Near-Optimal Clustering in Mixture of Markov Chains

Junghyun Lee
KAIST AI
jh_lee00@kaist.ac.kr

Yassir Jedra
ICL EEE
y.jedra@imperial.ac.uk

Alexandre Proutière
KTH EECS, Digital Futures
alepro@kth.se

Se-Young Yun
KAIST AI
yunseyoung@kaist.ac.kr

Abstract

We study the problem of clustering T trajectories of length H, each generated by one of K unknown ergodic Markov chains over a finite state space of size S. We derive an instance-dependent, high-probability lower bound on the clustering error rate, governed by the stationary-weighted KL divergence between transition kernels. We then propose a two-stage algorithm: Stage I applies spectral clustering via a new injective Euclidean embedding for ergodic Markov chains, a contribution of independent interest enabling sharp concentration results; Stage II refines clusters with a single likelihood-based reassignment step. We prove that our algorithm achieves near-optimal clustering error with high probability under reasonable requirements on T and H. Preliminary experiments support our approach, and we conclude with discussions of its limitations and extensions.

1 INTRODUCTION

Clustering, or community detection for graphs, is a fundamental statistical problem with applications across diverse scientific disciplines, including social science, biology, and statistical physics (McLachlan et al., 2019; Kiselev et al., 2019; Ezugwu et al., 2022; Fortunato, 2010). The precise statistical characterization of clustering has been rigorously investigated under various probabilistic frameworks, such as stochastic block models (SBMs, Abbe (2018)), Gaussian mixture models (GMMs, Lu and Zhou (2016); Löffler et al. (2021); Chen and Zhang (2024)), and block Markov chains (BMCs, Sanders et al. (2020); Jedra et al. (2023)).

Proceedings of the 29th International Conference on Artificial Intelligence and Statistics (AISTATS) 2026, Tangier, Morocco. PMLR: Volume 300. Copyright 2026 by the author(s).

However, the aforementioned models primarily focus on clustering static data points or nodes, where individual elements often lack inherent informative structure in isolation: the clustering signal typically emerges only when considering the entire data. In contrast, many real-world applications involving multiple underlying processes, such as astronomy (Yang et al., 2020), mobile social networks (Tang et al., 2021), and human activity patterns (Zhou et al., 2021), deal with *trajectories* with a mixture of temporal information. The task is to cluster the trajectories based on their generating process or model. The longer each trajectory (e.g., a user's interaction sequence or a time series) is, the more information it potentially reveals about its generating model, facilitating clustering.

The **Mixture of Markov Chains (MMC)** provides a fundamental yet powerful probabilistic framework for the problem of trajectory clustering in the most basic scenario with no controllable actions. In this model, we are given T trajectories of length H, where each trajectory is generated by one of K unknown Markov chains defined over a finite state space of size S. This model has been studied in a variety of contexts, originally by Blumen et al. (1955) for modeling heterogeneous labor mobility patterns, and later utilized for learning underlying usage patterns from user trails of app usages, music playlists, web browsing, and more (Gupta et al., 2016; Spaeh and Tsourakakis, 2023; Maystre et al., 2022; Cadez et al., 2003; Poulsen, 1990; Girolami and Kabán, 2003; Zhou et al., 2021).

In **MMC**, two primary and closely intertwined objectives emerge: *learning* the underlying K Markov chain models (Gupta et al., 2016; Spaeh and Tsourakakis, 2023; Kausik et al., 2023) and *clustering* the observed trajectories according to their generative source (Kausik et al., 2023). Successful clustering can significantly simplify the learning task by allowing for information aggregation within each identified group of trajectories. Conversely, while learning each model accurately can indeed facilitate clustering, it is expected to necessitate stringent requirements on the trajectory length H, because each trajectory is solely responsible

for its own amount of information helpful for clustering. Thus, beyond its intrinsic value as an exploratory data analysis technique, accurate clustering can lead to statistically more efficient inference than analyzing each trajectory independently.

As done in prior clustering literature, one can ask two critical questions: *(i) What is the fundamental limit on the misclassification rate? (ii) Given sufficiently large (but not excessively so) trajectories, is there a computationally tractable algorithm whose performance (clustering error rate) matches the lower bound? What is the requirement on T and H?*

Several studies have algorithmically addressed the second question (Gupta et al., 2016; Kausik et al., 2023; Spaeh and Tsourakakis, 2023, 2024; Spaeh et al., 2024), often without rigorous statistical guarantees. One notable exception is Kausik et al. (2023), where the authors proposed a clustering algorithm that provably performs *exact* clustering (i.e., zero misclassification error) in **MMC** when $T = \widetilde{\Omega}(K^2 S)$ and $H = \widetilde{\Omega}(K^{3/2} t_{\text{mix}})$. However, they do not provide the clustering error rate across various regimes of T, H and appropriate separation parameters. Perhaps more importantly, they lack a corresponding lower bound and related discussions, making it difficult to assess the optimality of their algorithm. Furthermore, their algorithm relies on explicit knowledge of problem-specific quantities, often unavailable in practice. We will elaborate on this comparison in Section 5. Consequently, despite their importance, both of the aforementioned questions have remained elusive in the literature.

Contributions. In short, we answer both questions fully. Specifically: *(a)* We prove an instance-specific high-probability lower bound on the clustering error rate for **MMC**. This reveals the problem-difficulty quantity $\mathcal{D}$: the minimum weighted KL divergence between the transition kernels (Section 3). *(b)* We propose a two-stage clustering algorithm that achieves near-optimal clustering error. Notably, it does not require any *a priori* knowledge of the underlying model, yet fully adapts to the given problem difficulty (Section 4). Especially for Stage I, we introduce a new injective Euclidean embedding (L-embedding, Section 4.1) specifically designed for ergodic Markov chains. This embedding, a contribution of independent interest, facilitates sharp concentration results for spectral clustering analysis (Section 4.1). *(c)* Our upper and lower bounds reveal gaps in misclassification errors and the required trajectory length H. Building on recent advances in concentration inequalities (Paulin, 2015; Fan et al., 2021) and estimation techniques (Wolfer and Kontorovich, 2021) for Markov chains, we elucidate the inherent complexities of clustering in **MMC** that currently render these gaps unavoidable (Section 5).

