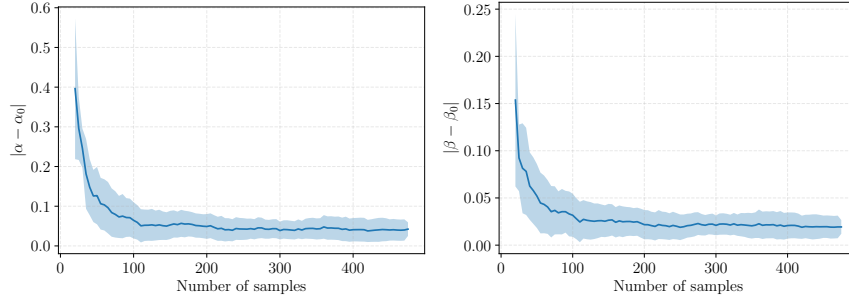

Figure 3: The estimation error of α (left) and β (right), for different training sizes over the synthetic data with $\alpha_0 = 1.5$ and $\beta_0 = 0.5$. The shaded area represents the standard deviation, computed over 25 independent trials. Our results show negligible estimation bias, despite using the truncation algorithm.

The synthetic experiments, where the ground truth parameters are known, confirm the accuracy and robustness of our parameter recovery.

Overall, our experiments validate that the L_α-Mallows model provides superior predictive accuracy, meaningful parameter interpretability, and computational efficiency, positioning it as an attractive method for ranking analysis across diverse datasets.

4.2 Validation of the Sampling Algorithm

Partition function. We first validate the accuracy of our approximation for the partition function $Z_n(\alpha, \beta)$. We quantify approximation quality by measuring the relative error: $\frac{|Z_n(\alpha,\beta) - \hat{Z}_n(\alpha,\beta)|}{Z_n(\alpha,\beta)}$, across different values of n, with a fixed truncation parameter $k = 3$. Additionally, we report the efficiency of our estimation by comparing our approximation's runtime against the exact computation, which computes the partition function constant exactly. We use Ryser's algorithm [41], which is regarded as one of the most efficient permanent computation algorithms. Despite using Ryser's algorithm, we still observe the overwhelming difference in the efficiency of our approximation, while only enduring negligible error. Table 6 summarizes our findings. We consistently observe small relative errors (on the order of 10^{-4}), validating the high accuracy of our approximation. More

11

importantly, our approximation significantly reduces computation time, especially as n increases. Figure 4 provides a visual comparison of runtime ratios and relative errors across different α and β.

n	Relative error	Algorithm 2 time (s)	Exact computation time (s)	Time ratio (Exact/Alg. 2)
6	$(1.25 \pm 0.03) \times 10^{-4}$	$0.011\ (\pm 0.002)$	$0.355\ (\pm 0.012)$	**32.0 ($\pm$5.9)**
8	$(2.10 \pm 0.05) \times 10^{-4}$	$0.027\ (\pm 0.003)$	$32.25\ (\pm 0.98)$	**1194.5 ($\pm$135.8)**
10	$(2.94 \pm 0.06) \times 10^{-4}$	$0.057\ (\pm 0.004)$	$6471.3\ (\pm 220.25)$	**113531.6 ($\pm$8166.4)**

Table 6: Relative error and runtime comparison between our approximation method (truncation order $k = 3$) and exact partition function computation via Ryser's algorithm, with parameters $\alpha = 1$ and $\beta = 2$. Results are averaged over 1000 trials, and 95% confidence intervals are reported.
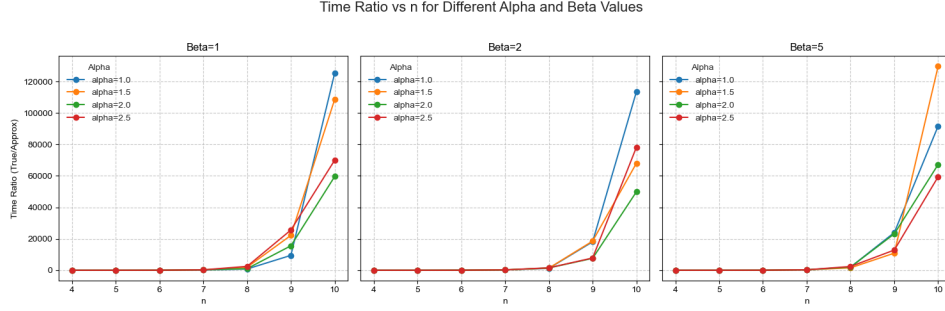

Figure 4: Runtime comparison between the permanent approximation ($k = 3$) and the full computation for varying values of n.

Limitations and Future Work. Our results highlight several promising directions, yet they also point clearly to opportunities for future work. A fundamental theoretical limitation lies in our efficient sampling algorithm, which currently requires a ($\alpha \geq 1$). Developing computationally efficient sampling for the broader class of distance functions with $\alpha < 1$ remains open. On the practical side, extending the empirical scope beyond sports rankings—potentially examining applications in recommendation systems, voting scenarios, or other domains—represents an exciting and natural next step, allowing further validation of the generalized L_α-Mallows framework.

References

[1] College Poll Archive. 2021 final ap men's basketball poll, 2025. Accessed: 2025-05-15.

[2] Kiana Asgari and Yeganeh Alimohammadi. Generalized mallow learning. `https://github.com/Kiana-Asgari/Generalized_Mallow_Learning`, 2025. Accessed: 2025-05-22.

[3] Pranjal Awasthi, Avrim Blum, Or Sheffet, and Aravindan Vijayaraghavan. Learning mixtures of ranking models. *Advances in Neural Information Processing Systems*, 27, 2014.

[4] John Bartholdi, Craig A Tovey, and Michael A Trick. Voting schemes for which it can be difficult to tell who won the election. *Social Choice and welfare*, 6:157–165, 1989.

[5] Ivona Bezáková, Daniel Štefankovič, Vijay V Vazirani, and Eric Vigoda. Accelerating simulated annealing for the permanent and combinatorial counting problems. *SIAM Journal on Computing*, 37(5):1429–1454, 2008.

[6] Ralph Allan Bradley and Milton E. Terry. Rank analysis of incomplete block designs: The method of paired comparisons. *Biometrika*, 39(3/4):324–345, 1976.

[7] Mark Braverman and Elchanan Mossel. Sorting from noisy information. *arXiv preprint arXiv:0910.1191*, 2009.

[8] Róbert Busa-Fekete, Dimitris Fotakis, Balázs Szörényi, and Manolis Zampetakis. Optimal learning of mallows block model. In *Conference on learning theory*, pages 529–532. PMLR, 2019.

[9] Haoxian Chen, Hanyang Zhao, Henry Lam, David Yao, and Wenpin Tang. Mallows-dpo: Fine-tune your llm with preference dispersions. *International Conference on Learning Representations, (ICLR)*, 2025.

[10] Antoine Désir, Vineet Goyal, Srikanth Jagabathula, and Danny Segev. Assortment optimization under the mallows model. *Advances in Neural Information Processing Systems*, 29, 2016.

[11] Antoine Désir, Vineet Goyal, Srikanth Jagabathula, and Danny Segev. Mallows-smoothed distribution over rankings approach for modeling choice. *Operations Research*, 69(4):1206–1227, 2021.

[12] Persi Diaconis. Group representations in probability and statistics. *Lecture notes-monograph series*, 11:i–192, 1988.

[13] Jean-Paul Doignon, Aleksandar Pekeč, and Michel Regenwetter. The repeated insertion model for rankings: Missing link between two subset choice models. *Psychometrika*, 69(1):33–54, 2004.

[14] Ran Duan and Seth Pettie. Approximating maximum weight matching in near-linear time. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pages 673–682. IEEE, 2010.

[15] Yifan Feng and Yuxuan Tang. On a mallows-type model for (ranked) choices. *Advances in Neural Information Processing Systems*, 35:3052–3065, 2022.

[16] Michael A. Fligner and James S. Verducci. Distance based ranking models. *Journal of the Royal Statistical Society: Series B (Methodological)*, 48(3):359–369, 1986.

[17] Charles J Geyer and Elizabeth A Thompson. Constrained monte carlo maximum likelihood for dependent data. *Journal of the Royal Statistical Society: Series B (Methodological)*, 54(3):657–683, 1992.

[18] E Irurozki, B Calvo, and A Lozano. Sampling and learning the mallows and generalized mallows models under the hamming distance. *Bernoulli (submitted)*, 18:19, 2014.

[19] Mark Jerrum, Alistair Sinclair, and Eric Vigoda. A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries. *J. ACM*, 51(4):671–697, 2004.

[20] Toshihiro Kamishima. Sushi preference dataset. `http://www.kamishima.net/sushi/`, 2005.

[21] Maurice G Kendall. A new measure of rank correlation. *Biometrika*, 30(1-2):81–93, 1938.

[22] Jon Kleinberg and Manish Raghavan. Algorithmic monoculture and social welfare. *Proceedings of the National Academy of Sciences*, 118(22):e2018340118, 2021.

[23] Guy Lebanon and Yi Mao. Non-parametric modeling of partially ranked data. *Advances in neural information processing systems*, 20, 2007.

[24] Allen Liu and Ankur Moitra. Efficiently learning mixtures of mallows models. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 627–638. IEEE, 2018.

[25] Dong C. Liu and Jorge Nocedal. On the limited memory bfgs method for large scale optimization. *Mathematical Programming*, 45(1-3):503–528, 1989.

[26] Tyler Lu and Craig Boutilier. Effective sampling and learning for mallows models with pairwise-preference data. *J. Mach. Learn. Res.*, 15(1):3783–3829, 2014.

[27] R Duncan Luce. *Individual choice behavior*, volume 4. Wiley New York, 1959.

[28] Chris Maddison, Daniel Tarlow, and Tom Minka. A* sampling, 2014.

[29] Colin L. Mallows. Non-null ranking models. i. *Biometrika*, 44(1–2):114–130, 1957.

[30] Cheng Mao and Yihong Wu. Learning mixtures of permutations: Groups of pairwise comparisons and combinatorial method of moments. *The Annals of Statistics*, 50(4):2231–2255, 2022.

[31] Marina Meilă, Annelise Wagner, and Christopher Meek. Recursive inversion models for permutations. *Statistics and Computing*, 32(4):54, 2022.

[32] Sumit Mukherjee. Estimation in exponential families on permutations. *The Annals of Statistics*, pages 853–875, 2016.

[33] Sumit Mukherjee and Daiki Tagami. Inference on a class of exponential families on permutations. *arXiv preprint arXiv:2304.04074*, 2023.

[34] James E Newman and Moshe Y Vardi. Fpras approximation of the matrix permanent in practice. *arXiv e-prints*, page : 2012.03367, 2020.

[35] Robin L Plackett. The analysis of permutations. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 24(2):193–202, 1975.

[36] Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D Manning, Stefano Ermon, and Chelsea Finn. Direct preference optimization: Your language model is secretly a reward model. *Advances in Neural Information Processing Systems*, 36, 2024.

[37] Label Ranking. Label ranking methods based on the plackett-luce model. *Poster presented at ICML*, page 27, 2010.

[38] Massey Ratings. College football/basketball/baseball rankings, 2025. Accessed: 2025-05-15.

[39] Bleacher Report. Ap college football poll: Final top 25 rankings after cfp championship, 2022. Accessed: 2025-05-15.

[40] Aviad Rubinstein and Shai Vardi. Sorting from noisier samples. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 960–972. SIAM, 2017.

[41] Herbert John Ryser. *Combinatorial Mathematics*, volume 14. Mathematical Association of America, 1963.