Notation. For a positive integer $n \geq 1$, let $[n] := \{1, 2, \cdots, n\}$. For a set X, let $\Delta(X)$ be the set of probability distributions over X. Let $a \vee b := \max\{a, b\}$ and $a \wedge b := \min\{a, b\}$. We will freely utilize the asymptotic notations $\mathcal{O}, o, \Omega, \omega, \Theta$, and for aesthetic purpose, we will also use $f \gtrsim g$, $f \lesssim g$, $f \asymp g$, defined as $f = \Omega(g)$, $f = \mathcal{O}(g)$, $f = \Theta(g)$, respectively; log in the subscript (e.g., $\lesssim_{\log}$) indicates up to logarithmic factors, corresponding to $\widetilde{\Omega}$, $\widetilde{\mathcal{O}}$, and $\widetilde{\Theta}$.

2 PROBLEM SETTING

Mixture of Markov chains (MMCs). There are K unknown Markov chains. The k-th Markov chain is denoted as $\mathcal{M}^{(k)} = (\mathcal{S}, H, \mu^{(k)}, p^{(k)})$. $\mathcal{S}$ is a finite state space of cardinality S, $H \geq 2$ is the horizon or (episode length), $\mu^{(k)}$ is a initial state distribution, and $p^{(k)}(\cdot|\cdot)$ and $\boldsymbol{P}^{(k)}$ are the transition kernel and matrix.

We now recall an essential concept for spectral analysis of Markov chains. For an ergodic Markov chain with transition matrix $\boldsymbol{P}$, its **pseudo-spectral gap** is defined as $\gamma_{\text{ps}} := \max_{m \geq 1} \frac{1}{m} \gamma\left((\boldsymbol{P}^*)^m \boldsymbol{P}^m\right)$, where $\gamma(\cdot)$ is the spectral gap of the self-adjoint operator (Paulin, 2015, Section 3). With this, we assume the following:

Assumption 1. *Each $\mathcal{M}^{(k)}$ is ergodic,*[1] *with pseudo-spectral gap $\gamma_{\text{ps}}^{(k)} > 0$ and stationary distribution $\pi^{(k)}$. Also, $\gamma_{\text{ps}} := \min_k \gamma_{\text{ps}}^{(k)}$ and $\pi_{\min} := \min_{k,s} \pi^{(k)}(s) > 0$.*

Learner's Objective. We first clarify that the learner knows $\mathcal{S}$, and for the simplicity of exposition, we also assume that the learner knows *either* the number of clusters K *or* the minimum pseudo-spectral gap γ_{ps}.[2] With this, the learner observes T trajectories generated as follows: for each $t \in [T]$, a trajectory of length H is sampled from $\mathcal{M}^{(f(t))}$. Specifically, trajectory $\mathcal{T}_t = (s_{t,1}, \ldots, s_{t,H})$ is generated as $s_{t,1} \sim \mu^{(f(t))}$ and $s_{t,h+1} \sim p^{(f(t))}(\cdot|s_{t,h})$, which we denote as $\mathcal{T}_t \sim \mathcal{M}^{(f(t))}$. $f : [T] \to [K]$ is the unknown, ground-truth decoding function that maps each trajectory to its generating Markov chain.

The learner's objective is to recover f—that is, to cluster the T trajectories according to their generating models. Let $\widehat{f} : [T] \to [K]$ denote the learner's estimated decoding function. The number of misclassified

[1] We assume uniform ergodicity: there exist $M > 0, \rho \in (0, 1)$ such that $\max_{s \in \mathcal{S}} \mathsf{TV}(P^H(s, \cdot), \pi) \leq M\rho^H$ for all $H \in \mathbb{N}$, which is implied by aperiodicity and irreducibility (Levin and Peres, 2017, Theorem 4.9).

[2] A fully "parameter-free" algorithm would require neither. In fact, by leveraging recent advances in tight estimation of γ_{ps} (Wolfer and Kontorovich, 2024), our algorithm can be made parameter-free through an additional initialization step, as we elaborate below Theorem 4.1.

trajectories is measured as:

$$E_T(\widehat{f}, f) := \min_{\sigma \in \mathrm{Sym}(K)} \sum_{t=1}^{T} \mathbb{1}[\widehat{f}(t) \neq \sigma(f(t))],$$

where $\mathrm{Sym}(K)$ is the symmetric group over $[K]$. The goal is to design an algorithm such that, for any confidence level $\delta \in (0,1)$ and misclassification rate $\varepsilon \in (0,1)$, the following guarantee holds: $\mathbb{P}(E_T(\widehat{f}, f) \leq \varepsilon T) \geq 1 - \delta$, where the probability $\mathbb{P}$ is over the randomness of both the trajectory generation and the clustering algorithm. We note in advance that the error rate ε and the failure probability δ may both depend on factors such as T, H, S, K, and the mixing properties of the Markov chains.

3 FUNDAMENTAL LOWER BOUND

We define an instance of **clustering in MMC** as a tuple $\Phi_T := ((\mathcal{M}^{(k)})_{k\in[K]}, f, T)$, where $(\mathcal{M}^{(k)})_{k\in[K]}$ is a collection of K ergodic Markov chains, f is the decoding function, and $T \in \mathbb{N}$ is the total number of trajectories to be clustered.

To derive instance-specific lower bounds, we must consider algorithms that for a given instance, genuinely adapt to perturbations of the decoding function. Indeed, an algorithm that would return $\widehat{f} = f$, would have no misclassified trajectories, but would fail for other decoding functions. Hence, we now introduce a class of "good" algorithms that are robust to small instance perturbations with a given confidence level. Define $\boldsymbol{\alpha}(f) := (\alpha_k(f))_{k\in[K]} \in \Delta([K])$ with $\alpha_k(f) := |f^{-1}(k)|/T$.