[42] Piotr Sankowski. Alternative algorithms for counting all matchings in graphs. In Helmut Alt and Michel Habib, editors, *STACS 2003*, pages 427–438, Berlin, Heidelberg, 2003. Springer Berlin Heidelberg.

[43] Nihar B Shah and Martin J Wainwright. Simple, robust and optimal ranking from pairwise comparisons. *Journal of machine learning research*, 18(199):1–38, 2018.

[44] Rainer Storn and Kenneth Price. Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, 11:341–359, 1997.

[45] Wenpin Tang. Mallows ranking models: maximum likelihood estimate and regeneration. In *International Conference on Machine Learning*, pages 6125–6134. PMLR, 2019.

[46] Aad W Van der Vaart. *Asymptotic statistics*, volume 3. Cambridge university press, 2000.

[47] Valeria Vitelli, Øystein Sørensen, Marta Crispino, Arnoldo Frigessi, and Elja Arjas. Probabilistic preference learning with the mallows rank model. *Journal of Machine Learning Research*, 18(158):1–49, 2018.

[48] Jacques Wainer. A bayesian bradley-terry model to compare multiple ml algorithms on multiple data sets. *Journal of Machine Learning Research*, 24(341):1–34, 2023.

[49] John I Yellott. The relationship between luce's choice axiom, thurstone's theory of comparative judgment, and the double exponential distribution. *Journal of Mathematical Psychology*, 15(2):109–144, 1977.

[50] Chenyang Zhong. Mallows permutation models with l^1 and l^2 distances i: hit and run algorithms and mixing times. *arXiv preprint arXiv:2112.13456*, 2021.

A Proofs for Computing Central Ranking Estimator

We start by proving that normalizing constant Z_n is independent of choice of σ.

Proposition A.1 (Invariance of Partition Function). *For any fixed parameters $\alpha > 0$ and $\beta > 0$, the partition function*

$$Z_n(\beta, \alpha) = \sum_{\pi \in S_n} e^{-\beta d_\alpha(\pi, \sigma)}$$

is independent of the central ranking $\sigma \in S_n$.

Proof. Consider any two permutations $\sigma, \sigma' \in S_n$. Define the bijective mapping $f_{\sigma, \sigma'} : S_n \to S_n$ by

$$f_{\sigma, \sigma'}(\pi) = \pi \circ \sigma^{-1} \circ \sigma',$$

which satisfies $f_{\sigma, \sigma'}(\sigma) = \sigma'$. Since $f_{\sigma, \sigma'}$ is a bijection, suming over permutations is invariant under re-indexing. Thus,

$$Z_n(\beta, \alpha, \sigma') = \sum_{\pi \in S_n} e^{-\beta d_\alpha(f_{\sigma, \sigma'}(\pi), \sigma')} = \sum_{\pi \in S_n} e^{-\beta d_\alpha(\pi \circ \sigma^{-1} \circ \sigma', \sigma')}.$$

Now, observe by definition of the distance d_α (which depends only on relative positions of elements), we have

$$d_\alpha(\pi \circ \sigma^{-1} \circ \sigma', \sigma') = d_\alpha(\pi, \sigma).$$

Hence,

$$Z_n(\beta, \alpha, \sigma') = \sum_{\pi \in S_n} e^{-\beta d_\alpha(\pi, \sigma)} = Z_n(\beta, \alpha, \sigma).$$

Since this equality holds for any pair of permutations σ, σ', the partition function is invariant and thus independent of the choice of the central ranking. $\square$

Next we bring the proof of Proposition 2.2, which reduces the optimization problem regarding σ to minimum matching problem.

Proof of Proposition 2.2. Consider a complete bipartite graph $G = (U \cup V, E)$ with $|U| = |V| = n$. Label the vertices in U by $u_1, u_2, \ldots, u_n$ and those in V by $v_1, v_2, \ldots, v_n$. For each edge $e = (u_i, v_j)$, assign the weight

$$w(i, j) = -\sum_{l=1}^{m} |\pi^{(l)}(i) - j|^\alpha.$$

Any perfect matching $M \subseteq U \times V$ corresponds to a unique permutation $\sigma \in S_n$ via the rule $\sigma(i) = j$ whenever the edge $(u_i, v_j) \in M$. Indeed, M must match every u_i to exactly one v_j, and thus σ is well-defined and bijective on $\{1, \ldots, n\}$.

Under this correspondence, the matching M has total weight $\sum_{(u_i, v_j) \in M} w(i, j)$, which is exactly

$$\sum_{i=1}^{n} \sum_{l=1}^{m} |\pi^{(l)}(i) - \sigma(i)|^\alpha = -\sum_{l=1}^{m} d_\alpha(\pi^{(l)}, \sigma).$$

Hence finding optimal σ is equivalent to finding a minimum weight perfect matching. $\square$

B Proof of Theorem 2.4

B.1 Consistency of $\hat{\sigma}_m$ (Proof of Lemma 2.1)

We start with the first part of the theorem which is consistency of $\hat{\sigma}_m$. Recall that $\hat{\sigma}_m$ is answer to the following optimiztion

$$\hat{\sigma}_m = \arg \min_{\sigma \in S_n} \frac{1}{m} \sum_{l=1}^{m} d_\alpha(\pi^{(l)}, \sigma).$$

Note that as the sample size grows, by the law of large number the right hand side converges to $\mathbb{E}_{\pi \sim \mathbb{P}_{\alpha_0,\beta_0,\sigma_0}}[d_{\alpha_0}(\pi,\sigma)]$. We start the proof by showing that in the limit, we can change the distance function inside, and for any choice of $\alpha > 0$ the unique minimizer of this is σ_0.

So, we start by proving the second part of Lemma 2.1, i.e., our goal is to show that given any $\tilde{\alpha} > 0$, the following holds

$$\arg\min_{\sigma' \in S_n} \mathbb{E}_{\Pi \sim \mathbb{P}_{\alpha_0,\beta_0,\sigma_0}}[d_{\tilde{\alpha}}(\Pi,\sigma')] = \{\sigma_0\}.$$

Proof. Define

$$F(\mathbf{x}) = \sum_{\pi \in S_n}\left(\sum_{u=1}^{n}|\pi(u) - x_u|^{\tilde{\alpha}}\right)\mathbb{P}_{\alpha_0,\beta_0,\sigma_0}(\pi) = \sum_{u=1}^{n}\mathbb{E}\big[\,|\Pi(u) - x_u|^{\tilde{\alpha}}\,\big],$$

where Π is sampled from $\mathbb{P}_{\alpha_0,\beta_0,\sigma_0}$. Notice that if we restrict x to permutations in S_n, then we recover the original problem. We show that whenever $x \neq \sigma_0$, performing a sequence of transpositions that brings $\mathbf{x}$ closer to σ_0 strictly decreases $F(\mathbf{x})$, and hence the unique minimizer is $\mathbf{x} = \sigma_0$.

For simplicity, assume σ_0 is the identity permutation, and assume there is another σ' minimizing F. We will prove uniqueness of the minimizer in two stages: (i) First we show, swapping any inverted pair of positions strictly decreases the objective. (ii) In the second step, by repeatedly correcting inversions we decrease $F(x)$ down to $F(\sigma_0)$.

Step 1: Inversion swap. Suppose $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and there is an inverted pair of positions $k < i$ with $x_k > x_i$. Let $\mathbf{x}' = (x_1', \ldots, x_n')$ be the vector obtained by swapping x_k and x_i:

$$x_k' = x_i, \quad x_i' = x_k, \quad x_u' = x_u \text{ for } u \notin \{k,i\}.$$

Then we claim that

$$F(\mathbf{x}') < F(\mathbf{x}).$$

Set $s = x_i, t = x_k$ so that $s < t$. Observe that only the kth and ith terms in F change under the swap. Indeed,

$$F(\mathbf{x}) - F(\mathbf{x}') = \mathbb{E}\big[|\Pi(k) - t|^{\tilde{\alpha}}\big] + \mathbb{E}\big[|\Pi(i) - s|^{\tilde{\alpha}}\big] - \left(\mathbb{E}\big[|\Pi(k) - s|^{\tilde{\alpha}}\big] + \mathbb{E}\big[|\Pi(i) - t|^{\tilde{\alpha}}\big]\right).$$

Rearrange:

$$F(\mathbf{x}) - F(\mathbf{x}') = \left(\mathbb{E}[|\Pi(i) - s|^{\tilde{\alpha}}] - \mathbb{E}[|\Pi(i) - t|^{\tilde{\alpha}}]\right) - \left(\mathbb{E}[|\Pi(k) - s|^{\tilde{\alpha}}] - \mathbb{E}[|\Pi(k) - t|^{\tilde{\alpha}}]\right).$$

Define

$$D(j) = |j - s|^{\tilde{\alpha}} - |j - t|^{\tilde{\alpha}}, \quad j = 1, \ldots, n.$$

Since $\tilde{\alpha} > 0$, $x \mapsto |x - s|^{\tilde{\alpha}} - |x - t|^{\tilde{\alpha}}$ is strictly increasing when $x \in [s,t]$, so

$$\Delta_j = D(j) - D(j-1) = |j-s|^{\tilde{\alpha}} - |j-1-s|^{\tilde{\alpha}} - |j-t|^{\tilde{\alpha}} + |j-1-t|^{\tilde{\alpha}} > 0 \quad \text{for } j = s+1, \ldots, t,$$

and $\Delta_j \geq 0$ otherwise (in fact $\Delta_j = 0$ for $\tilde{\alpha} = 1$ when $j \notin \{s+1, s+2, \ldots, t\}$).

Then note that

$$\mathbb{E}[|\Pi(u) - s|^{\tilde{\alpha}}] - \mathbb{E}[|\Pi(u) - t|^{\tilde{\alpha}}] = \sum_{j=1}^{n} D(j)\,\mathbb{P}_{\alpha_0,\beta_0,\sigma_0}(\Pi(u) = j)$$

$$= \sum_{j=2}^{n}[D(j) - D(j+1)]\,\mathbb{P}_{\alpha_0,\beta_0,\sigma_0}(\Pi(u) < j)$$

$$= -\sum_{j=s+1}^{t}\Delta_{j+1}\mathbb{P}_{\alpha_0,\beta_0,\sigma_0}(\Pi(u) < j).$$

Applying this for $u = i$ and $u = k$ gives

$$F(x) - F(x') = \sum_{j=s+1}^{t}\Delta_{j+1}\left[\mathbb{P}_{\alpha_0,\beta_0,\sigma_0}(\Pi(k) < j) - \mathbb{P}_{\alpha_0,\beta_0,\sigma_0}(\Pi(i) < j)\right].$$

17

Since $\Delta_j \geq 0$ and it strictly positive for $j \in s+1, \ldots, t$, it remains to show $\mathbb{P}_{\alpha_0, \beta_0, \sigma_0}(\Pi(k) < j) \geq \mathbb{P}_{\alpha_0, \beta_0, \sigma_0}(\Pi(i) < j)$ for each j. To see this, define an involution ϕ on S_n by swapping the values in positions k and i, where $k < i$. Whenever π satisfies $\pi(k) < \pi(i)$, by repeating the argument above,

$$d_{\alpha_0}\big(\phi(\pi), \mathrm{id}\big) - d_{\alpha_0}(\pi, \mathrm{id}) = \big|\pi(i) - k\big|^{\alpha_0} - \big|\pi(i) - i\big|^{\alpha_0} - \big(\big|\pi(k) - k\big|^{\alpha_0} - \big|\pi(k) - i\big|^{\alpha_0}\big) > 0,$$

so $\mathbb{P}_{\alpha_0, \beta_0, \sigma_0}(\Pi = \phi(\pi)) < \mathbb{P}_{\alpha_0, \beta_0, \sigma_0}(\Pi = \pi)$. Hence summing over all π with $\pi(k) < j \leq \pi(i)$ yields

$$\mathbb{P}_{\alpha_0, \beta_0, \sigma_0}\big(\Pi(k) < j, \ \Pi(i) \geq j\big) > \mathbb{P}_{\alpha_0, \beta_0, \sigma_0}\big(\Pi(i) < j, \ \Pi(k) \geq j\big).$$

Adding $\mathbb{P}_{\alpha_0, \beta_0, \sigma_0}\big(\Pi(k) < j, \Pi(i) < j\big)$ to both sides gives $\mathbb{P}_{\alpha_0, \beta_0, \sigma_0}(\Pi(k) < j) > \mathbb{P}_{\alpha_0, \beta_0, \sigma_0}\big(\Pi(i) < j\big)$. Therefore, each term in the sum is nonnegative, and at least one is strictly positive, so $F(x) - F(x') > 0$. This completes the proof of the first part.