Definition 3.1: Stable Algorithms

Let $(\varepsilon, \beta, \delta) \in [0,1] \times \mathbb{R}_{\geq 0} \times (0, 1/2]$. A clustering algorithm $\mathcal{A}$ is $(\varepsilon, \beta, \delta)$**-locally stable** at $\Phi_T := ((\mathcal{M}^{(k)})_{k\in[K]}, f, T)$ if the following holds: for all $\Phi'_T = ((\mathcal{M}^{(k)})_{k\in[K]}, f', T)$ such that $\|\boldsymbol{\alpha}(f') - \boldsymbol{\alpha}(f)\|_2 \leq \beta$,

$$\mathbb{P}_{\Phi'_T, \mathcal{A}}\left(E_T(\widehat{f}_{\mathcal{A}}, f') > \varepsilon T\right) \leq \delta, \qquad (1)$$

where $\widehat{f}_{\mathcal{A}}$ is the outputted clustering function from our algorithm $\mathcal{A}$ that takes as input the T trajectories whose ground-truth clustering is given by f'. Furthermore, we say that $\mathcal{A}$ is (ε, β)**-asymptotically locally stable** at $\Phi := ((\mathcal{M}^{(k)})_{k\in[K]}, f)$ if there exists a sequence $(\delta_T)_{T\in\mathbb{N}}$ with $\lim_{T\to\infty} \delta_T = 0$ such that $\mathcal{A}$ is $(\varepsilon, \beta, \delta_T)$-locally stable at Φ_T for all $T \in \mathbb{N}$.

Intuitively, perturbations are (slight) changes in the relative cluster sizes $\boldsymbol{\alpha}(f)$, inspired by the prior definitions of locally stable algorithms for clustering in SBMs (Yun and Proutière, 2019, Definition 1), block Markov chains (Sanders et al., 2020), and recently, block Markov Decision Processes (Jedra et al., 2023).

We now present our *instance-specific, high-probability* lower bound for clustering in mixture of Markov chains:

Theorem 3.1: Error Rate Lower Bound

Let $(\varepsilon, \delta) \in [0,1] \times (0, 1/2]$. Then, a necessary condition for the existence of a $(\varepsilon, \beta, \delta)$-locally stable algorithm at $\Phi_T := ((\mathcal{M}^{(k)})_{k\in[K]}, f, T)$ with $\beta \geq 2\sqrt{2}\varepsilon$ is as follows: denoting $\alpha_{\min} = \min_{k\in[K]} \alpha_k(f)$,

$$\delta \geq \frac{1}{2}\left(\frac{\alpha_{\min}}{16e\varepsilon}\right)^{\varepsilon T} \exp\left(-4\varepsilon T(H-1)\mathcal{D}\right)$$
$$\iff 4(H-1)\mathcal{D} \geq \frac{1}{\varepsilon T}\log\frac{1}{2\delta} + \log\frac{\alpha_{\min}}{16e\varepsilon}. \qquad (2)$$

For (ε, β)-asymptotically locally stable algorithm with the same β as above and $\varepsilon = o(1)$, we have the following necessary condition:

$$\liminf_{T\to\infty} \frac{2(H-1)\mathcal{D}}{\log \frac{\alpha_{\min}}{16e\varepsilon}} \geq 1. \qquad (3)$$

The **information-theoretic divergence** $\mathcal{D}$ is defined as $\mathcal{D} := \min_{k\neq k'\in[K]} \mathcal{D}^{(k,k')}$, where

$$\mathcal{D}^{(k,k')} := \frac{1}{H-1}\mathsf{KL}\left(\mu^{(k)}, \mu^{(k')}\right) + \sum_{s\in\mathcal{S}} \mathbb{P}_H^{(k)}(s)\,\mathsf{KL}\left(p^{(k)}(\cdot|s), p^{(k')}(\cdot|s)\right),$$

where $\mathbb{P}_H^{(k)}(s) := \frac{1}{H-1}\sum_{h=1}^{H-1} \mathbb{P}^{(k)}(s_h = s)$.

Proof sketch. The proof proceeds by first constructing a large number of hypotheses, each corresponding to different cluster allocations of the observations, following the combinatorial arguments of Yun and Proutière (2019). These hypotheses are slight perturbations of the given instance of clustering in MMC. We then apply the change-of-measure argument (Lai and Robbins, 1985) and the data-processing inequality (Garivier et al., 2019, Lemma 1) to relate the error of any "good" clustering algorithm to the KL divergences between true and alternate models. Finally, we optimize the resulting lower bound over the different allocations. The approach is inspired by Yun and Proutière (2019, Theorem 1) and Jedra and Proutière (2023, Theorem 2), which establish high-probability lower bounds for clustering in SBMs and for linear system identification, respectively. The full proof is deferred to Appendix A. □

Remark 1 (Chernoff Information)**.** *Dreveton et al. (2024) recently highlighted Chernoff information as a universal measure of clustering difficulty for sub-exponential mixtures, extending earlier results in SBM clustering (Abbe and Sandon, 2015a,b; Dreveton et al., 2023). While we use* $\mathsf{KL}(\cdot,\cdot)$*, the two quantities are closely related (van Erven and Harremos, 2014) and become equivalent as* $T \to \infty$ *in many settings, including labeled SBMs (Yun and Proutière, 2016, Claim 4).*

Remark 2 (Comparison to Wang and Choi (2023))**.** *One notable related result is Wang and Choi (2023, Theorem 4.1), which establishes an asymptotic equipartition property (Cover and Thomas, 2006, Theorem 11.8.1) of binary hypothesis testing between two Markov chains from* $\{(X_{h,1}, X_{h,2})\}_{h\in[H]}$*, where* $X_{h,1} \overset{i.i.d.}{\sim} \nu$ *and* $X_{h,2} \sim P(\cdot|X_{h,1})$*. Our asymptotic lower bound is a generalization of theirs, as we consider the full Markov chain instead of a collection of single hops.*

Asymptotically Exact Recovery. The necessary condition for asymptotically exact recovery ($\varepsilon = \mathcal{O}(1/T)$) is $H\mathcal{D} = \Omega(\log T)$. This mirrors the minimax lower bound for GMM clustering (Lu and Zhou, 2016, Theorem 3.3), where strong consistency demands that the squared signal-to-noise ratio (SNR) is at least logarithmic w.r.t. the number of data points n: $\Delta^2/\sigma^2 = \Omega(\log n)$, with Δ denoting the minimum ℓ_2-separation of the mean vectors and σ^2 the isotropic variance of the GMM. Analogously, in our setting, $\mathcal{D}$ corresponds to Δ^2, and H plays a role similar to $1/\sigma^2$: longer trajectories reduce variance, much like smaller σ^2 does in the GMM.