Step 2: Eliminating inversions via successive swaps. We now transform any $\mathbf{x} \neq \sigma_0$ to σ_0 by a finite sequence of single-inversion swaps, each of which strictly decreases $F(x)$. To see this, not that if $\mathbf{x} \neq \sigma_0$ (recall that we assume σ_0 is the identity), then the sequence $(x_1, \ldots, x_n)$ is not increasing, so there exists at least one pair of indices

$$1 \leq k < i \leq n \quad \text{with} \quad x_k > x_i.$$

By Step 1, swapping the two entries at positions k and i produces a new permutation $\mathbf{x}'$ with $F(\mathbf{x}') < F(\mathbf{x})$, because (k, i) was an inversion. Now, by repeating this argument, since there are only finitely many permutations and F strictly decreases at each step, the process terminates in a finite number of swaps, necessarily at the unique no-inversion state $\mathbf{x} = \sigma_0$.

Putting this all together, we conclude that any $\mathbf{x} \neq \sigma_0$ can be changed to get a lower F value, so

$$\sigma_0 = \arg\min_{\mathbf{x} \in S_n} F(\mathbf{x}) = \arg\min_{\sigma' \in S_n} \mathbb{E}_{\alpha_0, \beta_0, \sigma_0}\big[d_{\tilde{\alpha}}(\Pi, \sigma')\big],$$

and the minimizer is unique. $\qquad\qquad\square$

Now, we are ready to finish the proof of Lemma 2.1 (and hence the first part of Theorem 2.4). Define the empirical objective

$$F_m(\sigma') = \frac{1}{m}\sum_{j=1}^{m} d_1(\pi^{(j)}, \sigma'),$$

and let

$$\hat{\sigma}_m = \arg\min_{\sigma' \in S_n} F_m(\sigma') \quad \text{(ties broken arbitrarily)}.$$

We want to prove that $\hat{\sigma}_m \to \sigma$ almost surely.

Since S_n is finite, the Strong Law of Large Numbers gives, for each $\sigma' \in S_n$,

$$F_m(\sigma') = \frac{1}{m}\sum_{j=1}^{m} d_{\tilde{\alpha}}(\pi^{(j)}, \sigma') \xrightarrow{\text{a.s.}} F(\sigma') = \mathbb{E}_{\Pi \sim \mathbb{P}_{\alpha_0, \beta_0, \sigma_0}}\big[d_{\tilde{\alpha}}(\Pi, \sigma')\big].$$

Since there are only $n!$ different permutations, the convergence is uniform:

$$\sup_{\sigma' \in S_n} \big|F_m(\sigma') - F(\sigma')\big| \xrightarrow{\text{a.s.}} 0.$$

Since F has a unique minimizer at σ_0 (as proved above), set $\delta = \min_{\sigma' \neq \sigma_0}\big[F(\sigma') - F(\sigma_0)\big] > 0$. By uniform convergence, for large enough m we have $\sup_{\sigma'} |F_m(\sigma') - F(\sigma')| < \delta/2$. Therefore, for all $\sigma' \neq \sigma_0$,

$$F_m(\sigma_0) \leq F(\sigma_0) + \tfrac{\delta}{2} < F(\sigma') - \tfrac{\delta}{2} \leq F_m(\sigma').$$

Thus for large enough m the empirical minimizer must converge to the true ranking ($\hat{\sigma}_m \to \sigma_0$). This proves almost sure convergence $\hat{\sigma}_m \xrightarrow{a.s.} \sigma_0$.

For the tail bound, we have a crude bound of $0 \leq d_{\tilde{\alpha}}(\pi, \sigma') \leq n^{\tilde{\alpha}+1}$, so by Hoeffding's inequality, for any fixed $\sigma' \neq \sigma_0$,

$$\mathbb{P}\Big(\big|F_m(\sigma') - F(\sigma')\big| \geq \tfrac{\delta}{2}\Big) \leq 2\exp\Big(-\tfrac{2m(\delta/2)^2}{(n^{\tilde{\alpha}+1})^2}\Big) = 2\exp\Big(-\tfrac{m\delta^2}{2\, n^{2\tilde{\alpha}+2}}\Big),$$

18

here the probability is over m i.i.d samples from $\mathbb{P}_{\alpha_0,\beta_0,\sigma_0}$ Applying the union bound over all $|S_n| = n!$ choices of σ',

$$\mathbb{P}\big(\sup_{\sigma'} |F_m(\sigma') - F(\sigma')| > \delta/2\big) \;\leq\; n! \;\cdot\; \exp\Big(-\frac{m\,\delta^2}{2\,n^{2\tilde{\alpha}+2}}\Big)$$

As a result,

$$\mathbb{P}\big(\hat{\sigma}_m \neq \sigma_0\big) \;\leq\; n! \;\cdot\; \exp\Big(-\frac{m\,\delta^2}{2\,n^{2\tilde{\alpha}+2}}\Big)$$

which in particular decays to 0 exponentially in m.

To finish the proof of the first part, it remains to find lower bound for

$$\delta = \min_{\sigma' \neq \sigma_0} \big[F(\sigma') - F(\sigma_0)\big] > 0.$$

Without loss of generality, assume that σ_0 is the identity permutation. As shown earlier, each transposition strictly decreases F, so it suffices to consider only transpositions. Let $\sigma' = (i\ j)$ be a transposition, and define $h := |i - j| \geq 1$ as the separation between the swapped positions. The expected cost difference decomposes as:

$$\delta_{i,j} := \mathbb{E}\big[|\Pi(i) - j|^{\tilde{\alpha}} - |\Pi(i) - i|^{\tilde{\alpha}}\big] + \mathbb{E}\big[|\Pi(j) - i|^{\tilde{\alpha}} - |\Pi(j) - j|^{\tilde{\alpha}}\big].$$

Then $\delta \geq \min_{i<j} \delta_{i,j}$.

Since $\Pi(i), \Pi(j), i, j \in \{1, \ldots, n\}$ are integers, the function $|\cdot|^{\tilde{\alpha}}$ is evaluated only on integer arguments. The minimum nonzero difference of the form $|x - j|^{\tilde{\alpha}} - |x - i|^{\tilde{\alpha}}$ must be at least

$$\min_{ij} \delta_{i,j} \geq \min_{0 \leq z \leq n-1} \big[(z+1)^{\tilde{\alpha}} - z^{\tilde{\alpha}}\big] = n^{\tilde{\alpha}} - (n-1)^{\tilde{\alpha}},$$

since the function $z \mapsto (z+1)^{\tilde{\alpha}} - z^{\tilde{\alpha}}$ is decreasing in z for $\tilde{\alpha} > 0$. By using the fact that $n^{\tilde{\alpha}} - (n-1)^{\tilde{\alpha}} \geq \tilde{\alpha}(n-1)^{\tilde{\alpha}-1}$ and that $(1 - \frac{1}{n})^{2\alpha} \geq 1 - \frac{2\alpha}{n} - \frac{2\alpha^2}{n^2}$, we the desired inequality:

$$\mathbb{P}\big(\hat{\sigma}_m \neq \sigma_0\big) \;\leq\; n! \;\cdot\; \exp\Big(-m\big(\tfrac{\tilde{\alpha}^2}{2\,n^4}(1 - \frac{2\tilde{\alpha}}{n})\big)\Big).$$

If we let $\tilde{\alpha} = 1$, we get the desired inequality.

B.2 Consistency of the MLE for $\hat{\alpha}_m, \hat{\beta}_m$ (Part 2 of Theorem 2.4)

Throughout this section, we assume $\hat{\sigma}_m$ is estimated in the first step of MLE. Define the normalized likelihood function:

$$\ell_m(\alpha, \beta) \;=\; \frac{1}{m} \sum_{j=1}^{m} \log f\big(\pi^{(j)}; \alpha, \beta, \hat{\sigma}_m\big) \;=\; -\beta \frac{1}{m} \sum_{j=1}^{m} d_\alpha(\pi^{(j)}, \hat{\sigma}_m) \;-\; \ln Z(\alpha, \beta),$$

and let $\hat{\theta}_m = (\hat{\alpha}_m, \hat{\beta}_m)$ be any maximizer of ℓ_m over $\Theta = (0, \infty) \times (0, \infty)$. Further define,

$$\Psi(\theta) := \mathbb{E}[\Psi_m(\theta)], \quad \text{where } \theta = (\alpha, \beta).$$

Recall the definition of Ψ_m from Step 2 of Algorithm 1. We verify two key properties:

1. *Uniform convergence:*
$$\sup_{\theta \in \Theta} \big\|\Psi_m(\theta) - \Psi(\theta)\big\| \xrightarrow{\mathbb{P}} 0.$$

2. *Identifiability of the zero:* Let $\theta_0 = (\alpha_0, \beta_0)$ $\Psi(\theta_0) = 0$, and for every $\varepsilon > 0$,
$$\inf_{\|\theta - \theta_0\| \geq \varepsilon} \big\|\Psi(\theta)\big\| > 0.$$

Once these two properties are proved, Theorem 5.9 of [46] yields $\hat{\theta}_m \xrightarrow{\mathbb{P}} \theta_0 = (\alpha_0, \beta_0)$. We start to prove identifiability.

Lemma B.1. *(Identifiability) The population score $\Psi(\theta)$ satisfies $\Psi(\theta_0) = 0$ and for every $\varepsilon > 0$,*

$$\inf_{\|\theta - \theta_0\| \geq \varepsilon} \|\Psi(\theta)\| > 0.$$

Proof. A straightforward calculation gives

$$\Psi(\theta) = \begin{pmatrix} \mathbb{E}_{\alpha_0,\beta_0}\left[\dot{d}_\alpha(\pi)\right] - \mathbb{E}_{\alpha,\beta}\left[\dot{d}_\alpha(\pi)\right] \\ \mathbb{E}_{\alpha_0,\beta_0}\left[d_\alpha(\pi)\right] - \mathbb{E}_{\alpha,\beta}\left[d_\alpha(\pi)\right] \end{pmatrix},$$

here since $\hat{\sigma}_m$ is fixed, with abuse of notation we write $d_\alpha(\pi)$ instead of $d_\alpha(\pi, \hat{\sigma}_m)$ and $\mathbb{E}_{\alpha',\beta'}$ instead of $\mathbb{E}_{\mathbb{P}'_\alpha,\beta',\hat{\sigma}_m}$. With this, proving identifiability reduces to showing that the system of equations

$$\mathbb{E}_{\alpha,\beta}[d_\alpha(\pi)] = \mathbb{E}_{\alpha_0,\beta_0}[d_\alpha(\pi)], \quad \mathbb{E}_{\alpha,\beta}[\dot{d}_\alpha(\pi)] = \mathbb{E}_{\alpha_0,\beta_0}[\dot{d}_\alpha(\pi)]$$

admits a unique solution $(\alpha, \beta) = (\alpha_0, \beta_0)$.

To facilitate the argument, we fix (α, β) and view the left-hand sides as functions of (α_0, β_0). In this formulation, identifiability reduces to proving that for each fixed (α, β), there exists a unique (α_0, β_0) satisfying the system.

For this purpose, we introduce the auxiliary map

$$\boldsymbol{f}_\alpha(\alpha_0, \beta_0) = \left(\mathbb{E}_{\alpha_0,\beta_0}[d_\alpha(\pi)], \ \mathbb{E}_{\alpha_0,\beta_0}[\dot{d}_\alpha(\pi,)] \right),$$

where α is fixed. That is, while the distance function d_α is fixed, the true data-generating parameters (α_0, β_0) varies. The key idea is that if $\boldsymbol{f}_\alpha$ is injective, then the solution to the system is unique. We now proceed to prove that $\boldsymbol{f}_\alpha$ is a global diffeomorphism onto its image. To this end, we first establish local invertibility and properness, which together imply global invertibility by the Hadamard–Caccioppoli theorem.

Lemma B.2 (Local Invertibility)**.** *For any $\alpha > 0$, the function $\boldsymbol{f}_\alpha$ is locally invertible on $(0, \infty)^2$.*

Proof. It suffices to show that the Jacobian matrix $\boldsymbol{Jf}_\alpha(\alpha_0, \beta_0)$ is invertible at every point $(\alpha_0, \beta_0) \in (0, \infty)^2$. Direct calculation gives:

$$\boldsymbol{Jf}_\alpha(\alpha_0, \beta_0) = - \begin{bmatrix} \mathrm{Cov}_{\alpha_0,\beta_0}\left(d_\alpha(\pi), d_{\alpha_0}(\pi)\right) & \beta\,\mathrm{Cov}_{\alpha_0,\beta_0}\left(d_\alpha(\pi), \dot{d}_{\alpha_0}(\pi)\right) \\ \mathrm{Cov}_{\alpha_0,\beta_0}\left(\dot{d}_\alpha(\pi), d_{\alpha_0}(\pi)\right) & \beta\,\mathrm{Cov}_{\alpha_0,\beta_0}\left(\dot{d}_\alpha(\pi), \dot{d}_{\alpha_0}(\pi)\right) \end{bmatrix}.$$

Define the two-dimensional random vectors:

$$\boldsymbol{a}_1(\pi) = \left(d_{\alpha_0}(\pi), \dot{d}_{\alpha_0}(\pi) \right), \quad \boldsymbol{a}_2(\pi) = \left(d_\alpha(\pi), \dot{d}_\alpha(\pi) \right).$$

Then

$$\det\left(\boldsymbol{Jf}_{\tilde{\alpha}}(\alpha, \beta)\right) = \beta \cdot \det\left(\mathrm{Cov}_{\alpha,\beta}\left(\boldsymbol{a}_1(\pi), \boldsymbol{a}_2(\pi)\right)\right).$$

When $(\alpha_0, \beta_0) = (\alpha, \beta)$, we have $\boldsymbol{a}_1(\pi) = \boldsymbol{a}_2(\pi)$, and the covariance matrix reduces to the variance of $\boldsymbol{a}_1(\pi)$, which is positive definite.

When $\alpha_0 \neq \alpha$, suppose, for contradiction, that $\mathrm{Cov}_{\alpha_0,\beta_0}(\boldsymbol{a}_1(\Pi), \boldsymbol{a}_2(\Pi))$ is singular. Then there must exist nonzero vectors $\lambda^{(1)}, \lambda^{(2)} \in \mathbb{R}^2$ such that

$$\left\langle \lambda^{(1)}, \boldsymbol{a}_1(\Pi) \right\rangle + \left\langle \lambda^{(2)}, \boldsymbol{a}_2(\Pi) \right\rangle \quad \text{is constant almost surely.} \tag{2}$$

Since $\mathbb{P}_{\alpha_0,\beta_0}$ has full support on the finite set S_n, this equality holds for every $\pi \in S_n$.

To derive a contradiction, we consider a few examples. Let $\pi_1 = (2, 4)$ denote the transposition swapping elements 2 and 4, and π_2 denote the composition of π_1 with the transposition swapping 1 and 3, that is, $\pi_2 = (1\ 3)(2\ 4)$. Applying (2) to both π_1 and π_2, we have:

$$\left\langle \lambda^{(1)}, \boldsymbol{a}_1(\pi_1) \right\rangle + \left\langle \lambda^{(2)}, \boldsymbol{a}_2(\pi_1) \right\rangle = \left\langle \lambda^{(1)}, \boldsymbol{a}_1(\pi_2) \right\rangle + \left\langle \lambda^{(2)}, \boldsymbol{a}_2(\pi_2) \right\rangle.$$

By construction, we observe:

$$\boldsymbol{a}_1(\pi_2) = \boldsymbol{a}_1(\pi_1) + \boldsymbol{a}_1((1\ 3)), \quad \boldsymbol{a}_2(\pi_2) = \boldsymbol{a}_2(\pi_1) + \boldsymbol{a}_2((1\ 3)).$$

Substituting the expressions for π_2, we obtain:

$$\left\langle \lambda^{(1)}, \boldsymbol{a}_1((1\ 3)) \right\rangle + \left\langle \lambda^{(2)}, \boldsymbol{a}_2((1\ 3)) \right\rangle = 0.$$

Repeating this argument with other simple transpositions such as $(1\ 4)$, $(1\ 5)$, and $(1\ 6)$, we obtain a homogeneous linear system:

$$A \begin{pmatrix} \lambda^{(1)} \\ \lambda^{(2)} \end{pmatrix} = 0,$$

where

$$A = \begin{bmatrix} \boldsymbol{a}_1(1\ 3) & \boldsymbol{a}_2(1\ 3) \\ \boldsymbol{a}_1(1\ 4) & \boldsymbol{a}_2(1\ 4) \\ \boldsymbol{a}_1(1\ 5) & \boldsymbol{a}_2(1\ 5) \\ \boldsymbol{a}_1(1\ 6) & \boldsymbol{a}_2(1\ 6) \end{bmatrix}.$$

It can be checked that when $\alpha \neq \alpha_0$, the matrix $\boldsymbol{A}$ has full rank, Since the points $(d_{\alpha_0}(\pi), \dot{d}_{\alpha_0}(\pi), d_\alpha(\pi), \dot{d}_\alpha(\pi))$ for different transpositions $(1\ k)$ involve distinct powers of $|k-1|$ under α and α_0, and the functions $\alpha \mapsto |k-1|^\alpha$ and $\alpha \mapsto \log|k-1| \cdot |k-1|^\alpha$ are linearly independent for different k, the matrix $\boldsymbol{A}$ is full rank whenever $\alpha \neq \alpha_0$. As a result, $\lambda^{(1)} = \lambda^{(2)} = 0$. Thus, no nontrivial linear relation exists, contradicting the assumption that $\mathrm{Cov}_{\alpha_0,\beta_0}(\boldsymbol{a}_1(\Pi), \boldsymbol{a}_2(\Pi))$ is singular. This completes the proof of local invertibility. $\square$

Lemma B.3 (Properness). *For any α, the function $\boldsymbol{f}_\alpha$ is proper: the preimage of every compact set is compact.*

Proof. Let $K \subset \mathrm{Img}(\boldsymbol{f}_{\tilde{\alpha}})$ be any compact set. We must show that $\boldsymbol{f}_{\tilde{\alpha}}^{-1}(K)$ is compact. Suppose for contradiction that $\boldsymbol{f}_{\tilde{\alpha}}^{-1}(K)$ is not compact. Then there exists a sequence (α_m, β_m) in $\boldsymbol{f}_{\tilde{\alpha}}^{-1}(K)$ that escapes to infinity, meaning

$$\lim_{m \to \infty} \|(\alpha_m, \beta_m)\| = \infty.$$

As $(\alpha_m, \beta_m) \to \infty$, the corresponding distributions $\mathbb{P}_{\alpha_m,\beta_m}$ concentrate on the identity permutation. To see this as $(\alpha_m, \beta_m) \to \infty$, either $\alpha_m \to \infty$ or $\beta_m \to \infty$ or both. In any of these cases, for any permutation $\pi \neq \mathrm{id}$, we have

$$\frac{\mathbb{P}_{\alpha_m,\beta_m}(\pi)}{\mathbb{P}_{\alpha_m,\beta_m}(\mathrm{id})} = \exp\left(-\beta_m(d_{\alpha_m}(\pi) - d_{\alpha_m}(\mathrm{id}))\right) \to 0.$$

Thus, as m increases,

$$\boldsymbol{f}_{\tilde{\alpha}}(\alpha_m, \beta_m) \to \boldsymbol{0}.$$

But note that for any finite $(\alpha, \beta) \in (0,\infty)^2$, we always have:

$$\mathbb{E}_{\alpha,\beta}[d_{\tilde{\alpha}}(\Pi)] > 0, \quad \mathbb{E}_{\alpha,\beta}[\dot{d}_{\tilde{\alpha}}(\Pi)] > 0,$$

since $\mathbb{P}_{\alpha,\beta}$ has full support on S_n, and $d_{\tilde{\alpha}}(\Pi) > 0$ with positive probability. Therefore, $\boldsymbol{0} \notin \mathrm{Img}(\boldsymbol{f}_{\tilde{\alpha}})$. Now, since $\boldsymbol{f}_{\tilde{\alpha}}(\alpha_m, \beta_m) \in K$ for all m, and $\boldsymbol{f}_{\tilde{\alpha}}(\alpha_m, \beta_m) \to \boldsymbol{0}$, we conclude that $\boldsymbol{0} \in K$ by closedness of K. This contradicts $\boldsymbol{0} \notin \mathrm{Img}(\boldsymbol{f}_{\tilde{\alpha}})$. Therefore, the preimage $\boldsymbol{f}_{\tilde{\alpha}}^{-1}(K)$ must be compact. $\square$

Now, to finish the proof of uniqueness of the critical point, observe that the condition $\Psi(\theta) = 0$ is equivalent to

$$\boldsymbol{f}_\alpha(\alpha, \beta) = \boldsymbol{f}_\alpha(\alpha_0, \beta_0).$$

Since $\boldsymbol{f}_\alpha$ is a global diffeomorphism (by Lemmas B.2 and B.3), it follows that $(\alpha, \beta) = (\alpha_0, \beta_0)$, and hence $\Psi(\theta_0) = 0$ has a unique solution.

To complete the identifiability condition, we note that the global invertibility of $\boldsymbol{f}_\alpha$ implies that its inverse $\boldsymbol{f}_\alpha^{-1}$ is continuous on its image. Therefore, $\Psi(\theta) = \boldsymbol{f}_\alpha(\alpha_0, \beta_0) - \boldsymbol{f}_\alpha(\theta)$ is continuous, vanishes only at θ_0, and cannot tend to zero along any sequence staying at distance at least $\varepsilon > 0$ from θ_0. Thus, for every $\varepsilon > 0$,

$$\inf_{\|\theta - \theta_0\| \geq \varepsilon} \|\Psi(\theta)\| > 0,$$

as required. $\square$

Uniform Convergence of Ψ_m. Next we prove uniform convergence of Ψ_m to Ψ. Note that if the true-parameters of Mallows (α_0, β_0) where restricted to a compact set, then this was immediate. But the proof for general parameters needs careful analysis which we bring next.