Stationary Divergence. Without loss of generality, suppose that $\max_{k,k'} \mathsf{KL}(\mu^{(k)}, \mu^{(k')}) < \infty$. Then, under Assumptions 1 (ergodicity), by the Ergodic Theorem for Markov chains (Levin and Peres, 2017, Appendix C)), $\mathcal{D}^{(k,k')} \to \mathcal{D}_\pi^{(k,k')}$ almost surely as $H \to \infty$, where

$$\mathcal{D}_\pi^{(k,k')} := \sum_{s\in\mathcal{S}} \pi^{(k)}(s)\, \mathsf{KL}\left(p^{(k)}(\cdot|s), p^{(k')}(\cdot|s)\right). \quad (4)$$

The **stationary (information-theoretic) divergence** is then defined as $\mathcal{D}_\pi := \min_{k\neq k'} \mathcal{D}_\pi^{(k,k')}$, which quantifies the difficulty of clustering in MMC in the stationary regime. It aggregates the "local" difficulty at each state s—quantified by the KL-divergence between outgoing transition probabilities—weighted by the long-run frequency $\pi^{(k)}(s)$ of visiting that state. Thus, a state's contribution to distinguishability is small if *either* the chains behave similarly from that state or the state is rarely visited. This aggregation is information-geometrically optimal, as it weights the KL divergence – the optimal statistic for local hypothesis testing under the Neyman-Pearson lemma (Neyman and Pearson, 1933) – by the stationary distribution.

4 TWO-STAGE ALGORITHM

Motivated by prior clustering literature (Gao et al., 2017; Yun and Proutière, 2016; Lu and Zhou, 2016), our algorithm consists of two stages. The full pseudocodes are provided in Algorithms 1 and 2.

4.1 Stage I. Initial Spectral Clustering

L-Embedding. We first introduce a new Euclidean embedding of ergodic Markov chains:

Definition 4.1: L-Embedding

For any ergodic Markov chain $\mathcal{M}$ with transition probability matrix $\boldsymbol{P}$ and stationary distribution π, its **L-embedding** and its empirical version $\widehat{L}$ are defined as

$$L : \mathcal{M} \mapsto \mathrm{vec}\left(\mathrm{diag}(\pi)^{1/2}\boldsymbol{P}\right),$$

$$\widehat{L} : \mathcal{T} = (s_1, s_2, \cdots, s_H) \mapsto \left(\frac{\widehat{N}(s,s')}{\sqrt{H\widehat{N}(s)}}\right)_{s,s'\in\mathcal{S}},$$

where $\mathcal{T} \sim \mathcal{M}$, $\widehat{N}(s) := \sum_{h=1}^{H} \mathbb{1}[s_h = s]$, and $\widehat{N}(s,s') := \sum_{h=1}^{H-1} \mathbb{1}[s_h = s, s_{h+1} = s']$.

With this, we define the ground-truth and empirical data matrices of size $T \times S^2$ by row-concatenating the L-embeddings as follows:

$$\boldsymbol{W} = \left[L(\mathcal{M}^{(f(t))})\right]_{t\in[T],:}, \quad \widehat{\boldsymbol{W}} = \left[\widehat{L}(\mathcal{T}_t)\right]_{t\in[T],:}. \quad (5)$$

Remark 3. *To the best of our knowledge, such a Euclidean embedding for ergodic Markov chains has not been previously reported in the literature. In Appendix G, we discuss its connections and distinctions to existing spectral-related representations, including diffusion maps (Coifman and Lafon, 2006), the (weighted) Laplacian (Chung, 1997), the symmetrized form (Levin and Peres, 2017), and doublet frequencies (Wolfer and Kontorovich, 2021; Vidyasagar, 2014).*

We now discuss three desirable properties of the L-embedding that make it suitable for our purpose.

First, our $L(\cdot)$ uniquely determines $\mathcal{M}$, up to the initial distribution:

Proposition 4.1. *Over the space of ergodic Markov chains, L is an injective mapping up to the initial distribution: for $\mathcal{M} = (\mathcal{S}, H, \mu, p)$ and $\mathcal{M}' = (\mathcal{S}, H, \mu', p')$, $L(\mathcal{M}) = L(\mathcal{M}') \iff p = p'$.*

Proof. By definition, we must have that for all $s, s' \in \mathcal{S}$, $\sqrt{\pi(s)}p(s'|s) = \sqrt{\pi'(s)}p'(s'|s)$. Summing over $s' \in \mathcal{S}$,

we have that $\pi(s) = \pi'(s) > 0$ for all $s \in \mathcal{S}$, where the ergodicity of the chains implies strict positivity. This then implies that $p = p'$. □

Second, the L-embedding possesses desirable concentration properties. The empirical mapping $\widehat{L}$ normalizes the transition counts by $\sqrt{\widehat{N}(s)}$. To understand why this is advantageous, note that we can rewrite each entry of the empirical L-embedding as

$$\frac{\widehat{N}(s,s')}{\sqrt{H\widehat{N}(s)}} = \underbrace{\frac{\widehat{N}(s,s')}{\widehat{N}(s)}}_{\approx p(s'|s)} \underbrace{\sqrt{\frac{\widehat{N}(s)}{H}}}_{\approx\sqrt{\pi(s)}}. \tag{6}$$

Because the concentration bound of the empirical transition vector for state s, $\left(\widehat{N}(s,s')/\widehat{N}(s)\right)_{s'\in\mathcal{S}}$, scales as $\tilde{\mathcal{O}}(1/\sqrt{H\pi(s)})$ (see Eqn. (20) in Appendix B), multiplying by $\sqrt{\pi(s)}$—which is effectively what the empirical $\sqrt{\widehat{N}(s)}$ normalization approximates—homogenizes the rate of concentration across all states $s \in \mathcal{S}$. This variance stabilization improves the conditioning of the spectral representation and yields a sharp concentration rate for $\left\|\widehat{\boldsymbol{W}} - \boldsymbol{W}\right\|_{2\to\infty}$ that completely avoids a detrimental dependence on $\pi_{\min}^{-1}$. This tightly controlled error is critical for the theoretical guarantees of our initial spectral clustering, as we will highlight in the proof sketch of Theorem 4.1.