First, fix any compact rectangle

$$\Theta_R := [1/R, R] \times [1/R, R] \subset (0, \infty) \times (0, \infty),$$

for some $R > 1$. Over Θ_R, $\Psi_m(\theta)$ and $\Psi(\theta)$ are uniformly bounded and equicontinuous as functions of θ. Thus, by pointwise convergence along with the law of large numbers, we have

$$\sup_{\theta \in \Theta_R} \|\Psi_m(\theta) - \Psi(\theta)\| \xrightarrow{\mathbb{P}} 0.$$

Now, outside Θ_R (that is, for $\theta \notin \Theta_R$), we claim that $\|\Psi(\theta)\|$ is bounded away from zero for large R, and hence the sequence θ_m lies inside Θ_R for some large R. Indeed, as $\|\theta\| \to \infty$, we know from the properness of $\boldsymbol{f}_\alpha$ (Lemma B.3) that $\|\Psi(\theta)\| \to \infty$ which implies that

$$\inf_{\theta \notin \Theta_R} \|\Psi(\theta)\| \geq \delta_R > 0$$

for some $\delta_R > 3\varepsilon$ (depending on R). Now, consider any $\theta \notin \Theta_R$. By the pointwise convergence of $\Psi_m(\theta) \to \Psi(\theta)$ as $m \to \infty$ for each fixed θ. Thus, for each fixed $\theta \notin \Theta_R$, given any small $\delta > 0$, there exists $m_0(\theta, \delta)$ such that for all $m \geq m_0$,

$$\mathbb{P}\left(\|\Psi_m(\theta) - \Psi(\theta)\| \leq \varepsilon\right) \geq 1 - \delta.$$

Moreover, since $\|\Psi(\theta)\| \geq 3\varepsilon$, if $\|\Psi_m(\theta) - \Psi(\theta)\| \leq \varepsilon$, then by the triangle inequality:

$$\|\Psi_m(\theta)\| \geq \|\Psi(\theta)\| - \|\Psi_m(\theta) - \Psi(\theta)\| \geq 2\varepsilon.$$

Therefore $\Psi_m(\theta) \neq 0$ for $\theta \notin \Theta_R$ and all m large enough. But by the first order optimality conditions we must have $\Psi_m(\theta_m) = 0$. As a result we must have the sequence is contained in a compact set i.e., for some large enough R, we have that $\theta_m \in \Theta_R$ with high probability as $m \to \infty$. Since Ψ_m has no minimizer outside Θ_R then by applying Theorem 5.9 of van der Vaart [46], we conclude that

$$\hat{\theta}_m = (\hat{\alpha}_m, \hat{\beta}_m) \xrightarrow{\mathbb{P}} (\alpha_0, \beta_0).$$

Asymptotic Normality. The final step to prove asymptotic normality by Theorem 5.23 of van der Vaart [46]. We verify the conditions of this theorem for the M-estimator $(\hat{\alpha}_m, \hat{\beta}_m)$, defined as the maximizer of the empirical objective $\ell_m(\theta)$.

(i) Consistency: Established previously using Theorem 5.9.

(ii) Differentiability: The log-likelihood $\ell_m(\theta)$ is twice continuously differentiable in $\theta = (\alpha, \beta)$ because both $d_\alpha(\pi, \sigma_0)$ and $Z(\alpha, \beta)$ are smooth in α and β.

(iii) Central Limit Theorem for the Score: By the classical central limit theorem, the empirical average of the first term of Ψ_m converges in distribution:

$$\sqrt{m}\left(\frac{1}{m}\sum_{j=1}^{m}\begin{pmatrix}\beta_0\,\dot{d}_{\alpha_0}(\pi^{(j)}) \\ d_{\alpha_0}(\pi^{(j)})\end{pmatrix} - \mathbb{E}\left[\begin{pmatrix}\beta_0\,\dot{d}_{\alpha_0}(\Pi) \\ d_{\alpha_0}(\Pi)\end{pmatrix}\right]\right) \xrightarrow{d} \mathcal{N}(0, B),$$

where B is the variance of the score under $\mathbb{P}_{\alpha_0, \beta_0, \sigma_0}$.

(iv) Hessian convergence: The Hessian $\nabla_\theta^2 \ell_m(\theta)$ is the sum of an empirical average and $\nabla^2 \ln Z(\theta)$, both of which converge uniformly by the law of large numbers and analytic smoothness of Z.

(v) Invertibility: The matrix $\mathcal{I}(\alpha_0, \beta_0) = -\nabla^2 \ell(\theta_0)$ is positive definite due to identifiability of Ψ (Lemma B.1). Thus, all conditions of Theorem 5.23 are satisfied, and the result follows. $\quad\square$

B.3 Asymptotic Normality under Approximate Score

Proof of Theorem 2.6. Let $\hat{\theta}_m$ be the exact MLE, satisfying $\Psi_m(\hat{\theta}_m) = 0$. Using Theorem 2.4,

$$\sqrt{m}(\hat{\theta}_m - \theta_0) \xrightarrow{d} \mathcal{N}(0, \mathcal{I}^{-1}).$$

22

We now analyze the perturbation due to the approximate score:

$$\widetilde{\Psi}_m(\tilde{\theta}_m) = \Psi_m(\tilde{\theta}_m) + \Delta(\tilde{\theta}_m) = 0.$$

Since $\Psi_m(\hat{\theta}_m) = 0$ and $\widetilde{\Psi}_m(\tilde{\theta}_m) = 0$, we write:

$$\widetilde{\Psi}_m(\tilde{\theta}_m) = \Psi_m(\tilde{\theta}_m) + \Delta(\tilde{\theta}_m) = 0,$$

which implies:

$$\Psi_m(\tilde{\theta}_m) = -\Delta(\tilde{\theta}_m).$$

Now apply a first-order Taylor expansion of Ψ_m around $\hat{\theta}_m$:

$$\Psi_m(\tilde{\theta}_m) = \Psi_m(\hat{\theta}_m) + \nabla\Psi_m(\hat{\theta}_m)(\tilde{\theta}_m - \hat{\theta}_m) + R_m,$$

where R_m is the second-order remainder:

$$R_m = o(\|\tilde{\theta}_m - \hat{\theta}_m\|).$$

Since $\Psi_m(\hat{\theta}_m) = 0$, this gives:

$$\nabla\Psi_m(\hat{\theta}_m)(\tilde{\theta}_m - \hat{\theta}_m) = -\Delta(\tilde{\theta}_m) + o(\|\tilde{\theta}_m - \hat{\theta}_m\|).$$

By assumption, $\nabla\Psi_m(\hat{\theta}_m) \xrightarrow{P} \mathcal{I}(\theta_0)$, which is invertible. Thus, for large enough m, we can solve:

$$\tilde{\theta}_m - \hat{\theta}_m = -\nabla\Psi_m(\hat{\theta}_m)^{-1}\Delta(\tilde{\theta}_m) + o(\|\tilde{\theta}_m - \hat{\theta}_m\|).$$

Using the bound $\|\Delta(\tilde{\theta}_m)\| = o(m^{-1/2})$, we conclude:

$$\|\tilde{\theta}_m - \hat{\theta}_m\| = o_P(m^{-1/2}) \Rightarrow \sqrt{m}(\tilde{\theta}_m - \hat{\theta}_m) = o_P(1).$$

Combining with the asymptotic normality of $\hat{\theta}_m$,

$$\sqrt{m}(\tilde{\theta}_m - \theta_0) = \sqrt{m}(\hat{\theta}_m - \theta_0) + o_P(1) \xrightarrow{d} \mathcal{N}(0, \mathcal{I}^{-1}).$$

$\square$

C   Proofs for Efficient Sampling

C.1   Geometric Decay of Marginal Distributions

The main goal of this section is proving Lemma 3.2. We break the argument into a few steps. The first observation is that the marginal probability is symmetric in i and j, i.e., we claim that

$$P_n(i,j) = P_n(n+1-j, n+1-i).$$

This is true because of a standard relabeling argument implies that the probability of having $\pi(i) = j$ is the same as the probability of having $\pi(n+1-j) = n+1-i$ once we account for symmetry in d_α. Therefore, without loss of generality, it is enough to prove that for $j - i > k$,

$$\frac{P_n(i, j+1)}{P_n(i,j)} \le c(\alpha, \beta).$$

To proceed with the rest of the proof, we need the following definitions. Define the sets of permutations:

$$S^{(i,j)} := \{\pi \in S_n : \pi(i) = j\} \quad \text{and} \quad S^{(i,j+1)} := \{\pi \in S_n : \pi(i) = j+1\}.$$

and the weight function $w : S_n \to \mathbb{R}^+$ for a permutation π, which assigns to each permutation π its corresponding probability weight:

$$w(\pi) := e^{-\beta d_\alpha(\pi, \mathrm{id})} = \prod_{i=1}^n e^{-\beta|i-\pi(i)|^\alpha}. \tag{3}$$
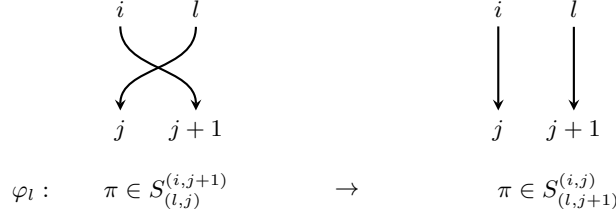
Figure 5: A schematic of the first mapping.

For any subset $S \subseteq S_n$, we extend this definition, with a slight abuse of notation, to the total weight of S as $w(S) = \sum_{\pi \in S} w(\pi)$.

In broad terms, the proof compares the weights of two sets of permutations $S^{(i,j)}$ and $S^{(i,j+1)}$. In fact, $\frac{P_n(i,j+1)}{P_n(i,j)} = \frac{w(S^{(i,j+1)})}{w(S^{(i,j)})}$. Thus, our objective is to prove $\frac{w(S^{(i,j+1)})}{w(S^{(i,j)})} < c(\alpha, \beta)$ for large $j - i$.

For this purpose we find mappings from permutations in $w(S^{(i,j+1)})$ to $w(S^{(i,j)})$ that shows the total weight ratio is bounded.

- **First Round of Mappings (Proposition C.1).** We break down the set $S^{(i,j+1)}$ by looking at which index was mapped to j. We build a mapping (a "swap") that sends permutations in $S^{(i,j+1)}_{(l,j)} := \{\pi \in S_n : \pi(i) = j + 1, \pi(l) = j\}$ to the permutations in $S^{(i,j)}_{(l,j+1)} := \{\pi \in S_n : \pi(i) = j, \pi(l) = j + 1\}$.
  - If $l > j$, we show the weight increases – directly helping us prove $S^{(i,j+1)}$ is smaller in total weight.
  - If $l \leq j$, the mapping might decrease the weight, so we cannot immediately conclude a reduction.
- **Second Round of Mappings (Proposition C.2).** For the "problematic" subset where $l \leq j$, we employ a second mapping by identifying the first index (larger than j) mapped to a value below j. Swapping this index with l brings the resulting permutation closer to *identity*, leading to an increased weight.

Finally, by uniting these two mapping arguments, we conclude $\frac{P_n(i,j+1)}{P_n(i,j)} < 1$. We now present the two mapping steps in detail.

C.1.1 First Round of Mappings

The first round of mapping is described above is done by swapping inverse of j and $j + 1$ which gives the following bounds.

Proposition C.1. *Suppose $j > i$, and let $S^{(i,j+1)}_{(l,j)}$ and $S^{(i,j)}_{(l,j+1)}$ be defined as above. Then for any $l \in [n] \setminus \{i\}$, and any $k \leq j - i$,*

$$\frac{w(S^{(i,j+1)}_{(l,j)})}{w(S^{(i,j)}_{(l,j+1)})} \leq \begin{cases} e^{-\beta((k+1)^\alpha - k^\alpha)} e^{-\beta((j+1-l)^\alpha - (j-l)^\alpha)} & l > j \\ e^{-\beta((k+1)^\alpha - k^\alpha)} e^{\beta((j+1-l)^\alpha - (j-l)^\alpha)} & l \leq j \end{cases}.$$

Proof. Define the map $\varphi_\ell : S^{(i,j+1)}_{(\ell,j)} \to S^{(i,j)}_{(\ell,j+1)}$ by swapping the *images* of i and ℓ in any permutation π (see Figure 5). More precisely,

$$(\varphi_\ell \circ \pi)(i) = \pi(\ell), \quad (\varphi_\ell \circ \pi)(\ell) = \pi(i), \quad (\varphi_\ell \circ \pi)(m) = \pi(m) \text{ for } m \neq i, \ell.$$

By direct inspection,

$$\log \frac{w(\pi)}{w(\varphi_l(\pi))} = -\beta \left(|j + 1 - i|^\alpha + |j - l|^\alpha \right) + \beta \left(|j - i|^\alpha + |j + 1 - l|^\alpha \right) \tag{4}$$

$$\leq \left(-\beta(k + 1)^\alpha + \beta(k)^\alpha \right) + \left(-\beta|j - l|^\alpha + \beta|j + 1 - l|^\alpha \right). \tag{5}$$

24

A simple bounding argument shows that the second term $(-\beta|j-l|^\alpha + \beta|j+1-l|^\alpha)$ is ≤ 0 if $\ell > j$. Since the upper bound is independent of π, by summing over all such $\pi \in S_{(\ell,j)}^{(i,j+1)}$ we finish the proof. $\qquad\square$

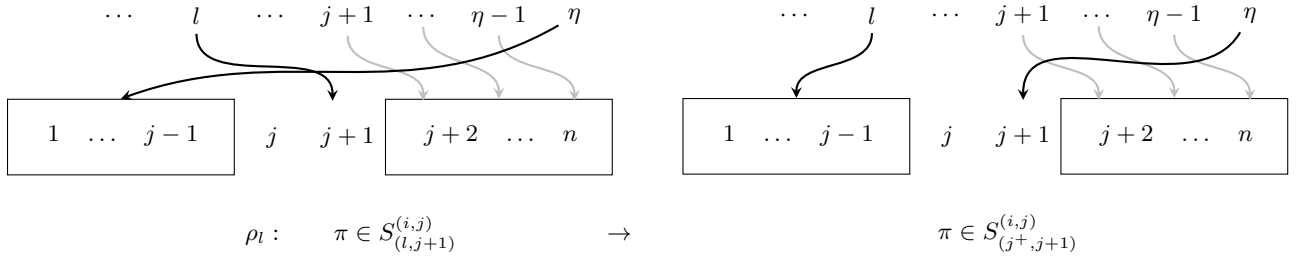### C.1.2 Handling the Problematic Cases with a Second Mapping

Figure 6: A schematic of the second mapping.

For $l \leq j$, we handle those cases by constructing a second mapping that locates a suitable index that swapping it with l will get a closer mapping to the central ranking (here identity). See Figure 6.

Proposition C.2. *Let $l \leq j$. Also define, $S_{(j^+,j+1)}^{(i,j)} = \bigcup_{k \geq j+1} S_{(k,j+1)}^{(i,j)}$. Then*

$$\frac{w(S_{(l,j+1)}^{(i,j)})}{w(S_{(j^+,j+1)}^{(i,j)})} \leq e^{-\beta(j-l)^\alpha + \beta|j-1-l|^\alpha}.$$

Proof. Fix a permutation $\pi \in S_{(\ell,j+1)}^{(i,j)}$. The first step is to find a mapping that maps π to another permutation in $S_{(j^+,j+1)}^{(i,j)}$ For each permutation $\pi \in S_{(\ell,j+1)}^{(i,j)}$, there exists at least one index $\eta \in \{j+1,\ldots,n\}$ such that $\pi(\eta) \leq j-1$. This is true because there are $n-j$ indices in $\{j+1,\ldots,n\}$, but $\pi(\ell) = j+1$ forces at most $n-j-1$ of those indices to map above $j+1$. Define η_π to be the smallest such index i.e., $\eta_\pi := \min\{m \in \{j+1,\ldots,n\} : \pi(m) \leq j-1\}$.

Using this, we define the projection $\rho_l : S_{(l,j+1)}^{(i,j)} \to S_{(j^+,j+1)}^{(i,j)}$, by swapping (l,η_π) in the permutation π, i.e.,

$$(\rho_\ell \circ \pi)(l) = \pi(\eta_\pi), \quad (\rho_\ell \circ \pi)(\eta_\pi) = \pi(l), \quad (\rho_\ell \circ \pi)(m) = \pi(m) \text{ for } m \neq \eta_\pi, \ell.$$

Injectivity. The map ρ_l is one-to-one, but not onto. In fact, for any $k \geq j+1$, any permutation $\pi' \in S_{(k,j+1)}^{(i,j)}$, can be uniquely inverted by swapping l and the corresponding k if it satisfies being in the image of ρ_l, i.e.,

$$\rho_l^{-1}(\pi') = \begin{cases} (l,k) \times \pi' & \text{if } \pi'(l) \leq j-1 \text{ and } \pi'(s) > j+1 \text{ for } s \in \{j+1, j+2, \ldots, k\} \\ null & \text{otherwise.} \end{cases} \tag{6}$$

See Figure 6. Hence, summing over $\pi \in S_{(l,j+1)}^{(i,j)}$ and comparing to $\rho_l(\pi)$ in the codomain involves no collisions.

Weight Change. Under the swap ρ_k, the weight $w(\pi)$ can change by a factor

$$\frac{w(\pi)}{w(\rho_l(\pi))} = exp\Big(-\beta\Big((j+1-l)^\alpha + (\eta_\pi - \pi(\eta_\pi))^\alpha - (\eta_\pi - j - 1)^\alpha - |l - \pi(\eta_\pi)|^\alpha\Big)\Big).$$

Now, if we sum up over all possible values of $\eta_\pi \geq j+1$, we get

$$w(S^{(i,j)}_{(l,j+1)}) = \sum_{k=j+1}^{n} \sum_{\substack{\pi' \in S^{(i,j)}_{(k,j+1)}: \\ \rho_l^{-1}(\pi') \neq null}} w(\rho_l^{-1}(\pi'))$$

$$= \sum_{k=j+1}^{n} \sum_{\substack{\pi' \in S^{(i,j)}_{(k,j+1)}: \\ \rho_l^{-1}(\pi') \neq null}} w(\pi') \exp\Big(-\beta\Big((j+1-l)^\alpha + (k-\pi(k))^\alpha - (k-j-1)^\alpha$$

$$- |l-\pi(k)|^\alpha\Big)\Big).$$

Define $a(j,k,l,\pi(k)) := \exp\Big(-\beta\Big((j+1-l)^\alpha+(k-\pi(k))^\alpha-(k-j-1)^\alpha-|l-\pi(k)|^\alpha\Big)\Big)$. Observe that $a(j,k,l,\pi(k))$ achieves its maximum when $\pi(k)$ achieves its highest value, i.e., $\pi(k) = j-1$. Now, substituting $\pi(k) = j-1$, the function $a(j,k,l,j-1)$ is decreasing in k, and since $k \geq j+1$,

$$a(j,k,l,\pi(k)) \leq a(j,j+1,l,j-1) = a(j,l), \tag{7}$$

where $a(j,l) := \exp\left(-\beta(j+1-l)^\alpha + \beta|j-1-l|^\alpha\right)$. As a result, we prove the claim:

$$w(S^{(i,j)}_{(l,j+1)}) \leq a(j,l) \sum_{k=j+1}^{n} \sum_{\substack{\pi' \in S^{(i,j)}_{(k,j+1)}: \\ \rho_l^{-1}(\pi') \neq null}} w(\pi') \leq a(j,l) w(S^{(i,j)}_{j^+,j+1}).$$

$\square$

C.1.3 Geometric Decay Bound (Proof of Lemma 3.2)

Lemma C.3. *From Propositions C.1 and C.2, it follows that for sufficiently large k, for all n and j,i satisfying $j - i \geq k$,*

$$\frac{P_n(i, j+1)}{P_n(i, j)} \leq C(\alpha, \beta) < 1,$$

where $C(\alpha, \beta) = \frac{e^{-\beta \alpha k^{\alpha-1}}}{\alpha-1} \frac{\Gamma\left(\frac{1}{\alpha-1}\right)}{(\beta \alpha)^{\frac{1}{\alpha-1}}}$ for $\alpha > 1$, and $C(1, \beta) = \frac{2e^{-2\beta}}{1+e^{-2\beta}}$

Proof. By applying Proposition C.1, for $j - i \geq k$,

$$\frac{P_n(i, j+1)}{P_n(i, j)} \leq e^{-\beta(k+1)^\alpha + \beta(k)^\alpha} \frac{\sum_{l \in [n]} e^{-\beta|j-l|^\alpha + \beta|j+1-l|^\alpha} w(S^{(i,j)}_{(l,j+1)})}{\sum_{l \in [n]} w(S^{(i,j)}_{(l,j+1)})} \tag{8}$$

$$\leq e^{-\beta(k+1)^\alpha + \beta(k)^\alpha} \frac{w(S^{(i,j)}_{(j^+,j+1)}) + \sum_{l=1}^{j} e^{-\beta|j-l|^\alpha + \beta|j+1-l|^\alpha} w(S^{(i,j)}_{(l,j+1)})}{w(S^{(i,j)}_{(j^+,j+1)}) + \sum_{l=1}^{j} w(S^{(i,j)}_{(l,j+1)})} \tag{9}$$

$$= e^{-\beta(k+1)^\alpha + \beta(k)^\alpha} \frac{1 + \sum_{l=1}^{j} e^{-\beta|j-l|^\alpha + \beta|j+1-l|^\alpha} \frac{w(S^{(i,j)}_{(l,j+1)})}{w(S^{(i,j)}_{(j^+,j+1)})}}{1 + \sum_{l=1}^{j} \frac{w(S^{(i,j)}_{(l,j+1)})}{w(S^{(i,j)}_{(j^+,j+1)})}}, \tag{10}$$

where in the second inequality we used the fact that for $l > j$, the coefficient $e^{-\beta|j-l|^\alpha + \beta|j+1-l|^\alpha}$ is less than 1.

Now, we can use the second mapping (Proposition C.2) to derive the desired upper bound. To see this, note that the coefficients in the numerator are greater than 1, in fact, for $l \leq j$, we have $exp(-\beta(j-l)^\alpha + \beta(j+1-l)^\alpha) \geq exp(\beta)$. Therefore, we can apply Proposition C.2,

$$\frac{P_n(i, j+1)}{P_n(i, j)} \leq e^{-\beta(k+1)^\alpha + \beta(k)^\alpha} \frac{1 + \sum_{l=1}^{j} e^{-\beta|j-l|^\alpha + \beta|j+1-l|^\alpha} a(j,l)}{1 + \sum_{l=1}^{j} a(j,l)},$$

where as before $a(j,l) = \exp\left(-\beta(j+1-l)^\alpha + \beta|j-1-l|^\alpha\right)$. We now consider two cases based on the value of α.