Third, our L facilitates a more "direct" comparison between the ℓ_2-based separation $\Delta^2_{\boldsymbol{W}}$ and the KL-based separation $\mathcal{D}$. The minimum (row-wise) ℓ_2-separation of $\boldsymbol{W}$, which is a crucial quantity for ℓ_2-distance-based spectral clustering, is defined as:

$$\begin{aligned}\Delta^2_{\boldsymbol{W}} &:= \min_{k\neq k'} \left\|L(\mathcal{M}^{(k)}) - L(\mathcal{M}^{(k')})\right\|_2^2 \\ &= \min_{k\neq k'} \sum_{s\in\mathcal{S}} \left\|\sqrt{\pi^{(k)}(s)}p^{(k)}(\cdot|s) - \sqrt{\pi^{(k')}(s)}p^{(k')}(\cdot|s)\right\|_2^2.\end{aligned}$$

By Proposition 4.1, $\Delta_{\boldsymbol{W}} = 0$ iff at least two of the K Markov models are exactly the same. This also provides an intuitive motivation for our choice of L. Informally, $\Delta^2_{\boldsymbol{W}}$ can be viewed as roughly corresponding to $\min_{k\neq k'} \sum_{s\in\mathcal{S}} \pi^{(k)}(s) \left\|p^{(k)}(\cdot|s) - p^{(k')}(\cdot|s)\right\|_2^2$, which, up to a constant, is upper bounded by $\mathcal{D}_\pi$ (see Proposition 5.1). We discuss in more detail the relations between different gaps in Section 5**(3)**.

Algorithm Design. Stage I is then the usual (row-wise) spectral clustering (von Luxburg, 2007) applied to $\widehat{\boldsymbol{W}}$, divided into two cases. If K is given, we perform standard spectral clustering (lines 2–4) by computing the spectral embedding using the top-K singular vectors and then applying K-means row-wise.

Algorithm 1: Initial Spectral Clustering

Input: $\{(s_{t,1}, \cdots, s_{t,H})\}_{t\in[T]}$, K or γ_{ps}

1 $\widehat{\boldsymbol{U}}\widehat{\boldsymbol{\Sigma}}\widehat{\boldsymbol{V}}^\top \leftarrow$SVD of $\widehat{\boldsymbol{W}} \in [0,1]^{T\times S^2}$ (see Eqn. (5)), with $\widehat{\boldsymbol{\Sigma}} := \mathrm{diag}(\{\widehat{\sigma}_i\}_{i\in[\min(T,S^2)]})$;

2 **if** *K is given* **then**

3 $\quad\widehat{K} \leftarrow K$, $\widehat{\boldsymbol{X}} \leftarrow \widehat{\boldsymbol{U}}_{1:K}\widehat{\boldsymbol{\Sigma}}_{1:K}$;

4 $\quad\widehat{f}_0 \leftarrow K$-means clustering of the rows of $\widehat{\boldsymbol{X}}$;

5 **else**

$\quad$ /* 1. Adaptive Thresholding */

6 $\quad\widehat{R} \leftarrow \sum_{i=1}^{T\wedge S^2} \mathbb{1}\left[\widehat{\sigma}_i \geq \widehat{\sigma}_{\text{thres}} \triangleq \sqrt{\frac{64TS}{H\gamma_{\text{ps}}}\log\frac{TH}{\delta}}\right]$;

7 $\quad\widehat{\boldsymbol{X}} \leftarrow \widehat{\boldsymbol{U}}_{1:\widehat{R}}\widehat{\boldsymbol{\Sigma}}_{1:\widehat{R}}$

$\quad$ /* 2. Density-based Clustering */

8 $\quad$Define $\mathcal{Q}_t \leftarrow \left\{t' \in [T] : \left\|\widehat{\boldsymbol{X}}_{t',:} - \widehat{\boldsymbol{X}}_{t,:}\right\|_2^2 \leq (\widehat{\sigma}_{\text{thres}})^2\right\}$ for each $t \in [T]$;

9 $\quad$Initialize $\mathcal{S}_0 \leftarrow \emptyset$, $k \leftarrow 1$, and $\rho \leftarrow T$;

10 $\quad$**while** $\rho \geq \frac{32\widehat{R}T}{\log\frac{TH}{\delta}}$ **do**

11 $\quad\quad t_k^\star \leftarrow \arg\max_{t\in[T]} \left|\mathcal{Q}_t \setminus \bigcup_{\ell=0}^{k-1}\mathcal{S}_\ell\right|$;

12 $\quad\quad\mathcal{S}_k \leftarrow \mathcal{Q}_{t_k^\star} \setminus \bigcup_{\ell=0}^{k-1}\mathcal{S}_\ell$;

13 $\quad\quad\rho \leftarrow |\mathcal{S}_k|$, $k \leftarrow k+1$;

14 $\quad\widehat{K} := k-1$; `// Estimated # of clusters`

15 $\quad\widehat{f}_0(t) := k$ for $t \in \mathcal{S}_k$ and $k \in [\widehat{K}]$;

$\quad$ /* 3. ℓ_2-Distance-based Clustering */

16 $\quad$**for** $t \in [T] \setminus \bigcup_{k=1}^{\widehat{K}}\mathcal{S}_k$ **do**

17 $\quad\quad\widehat{f}_0(t) := \arg\min_{k\in[\widehat{K}]} \left\|\widehat{\boldsymbol{X}}_{t_k^\star,:} - \widehat{\boldsymbol{X}}_{t,:}\right\|_2$

Return: $\widehat{K}$, $\widehat{f}_0 : [T] \to [\widehat{K}]$

If K is unknown, the procedure consists of three substages (Kannan and Vempala, 2009; Yun and Proutière, 2016; van Vuren et al., 2024). First, we perform *adaptive thresholding*, where the rank of $\boldsymbol{W}$ is estimated via singular value thresholding as $\widehat{R}$,[3] and a low-rank spectral embedding is formed (lines 6–7). Second, we perform *density-based clustering*, where for each trajectory t, we construct a neighborhood $\mathcal{Q}_t$ based on ℓ_2 distance in the embedding space, and then greedily form clusters by selecting $t_{k^\star}$—the trajectory with the largest uncovered neighborhood—as a cluster "center" and assigning its neighbors (lines 8–15). Finally, for the remaining trajectories, we perform *assignment*, where each trajectory is assigned to its nearest cluster center in ℓ_2 distance (lines 16–17).

Theoretical Analysis. We now present the error rate of Stage I:

[3] It could be that $\mathrm{rank}(\boldsymbol{W}) \ll K$; see Appendix H.

Theorem 4.1: Error Rate of Stage I

Let $\delta \in (0,1)$. Suppose that Assumption 1 holds, and that the trajectories are sufficiently long in the following sense:

$$H \gtrsim \frac{1}{\pi_{\min}\gamma_{\mathrm{ps}}} \log \frac{1}{\pi_{\min}\delta} \vee \frac{S}{\Delta_{\boldsymbol{W}}^2 \gamma_{\mathrm{ps}}} \log \frac{H}{\delta} \log \frac{TH}{\delta}.$$

Then, we have: denoting $R = \mathrm{rank}(\boldsymbol{W})$,

$$\mathbb{P}\left(E_T(\widehat{f}_0, f) \lesssim \frac{TRS}{H\gamma_{\mathrm{ps}}\Delta_{\boldsymbol{W}}^2} \log \frac{TH}{\delta} \right) \geq 1 - \delta.$$

Proof sketch. We first prove $\|\boldsymbol{W} - \widehat{\boldsymbol{W}}\|_{2\to\infty} \lesssim \sqrt{\frac{S}{H\gamma_{\mathrm{ps}}} \log \frac{TH}{\delta}}$ (Lemma B.1), our key technical novelty. From here, the proof largely follows that of Yun and Proutière (2016, Algorithm 2). Note that the decay rate is *independent* of $\pi_{\min}^{-1}$, which is crucial in obtaining a tight requirement on H when combined with Stage II. Such rate is possible as the weighting by $\sqrt{\pi(s)}$ in our embedding L cancels out with both $\widetilde{\mathcal{O}}\left(1/\sqrt{H\pi(s)}\right)$ for ℓ_2-estimation of $p(\cdot|s)$ and $\widetilde{\mathcal{O}}\left(1/\sqrt{H\pi(s)\gamma_{\mathrm{ps}}}\right)$ for the estimation of $\sqrt{\pi(s)}$, which arises from triangle inequality. The full proof is deferred to Appendix B. □

Towards a Parameter-Free Algorithm. When K is known, our algorithm does not require knowledge of the pseudo-spectral gap γ_{ps}, and vice versa. However, to achieve a truly parameter-free algorithm where both K and γ_{ps} are unknown, one must first estimate γ_{ps}. This is because estimating the number of clusters K relies on spectral thresholding derived from Markovian concentration bounds, which inherently depend on γ_{ps} (line 6 of Algorithm 1). Since Stage II is already parameter-free as we will see later, obtaining a reliable estimator for γ_{ps} is the sole remaining hurdle. Theoretically, we only require a high-probability lower bound on γ_{ps}; specifically, an estimator $\hat{\gamma}_{\mathrm{ps}}$ satisfying $(1/2)\gamma_{\mathrm{ps}} \leq \hat{\gamma}_{\mathrm{ps}} \leq (3/2)\gamma_{\mathrm{ps}}$ would allow our current analysis to hold up to constant factors.

Crucially, this estimation introduces additional sample complexity requirements on the trajectory length H. A straightforward approach of taking the minimum over per-trajectory estimators, such as the one proposed by Wolfer and Kontorovich (2024, Theorem 1), can furnish the required lower bound. However, this necessitates a longer trajectory length of $H \gtrsim \gamma_{\mathrm{ps}}^{-3}(\log T)^2$ (ignoring other factors). This is strictly more demanding than our known-parameter regime by an extra $\gamma_{\mathrm{ps}}^{-2}$ factor and a $\log T$ term. The extra $\log T$ dependence arises from the union bound over T trajectories and can potentially be mitigated via logarithmic subsampling. Conversely, the extra $\gamma_{\mathrm{ps}}^{-2}$ factor is likely unavoidable in the fully parameter-free setting, given the minimax lower bounds for estimating the pseudo-spectral gap itself (Wolfer and Kontorovich, 2019, Theorem 4). We defer further details and extended discussions to Appendix I.

Algorithm 2: Likelihood Improvement

Input: $\widehat{f}_0 : [T] \to [K]$, $\{(s_{t,1}, \cdots, s_{t,H})\}_{t\in[T]}$

1 Estimate $p^{(k)}(\cdot|\cdot)$ for each $k \in [K]$ as follows: denoting $\widehat{\mathcal{C}}_0 := (\widehat{f}_0)^{-1}(k)$, for all $s, s' \in \mathcal{S}$,

$$\widehat{p}_0^{(k)}(s'|s) \leftarrow \frac{\sum_{t\in\widehat{\mathcal{C}}_0}\sum_{h\in[H-1]} \mathbb{1}[s_{t,h} = s, s_{t,h+1} = s']}{\sum_{t\in\widehat{\mathcal{C}}_0}\sum_{h\in[H-1]} \mathbb{1}[s_{t,h} = s]}.$$

2 Refine the cluster estimates via trajectory-wise maximum likelihood estimator: for each $t \in [T]$,

$$\widehat{f}(t) \leftarrow \arg\max_{k\in[K]} \left\{ \mathcal{L}(k;t) \triangleq \sum_{h=1}^{H-1} \log \widehat{p}_0^{(k)}(s_{t,h+1}|s_{t,h}) \right\}$$

Return: $\widehat{f} : [T] \to [K]$

4.2 Stage II. One-Shot Trajectory Likelihood Improvement

Algorithm Design. By the Neyman-Pearson lemma (Neyman and Pearson, 1933) and from our information-theoretic lower bound (Theorem 3.1), we must perform trajectory-wise likelihood testing to achieve the optimal performance. But, we do not know the transition kernels p beforehand. Hence, we first utilize the "good enough" $\widehat{f}_0 : [T] \to [K]$ from Stage I to estimate the transition kernels $\widehat{p}_0$ (line 1). Then, we perform likelihood-based reassignment of the cluster labels based on $\widehat{p}_0$ (line 2). A keen reader may note that this is essentially the hard Expectation-Maximization (EM) algorithm (Celeux and Govaert, 1992; Kearns et al., 1997).