Case $\alpha > 1$: In this case, we can rewrite the numerator

$$\sum_{l=1}^{j} e^{-\beta|j-l|^\alpha + \beta|j+1-l|^\alpha} a(j,l) = e^\beta + \sum_{i=0}^{j-1} e^{-\beta((i+1)^\alpha - i^\alpha)} \le e^\beta + \sum_{i=0}^{j-1} e^{-\beta\alpha i^{\alpha-1}}.$$

For the denominator, we use the trivial lower bound of 1. Therefore,

$$\frac{P_n(i,j+1)}{P_n(i,j)} \le e^{-\beta(k+1)^\alpha + \beta(k)^\alpha} \left(1 + \left(e^\beta + \sum_{i=0}^{j-1} e^{-\beta\alpha i^{\alpha-1}}\right)\right).$$

For $\alpha > 1$, the following limit exists: $\lim_{j\to\infty} \sum_{i=0}^{j-1} e^{-\beta\alpha i^{\alpha-1}} = C_\infty(\alpha,\beta)$. Also, note that $C_\infty(\alpha,\beta) \le \int_0^\infty e^{-\beta\alpha x^{\alpha-1}} dx = \frac{1}{\alpha-1} \frac{\Gamma\left(\frac{1}{\alpha-1}\right)}{(\beta\alpha)^{\frac{1}{\alpha-1}}}$. Also note that $e^{-\beta((k+1)^\alpha - k^\alpha)} \le e^{-\beta\alpha k^{\alpha-1}}$. Therefore,

$$\frac{P_n(i,j+1)}{P_n(i,j)} \le e^{-\beta\alpha k^{\alpha-1}} C_\infty(\alpha,\beta).$$

So, if we choose k large enough we can see that $\frac{P_n(i,j+1)}{P_n(i,j)}$ is bounded by a constant smaller than 1. In fact,

$$\frac{P_n(i,j+1)}{P_n(i,j)} \le \frac{e^{-\beta\alpha k^{\alpha-1}}}{\alpha-1} \frac{\Gamma\left(\frac{1}{\alpha-1}\right)}{(\beta\alpha)^{\frac{1}{\alpha-1}}}.$$

Case $\alpha = 1$: By directly applying $\alpha = 1$ in (8), we get

$$\frac{P_n(i,j+1)}{P_n(i,j)} \le e^{-\beta} \frac{e^{-\beta} w(S_{(j^+,j+1)}^{(i,j)}) + e^\beta w(S_{(j^-,j+1)}^{(i,j)})}{w(S_{(j^+,j+1)}^{(i,j)}) + w(S_{(j^-,j+1)}^{(i,j)})}.$$

Now, note that for $\alpha = 1$, the value of $a(j,l)$ is equal to $exp(-2\beta)$. Therefore, $w(S_{(j^-,j+1)}^{(i,j)}) \le e^{-2\beta} w(S_{(j^+,j+1)}^{(i,j)})$. Thus,

$$\frac{P_n(i,j+1)}{P_n(i,j)} \le e^{-\beta} \left(e^{-\beta} + \frac{(e^\beta - e^{-\beta}) w(S_{(j^-,j+1)}^{(i,j)})}{w(S_{(j^+,j+1)}^{(i,j)}) + w(S_{(j^-,j+1)}^{(i,j)})}\right)$$

$$\le e^{-\beta}(e^{-\beta} + (e^\beta - e^{-\beta}) \frac{e^{-2\beta}}{1 + e^{-2\beta}})$$

$$= \frac{e^{-2\beta}(2 - e^{-2\beta})}{1 + e^{-2\beta}} \le \frac{2e^{-2\beta}}{1 + e^{-2\beta}} < 1$$

Finishing the proof of the lemma in both cases. □

C.2 Proof of Lemma 3.3

We start by using Lemma 3.2 to prove the following claim.

Claim Given the conditions of Theorem 3.1, there exists constants $c < 1$, $k > 0$ and γ such that for any $i \in [n]$,

$$\sum_{j:|j-i|>k} P_n(i,j) \le \gamma c^k.$$

Proof. Consider a fixed index i. Let $C(\alpha,\beta) < 1$ and k be the constants from Lemma 3.2 such that for any j satisfying $j - i > k$,

$$P_n(i,j) \le C(\alpha,\beta) P_n(i,j-1) \le C(\alpha,\beta)^{|j-i|-k} P_n(i,i+k).$$

A similar bound holds for $P_n(i,j)$ when $j < i - k$.

To bound the sum of probabilities where the deviation exceeds k, we split the summation into deviations to the right and deviations to the left of i:

$$\sum_{j:|j-i|>k} P_n(i,j) = \sum_{j>i+k} P_n(i,j) + \sum_{j<i-k} P_n(i,j).$$

Applying the geometric decay bound to each term, we obtain

$$\sum_{j:|j-i|>k} P_n(i,j) \leq \sum_{j>i+k} C(\alpha,\beta)^{j-i}\frac{P_n(i,i+k)}{C(\alpha,\beta)^k} + \sum_{j<i-k} C(\alpha,\beta)^{i-j}\frac{P_n(i,i-k)}{C(\alpha,\beta)^k}.$$

Recognizing that both sums are identical in form, we can combine them:

$$\sum_{j:|j-i|>k} P_n(i,j) \leq \left(\frac{P_n(i,i-k)}{C(\alpha,\beta)^k} + \frac{P_n(i,i+k)}{C(\alpha,\beta)^k}\right) \sum_{k=k+1}^{\infty} C(\alpha,\beta)^k.$$

The infinite geometric series $\sum_{k=k+1}^{\infty} C(\alpha,\beta)^k$ converges to $\frac{C(\alpha,\beta)^{k+1}}{1-C(\alpha,\beta)}$. Therefore, with $\gamma = 2\frac{P_n(i,i+k)}{C(\alpha,\beta)^k(1-C(\alpha,\beta))}$ and $c = C(\alpha,\beta)$, we have the proof of claim. $\qquad\square$

Now, we are ready to prove Lemma 3.3.

Proof of Lemma 3.3. We start with the first part. Let $S_n^{(k)}$ be the set of permutations that maps at least one element further than k, i.e., $S_n^{(k)} = \{\pi \in S_n : \exists i \in [n] \text{ s.t. } |i - \pi(i)| > k\}$. Note that for any $\pi \in S_n^{(k)}$, $\mathbb{P}_{\alpha,\beta}^{(k)}(\pi) = 0 \leq \mathbb{P}_{\alpha,\beta}(\pi)$, and for any $\pi \in S_n \setminus S_n^{(k)}$, $\mathbb{P}_{\alpha,\beta}^{(k)}(\pi) \geq \mathbb{P}_{\alpha,\beta}(\pi)$. Hence, the set of permutation that have a higher value over $\mathbb{P}$ is exactly $S_n^{(k)}$, i.e.,

$$\{\pi \in S_n : \mathbb{P}_{\alpha,\beta}(\pi) > \mathbb{P}_{\alpha,\beta}^{(k)}(\pi)\} = S_n^{(k)}$$

Therefore, by the definition of the total-variation distance,

$$\|\mathbb{P}_{\alpha,\beta} - \mathbb{P}_{\alpha,\beta}^{(k)}\|_{tv} = 2(\mathbb{P}_{\alpha,\beta}(S_n^{(k)}) - \mathbb{P}_{\alpha,\beta}^{(k)}(S_n^{(k)})) = 2\mathbb{P}_{\alpha,\beta}(S_n^{(k)})$$

Define $S_{n,i}^{(k)} = \{\pi \in S_n : |i - \pi(i)| > k\}$. Therefore, $S_n^{(k)} = \bigcup_i S_{n,i}^{(k)}$. Let c be the constants given by the claim above. Then if we choose $k = \frac{\log(2n/\epsilon)}{-\log(c)}$, by applying the result of the claim,

$$\mathbb{P}_{\alpha,\beta}(S_{n,i}^{(k)}) \leq \sum_{|j-i|>k} P_n(i,j) \leq \frac{\epsilon}{2n}$$

As a result,

$$\mathbb{P}_{\alpha,\beta}(S_n^{(k)}) \leq \sum_{i=1}^{n} \mathbb{P}_{\alpha,\beta}(S_{n,i}^{(k)}) \leq n\frac{\epsilon}{2n} = \frac{\epsilon}{2}$$

So, we conclude that $\|\mathbb{P}_{\alpha,\beta} - \mathbb{P}_{\alpha,\beta}^{(k)}\|_{tv} \leq \epsilon$.

Now, we prove the bound on the permanent. Recall the function $w : S_n \to \mathbb{R}^+$ is the weight of each permutation:

$$w(\pi) = \prod_{i=1}^{n} e^{-\beta|i-\pi(i)|^\alpha}.$$

Note that

$$per(A_n) - per(A_n^{(k)}) = \sum_{\pi \in S_n^{(k)}} w(\pi) = per(A_n)\mathbb{P}_{\alpha,\beta}(S_n^{(k)}).$$

By applying total variation bound,

$$\mathbb{P}_{\alpha,\beta}(S_n^{(k)}) \leq \sum_{i=1}^{n} \mathbb{P}_{\alpha,\beta}(S_{n,i}^{(k)}) \leq n\frac{\epsilon}{n} \leq \epsilon,$$

we get the second part of lemma:

$$\frac{per(A_n) - per(A_n^{(k)})}{per(A_n)} \leq \mathbb{P}_{\alpha,\beta}(S_n^{(k)}) \leq \epsilon$$

$$\square$$

C.3 Details of the Sampling Algorithm

In this section, we present the omitted details of how the sampling algorithm operates, explicitly using dynamic programming (DP).

DP states. We structure our DP algorithm into layers, where each layer corresponds to assigning a row to a column. By assigning row i to column j, we mean the permutation π is constructed such that $\pi(i) = j$. Specifically, we define DP states and transitions as follows:

- We have $n + 1$ layers indexed by $i = 0, \ldots, n$, representing stages in which rows $1, \ldots, n$ are sequentially assigned to columns.
- Each DP state at layer i (for $i = 0, 1, \ldots, n$) captures which columns are still available to assign at that stage. Due to the truncation, element i can only be assigned to columns within a bandwidth of $\pm k$ around position i. Thus, each DP state is represented succinctly by a binary vector of length $2k$, where a '0' indicates an available column and a '1' indicates an assigned column.
- The initial DP state at layer -1 is set to "$1\ldots10\ldots0$", a binary string with k ones followed by k zeros. We initialize $DP[-1]["1\ldots10\ldots0"] = 1$.

DP Transitions and Updates. Transitions between DP states from layer $i - 1$ to layer i occur by:

1. **Left-shifting** the binary vector from the state at layer $i - 1$, appending a new '0' to the right to indicate a new available column for row i.
2. **Flipping exactly one** of the '0's to '1', representing the assignment of the current element i to a specific column. Precisely, assigning element i to column j corresponds to flipping the bit indexed by $i - j + k$. The edge's weight for this assignment is $\exp(-\beta|i - j|^{\alpha})$.

Each DP state at layer i thus has exactly k bits set to '1'. The cumulative weight of reaching state s_{new} at layer i is updated as:

$$DP[i][s_{\text{new}}] \;=\; \sum_{s_{\text{old}}} DP[i-1][s_{\text{old}}] \;\times\; \big(A_n^{(k)}[i, \text{col\_selected}(s_{\text{old}} \to s_{\text{new}})]\big),$$

where the sum is over all valid predecessor states s_{old}. Here, $\text{col\_selected}(s_{\text{old}} \to s_{\text{new}})$ denotes the bit flipped during the transition.

Sequential Sampling from the DP Table. The sampling begins from the terminal state at layer n, denoted by the binary vector "$1\ldots10\ldots0$". At each step $i = n, n - 1, \ldots, 0$, we move backward through the DP table, selecting a predecessor state at layer $i - 1$ with probability proportional to the corresponding DP state weight. The bit flipped during this transition identifies the column assigned to element i. This sequential backward pass continues until reaching the initial state at layer -1, thus completing the permutation.

The final permutation sampled this way precisely follows the truncated distribution $\hat{\mathbb{P}}_n^{(k)}$. Algorithms explicitly detailing these DP state constructions and sampling procedures are provided in Algorithms 2 and 3. Note the DP state space size at each layer is $\binom{2k}{k}$, with at most k predecessors per state. Therefore, the forward DP computation and backward sampling both run in time $O\left(nk\binom{2k}{k}\right)$.