Theoretical Analysis. For the analysis of Stage II, we consider the following additional assumption:

Assumption 2 (η-regularity *across* the chains)**.** *There exists $\eta_\pi, \eta_p > 1$ such that:*

$$\max_{s\in\mathcal{S}} \max_{k,k'} \frac{\pi^{(k)}(s)}{\pi^{(k')}(s)} \leq \eta_\pi, \quad \max_{s,s'\in\mathcal{S}} \max_{k,k'} \frac{p^{(k)}(s'|s)}{p^{(k')}(s'|s)} \leq \eta_p.$$

We now present the final error guarantee of Algorithm 2; the proof is deferred to Appendix C. Recalling that $\mathcal{D}_\pi$ is the stationary divergence (Eqn. (4)),

Theorem 4.2: Error Rate of Stage I+II

Suppose Assumptions 1 and 2 hold, and that T, H are sufficiently large in the following sense:

$$TH \gtrsim \frac{(S\log T)^2}{\gamma_{\mathrm{ps}}\alpha_{\min}\mathcal{D}_\pi^2}, \text{ and}$$

$$H \gtrsim \frac{1}{\gamma_{\mathrm{ps}}}\left(S^2\log T \vee \frac{\log T}{\pi_{\min}} \vee \frac{\eta_p\eta_\pi RS\log(TH)}{\alpha_{\min}\mathcal{D}_\pi\Delta_{\boldsymbol{W}}^2}\right).$$

Then, for some $C_\eta \asymp \frac{1}{\eta_p^2}$, we have that w.h.p.[a],

$$E_T(\widehat{f}, f) \lesssim T\exp\left(-C_\eta\gamma_{\mathrm{ps}}H\mathcal{D}_\pi\right). \quad (7)$$

[a] with probability tending to 1 as $T \to \infty$.

One can immediately see that the obtained upper bound on the clustering error rate nearly matches the lower bound (Theorem 3.1), but there are two gaps, which we elaborate on in the subsequent section.

Necessity of Assumption 2. We first clarify that Assumption 2 does not require the chains to be uniform, i.e., it does not impose $\pi^{(k)}(s) \asymp 1/S$, as is commonly assumed in stochastic block models (Yun and Proutière, 2014a) or block Markov chains (Sanders et al., 2020; Jedra et al., 2023). Instead, our assumption only requires that if one chain exhibits non-uniformity in certain states, the other chains must exhibit similar non-uniformity in those same states.

Crucially, this assumption is only required for the theoretical analysis of Stage II; Stage I's guarantee (Theorem 4.1) relies solely on Assumption 1. Theoretically, Assumption 2 acts as a proof artifact necessary to bound the summands when applying the Markovian Bernstein concentration inequality (Paulin, 2015) to the empirical likelihood ratios. Consequently, our current proof technique does not cover scenarios in which it fails, such as near-deterministic transitions in which a chain remains in a distinct state.

Empirically, however, this strong assumption is rarely necessary. When Assumption 2 fails, it often implies that certain chains are highly non-uniform and thus trivially easy to distinguish, a regime where we expect Stage II to remain robust. To handle potential numerical instabilities in such extreme scenarios – such as division by zero when candidate models assign near-zero probabilities to observed transitions – one can easily regularize the likelihood evaluation via Laplace smoothing (Laplace, 1812). Provided Stage I yields a sufficiently accurate initial clustering, this smoothed likelihood ratio faithfully captures the high information content, heavily penalizing incorrect models while inducing negligible bias. The precise theoretical analysis is left to future work.

What If Reversible? Conversely, one might ask whether additional structural assumptions, notably reversibility, a common property in Markov chain Monte Carlo (MCMC) methods (Robert and Casella, 2004), would simplify the analysis or yield sharper bounds.

Assuming reversibility would indeed simplify certain aspects of the proof. First, the classical spectral gap (Lezaud, 1998) becomes well-defined, eliminating the need to rely on the pseudo-spectral gap. Second, one could apply Bernstein's inequality specifically tailored for reversible Markov chains (Paulin, 2015, Theorem 3.9), which yields slightly improved constants.

However, beyond these simplifications, we do not expect the reversibility assumption to yield fundamentally sharper bounds. The order of Markovian concentration remains the same regardless of reversibility; only the governing gap parameter (spectral vs. pseudo-spectral) changes. Furthermore, Wolfer and Kontorovich (2021) established minimax-optimal rates for estimating Markov chain kernels, demonstrating that simple visitation-based empirical transition estimation yields a tight upper bound for any ergodic Markov chain, whereas their lower-bound construction involves reversible chains (Wolfer and Kontorovich, 2021, Section 6.4). Because our transition estimation in Stage II mirrors this approach (applied to grouped trajectories based on the initial clustering), restricting the scope to reversible chains is unlikely to improve the theoretical clustering error rate.

5 DISCUSSIONS

Throughout the discussions, for simplicity, we assume $\alpha_{\min} \asymp 1/K$, i.e., roughly balanced cluster sizes.

(1) Gaps Between Upper and Lower Bounds. By comparing our lower bound (Theorem 3.1) and upper bound (Theorem 4.2), we identify two key gaps.

The first concerns the clustering error rates: the lower bound scales as $\exp(-2H\mathcal{D})$, while the upper bound scales as $\exp(-C_\eta\gamma_{\mathrm{ps}}H\mathcal{D}_\pi)$, where C_η is a constant depending on the regularity parameters η_p and η_π. The upper bound thus contains an additional multiplicative factor of γ_{ps} in the exponential term. The second gap pertains to the requirement on the horizon length H. Specifically, the lower bound holds for any $H \geq 2$ and $T \geq 1$, while the upper bound requires $H = \widetilde{\Omega}(\gamma_{\mathrm{ps}}^{-1}(S^2 \vee \pi_{\min}^{-1} \vee KS))$ and $TH = \widetilde{\Omega}(\gamma_{\mathrm{ps}}^{-1}KS^2)$, omitting logarithmic and other factors.