C.4 Proof of Lemma 3.4

Proof. We proceed by induction on the layers, and we show that $DP[i][s]$ is equal to the cumulative weights of all partial assignments up to i given the availability of assignments corresponding to s.

Base case $i = 0$ holds by definition, since $DP[0][s_0] = 1$. Assume that for some $i \geq 0$, the DP table correctly stores the cumulative weights of all partial assignments up to layer i for every state s in layer i. That is, for each state s in layer i,

$$DP[i][s] = \sum_{\pi \in S_i(s)} \prod_{j=1}^{i} A_n^{(k)}[j, \pi(j)],$$

where $S_i(s)$ denotes the set of all partial permutations up to layer i that result in state s.

For the induction step, we need to show that the DP table correctly computes $DP[i+1][s']$ for each state s' in layer $i+1$. Consider a state s' in layer $i+1$. Let s be the predecessor state in layer i, and let j be the column to which row $i+1$ is assigned, corresponding to the flipped bit in s'. The cumulative weight for state s' is updated as:

$$DP[i+1][s'] = \sum_{s \in \text{Pred}(s')} DP[i][s] \times A_n^{(k)}[i+1, j],$$

where $\text{Pred}(s')$ denotes all predecessor states s that can transition to s'.

By induction hypothesis:

$$DP[i][s] = \sum_{\pi \in S_i(s)} \prod_{j=1}^{i} A_n^{(k)}[j, \pi(j)].$$

Substituting into the previous equation:

$$DP[i+1][s'] = \sum_{s \in \text{Pred}(s')} \left(\sum_{\pi \in S_i(s)} \prod_{j=1}^{i} A_n^{(k)}[j, \pi(j)] \right) \times A_n^{(k)}[i+1, j].$$

This can be rewritten as:

$$DP[i+1][s'] = \sum_{s \in \text{Pred}(s')} \sum_{\pi \in S_i(s)} \prod_{j=1}^{i+1} A_n^{(k)}[j, \pi(j)],$$

which is precisely:

$$DP[i+1][s'] = \sum_{\pi \in S_{i+1}(s')} \prod_{j=1}^{i+1} A_n^{(k)}[j, \pi(j)],$$

where $S_{i+1}(s')$ denotes the set of all partial permutations up to layer $i+1$ that result in state s'. Thus, the DP table correctly accumulates the weights for all states in layer $i+1$.

After processing all n layers, the final layer n contains the terminal state $s_n = \text{``}1 \cdots 1\, 0 \cdots 0\text{''}$. The value $DP[n][s_n]$ equals $\text{per}(A_n^{(k)})$, as it sums over all valid permutations $\pi \in S_n$ weighted by $\prod_{j=1}^{n} A_n^{(k)}[j, \pi(j)]$.

Sampling Correctness: The sampling procedure performs a backward traversal from the terminal state s_n to the initial state s_0. At each step $i = n, n-1, \ldots, 1$, the algorithm selects a predecessor state s with probability proportional to the weight of the edge connecting s to s'. Formally, the probability of transitioning from s to s' is:

$$P(s \to s') = \frac{DP[i-1][s'] A_n^{(k)}[i, j]}{DP[i][s]},$$

where j is the column assigned at this transition. By the inductive construction of the DP table, these transition probabilities ensure that the probability of sampling a path corresponding to π is:

$$\prod_{i=0}^{n} \frac{DP[i-1][s'] A_n^{(k)}[i, \pi(i)]}{DP[i][s]} = \frac{\prod_{j=1}^{n} A_n^{(k)}[j, \pi(j)]}{DP[n][s_0]} = \frac{\prod_{j=1}^{n} A_n^{(k)}[j, \pi(j)]}{\text{per}(A_n^{(k)})} = \hat{\mathbb{P}}_{\alpha,\beta}^{(k)}(\pi).$$

Therefore, the algorithm correctly samples permutations according to $\hat{\mathbb{P}}_{\alpha,\beta}^{(k)}$.

Time Complexity: The algorithm involves two main phases: building the DP table and performing the backward sampling. There are n layers and $\binom{2k}{k}$ nodes per layer. Each state has at most k predecessors (since exactly one '0' is flipped to '1' within a window of $2k$ positions), and each transition involved constant time multiplication and addition. Therefore, the total time to build the DP table is:

$$O\left(n \cdot \binom{2k}{k} \cdot k \right).$$

Using the approximation $\binom{2k}{k} \leq \frac{4^{(k)}}{\sqrt{\pi k}}$ (from Stirling's formula), the time complexity becomes:

$$O\left(n \cdot \frac{4^k}{\sqrt{k}} \cdot k\right) = O\left(n \cdot 4^k \cdot \sqrt{k}\right).$$

After the DP is constructed, Algorithm 3 generates samples from $\hat{\mathbb{P}}_{\alpha,\beta}^{(k)}$ in time $O(nk)$. Using Lemma 3.3, to achieve ϵ total variation distance with true Mallows we need to set $k = \frac{\log(n/\epsilon)}{\log(1/C(\alpha,\beta))}$ where $C(\alpha,\beta) < \frac{2exp(-2\beta)}{1+\exp(-2\beta)}$ is given by Lemma 3.2. So the runtime of computing permanent and preprocessing for sampling is with ϵ error is $O\left(n^{1-\frac{2}{\log(C(\alpha,\beta))}} \epsilon^{\frac{2}{\log(C(\alpha,\beta))}} \log(n/\epsilon)\right)$ $\qquad\square$

C.5 Pseudo-code: Sampling From Mallows Model

Algorithm 2: Build DP Table for Sampling

Input : $n, k \in \mathbb{N}$; Truncated matrix $A_n^{(k)}$ with bandwidth $2k$.
Output : DP States

$DP \leftarrow$ array of dictionaries indexed by $\ell = -1, 0, \ldots, n$;
Define $\texttt{init\_state} \leftarrow \underbrace{11\cdots1}_{k}\underbrace{00\cdots0}_{k}$;
$DP[-1][\texttt{init\_state}] \leftarrow \varnothing$;   // Initialization layer with one node.
for $\ell \leftarrow 0$ **to** n **do**
    foreach *state* s_{old} *in* $DP[\ell-1]$ **do**
        $s_{\text{shift}} \leftarrow s_{\text{old}}[1:] + 0$;   // Remove leftmost bit and append a '0'.
        ;
        for $i \leftarrow 0$ **to** $2k-1$ **do**
            if $s_{shift}[i] = 0$ **then**
                $s_{\text{new}} \leftarrow \text{FlipBit}(s_{\text{shift}}, i)$;
                $j \leftarrow \ell - (i - k)$;   // Column j assigned to row ℓ.
                $w_{\text{edge}} \leftarrow A_n^{(k)}[\ell, j]$;   // Edge weight.
                if $s_{new} \notin DP[\ell]$ **then**
                    $DP[\ell][s_{\text{new}}] \leftarrow \varnothing$;
                Append $(s_{\text{new}}, w_{\text{edge}})$ to $DP[\ell-1][s_{\text{old}}]$;
return DP

Algorithm 3: Sample a Permutation from the DP Table

Input : DP table $DP[\ell][s]$ containing cumulative weights.
Output: Sampled permutation $\pi \in S_n$.

Initialize $\pi[1..n]$ as an empty array;
$s_{\text{current}} \leftarrow \underbrace{11\cdots1}_{k}\underbrace{00\cdots0}_{k}$;   // Start from final state at layer n.

for $\ell \leftarrow n$ **to** 1 **do**

    $predList \leftarrow \varnothing$;   // Collect predecessors of s_{current}.
    foreach (s_{old}, w_{edge}) *in* $DP[\ell-1][\cdot]$ **do**
        if $(s_{old} \rightarrow s_{current})$ *is valid* **then**
            Let $j \leftarrow$ column assigned in this transition;
            $score \leftarrow DP[\ell-1][s_{\text{old}}] \times w_{\text{edge}}$;
            Append $(s_{\text{old}}, j, score)$ to $predList$;

    Normalize scores in $predList$ to form a probability distribution;
    Randomly select $(s_{\text{old}}, j, score)$ w.r.t. these probabilities;
    $\pi[\ell] \leftarrow j$;   // Assign column j to row ℓ.
    $s_{\text{current}} \leftarrow s_{\text{old}}$;

return π

D   Additional Implementation Details

Evaluation metrics. We evaluate predictive accuracy using the following metrics, comparing the test ranking π and the predicted ranking σ:

1. **Top-1/Top-5 Hit Rate:** Probability the top-ranked item in π appears in the top-1 or top-5 positions of σ:
$$\text{Top-}k \text{ Hit rate} = \Pr\left(\pi(1) \in \{\sigma(1), \cdots, \sigma(k)\}\right). \tag{11}$$

2. **Spearman's ρ:** Correlation based on squared rank differences:
$$\rho = 1 - \frac{6\sum_{i=1}^{n}(\pi(i) - \sigma(i))^2}{n(n^2 - 1)}.$$

3. **Kendall's τ:** Correlation based on pairwise agreements as defined by d_τ in Section 4.

4. **Hamming Distance:** Fraction of positions with differing ranks:
$$H = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\{\pi(i) \neq \sigma(i)\}.$$

5. **Pairwise Accuracy:** Fraction of item pairs with matching relative orders:
$$\text{Pairwise Accuracy} = \frac{\sum_{i<j} \mathbf{1}\{(\pi(i) - \pi(j))(\sigma(i) - \sigma(j)) > 0\}}{\binom{n}{2}}.$$

In practice, all evaluation metrics are computed using Monte Carlo methods. We generate 100 predicted ranking samples from each trained model and estimate the metrics empirically. These empirical estimates approximate the true population-level metrics.

Baseline Implementation Details As mentioned in 4.1, we compare against the Plackett-Luce (PL) and Mallow's τ model.

We fit the PL model via maximum likelihood estimation:
$$\mathcal{L}(\theta; \pi^{(1)}, \ldots, \pi^{(m)}) = \sum_{\ell=1}^{m}\sum_{i=1}^{n}\left[\theta_{\pi^{(\ell)}(i)} - \log\left(\sum_{j=i}^{n}\exp(\theta_{\pi^{(\ell)}(j)})\right)\right]. \tag{12}$$

We optimize the negative log-likelihood using the L-BFGS algorithm [25]. Additionally, we provide a fast sampler for the PL model using the Gumbel–Max [49, 28] to generate synthetic permutations:

$$\pi = \arg sort\left(-(\theta + \epsilon)\right), \quad \epsilon \sim \text{Gumbel}(0,1)^n, \tag{13}$$

where sorting is done in descending order to reflect best-to-worst preferences.

To fit the Mallows τ model, note that its normalizer has a closed form [16]:

$$Z(\beta) = \prod_{j=1}^{n} \frac{1 - e^{-j\beta}}{1 - e^{-\beta}}.$$

Again, we do maximum likelihood estimation where the log-likelihood of the Mallows τ model is:

$$\mathcal{L}(\sigma, \beta; \pi^{(1)}, \ldots, \pi^{(m)}) = -\beta \sum_{\ell=1}^{m} d_\tau(\pi^{(\ell)}, \sigma) - m \log Z(\beta).$$

We perform maximum likelihood estimation by:

- Setting the central ranking σ_0 to the Borda count aggregation over training rankings.
- Optimizing β using numerical minimization of the negative log-likelihood.

To generate samples from the Mallows–τ model, we use the exponential insertion algorithm by [26]. At each step, an item from σ_0 is inserted into a position within the partial permutation with probability proportional to $\exp(-\beta \cdot \text{reverse\_index})$, favoring positions near the end. This yields efficient forward sampling in $O(n^2)$ time.