Specifically, the first gap arises from employing the Markovian Bernstein concentration inequality of Paulin (2015), which relies on decomposing the trajectories of the Markov chain into nearly independent blocks of

length proportional to $\gamma_{\mathrm{ps}} H$ (Yu, 1994; Janson, 2004; Lezaud, 1998). Recent results confirm that this dependence on γ_{ps} in the exponent is indeed unavoidable for (reversible) Markov chains: there exist chains for which the concentration rate matches this dependency (Jiang et al., 2018; Fan et al., 2021; Rabinovich et al., 2020). The second gap stems from the concentration of $\widehat{\boldsymbol{W}}$ around $\boldsymbol{W}$, which in turn requires accurate estimation of the transition kernels. Wolfer and Kontorovich (2021, Theorem 3.2) show the following requirement on H: if $H = o\left(\frac{1}{\varepsilon^2 \pi_{\min}} + \frac{S}{\gamma_{\mathrm{ps}}}\right)$, then no estimator can achieve ε-accuracy in $\|\cdot\|_{2\to\infty}$.[4]

We believe these gaps indicate that the current lower bound is loose, as the lower bound analysis assumes complete knowledge of the p's; we leave it to future work to establish a tighter lower bound. Conceptually, this fits within a nonparametric testing viewpoint (Tsybakov, 2009): how does uncertainty in the underlying measures degrade testing power and, in turn, the achievable clustering error?

(2) Comparison with Kausik et al. (2023). From an algorithmic perspective, a notable limitation of Kausik et al. (2023) is that their approach requires prior knowledge of the problem parameters K, α, and Δ, whereas our proposed algorithm is entirely parameter-free. Turning to the statistical guarantees, we recall their condition for exact cluster recovery:

Assumption 3 (Assumption 2 of Kausik et al. 2023)**.** $\exists \alpha, \Delta > 0$ *s.t.* $\forall k \neq k' \in [K]$ $\exists s = s(k,k')$:

$$\pi^{(k)}(s), \pi^{(k')}(s) \geq \alpha \quad \textit{and} \quad \left\| p^{(k)}(\cdot|s) - p^{(k')}(\cdot|s) \right\|_2 \geq \Delta.$$

Theorem 5.1 (Theorem 3 of Kausik et al. 2023)**.** *Suppose that Assumptions 1 and 3 hold. Set the occurrence and histogram clustering thresholds as* $\beta = \frac{\alpha}{3}$ *and* $\tau = \frac{3\Delta^2}{4}$*. Then, with* $H \gtrsim K^{3/2} t_{\mathrm{mix}} \frac{\left(\log(T/(\alpha\delta))\right)^4}{\alpha^3 \Delta^6}$ *and* $T \gtrsim K^2 S \frac{\log(1/\delta)}{\alpha^3 \Delta^8}$*, their Algorithm 2 attains exact clustering with probability at least* $1-\delta$*.*

Kausik et al. (2023) treat α and Δ as constants, independent of the number of states S and the trajectory length H. This assumption allows their condition on H to avoid any polynomial dependence on S, and ensures that the requirement on T scales only linearly with S. In effect, they implicitly assume that for every pair $k \neq k'$, there exists a highly visited state $s(k,k')$ with constant separation and from which clusters can be identified. However, this assumption does not generally hold. For example, in uniform-like Markov chains where $\pi(s) \asymp 1/S$, they require[5] $H \gtrsim \gamma_{\mathrm{ps}}^{-1} K^{3/2} S^3 (\log T)^4$ and $T \gtrsim K^2 S^4 \log T$ (with $\delta = 1/T$), which are impractically large. In the same regime, our algorithm requires only $H \gtrsim \gamma_{\mathrm{ps}}^{-1}(S^2 \vee KS)\log T$ and $TH \gtrsim \gamma_{\mathrm{ps}}^{-1} K S^2$, significant improvements. An interesting future direction is to ask whether an *instance-wise near-optimal* algorithm can be designed – achieving the near-optimal error rate guarantee when the cluster recovery condition requires only $H \gtrsim \gamma_{\mathrm{ps}}^{-1} \log T$ under Assumption 3 with $\alpha\Delta^2 = \Omega(1)$, yet requires only $H \gtrsim \gamma_{\mathrm{ps}}^{-1}(S^2 \vee \pi_{\min}^{-1} \vee KS)\log T$ (or better) in the worst case, ignoring other factors.

(3) Different gaps: $\mathcal{D}_\pi$, $\Delta_{\boldsymbol{W}}$, and $\alpha\Delta^2$. Throughout this paper, we have introduced several notions of separability gaps: the KL gap $\mathcal{D}_\pi$, the "aggregated" ℓ_2 gap $\Delta_{\boldsymbol{W}}$, and the Δ, α from Kausik et al. (2023). The following proposition makes their relationships precise (see Appendix D for the proof):

Proposition 5.1. *We have the following: with* $p_{\max} := \max_{k\in[K]} \max_{s,s'\in\mathcal{S}} p^{(k)}(s'|s)$*,*

(a) $p_{\max}\mathcal{D}_\pi \gtrsim \alpha\Delta^2$*; (b)* $p_{\max}\mathcal{D}_\pi + \sqrt{\mathcal{D}_\pi/\pi_{\min}} \gtrsim \Delta_{\boldsymbol{W}}^2$*;*

(c) under Assumption 2, $\Delta_{\boldsymbol{W}}^2 \gtrsim \alpha\Delta^2 - \mathcal{O}((\sqrt{\eta_\pi}-1)^2)$*.*

The inequality *(b)* holds up to uniform ergodicity contraction constants (see Appendix D). Because of our particular choice of $\boldsymbol{W}$ (Eqn. (5)), the gaps $\Delta_{\boldsymbol{W}}^2$, $\alpha\Delta^2$, and $\mathcal{D}_\pi$ are comparable. Indeed, for η-regular ergodic Markov chains with $\alpha\Delta^2 \gtrsim (\sqrt{\eta_\pi}-1)^2$, we roughly have $\alpha\Delta^2 \lesssim \Delta_{\boldsymbol{W}}^2 \lesssim p_{\max}\mathcal{D}_\pi$.

Intuitively, Kausik et al. (2023)'s $\alpha\Delta^2$ accounts for separation from a single "good" state; $\Delta_{\boldsymbol{W}}$ aggregates information across all states but in a geometrically suboptimal manner; our $\mathcal{D}_\pi$ reflects the "correctly" aggregated information. Moreover, for uniform-like Markov chains where $p_{\max} \asymp \frac{1}{S}$, Proposition 5.1(b) implies a factor-S gap between $\mathcal{D}_\pi$ and the other two gaps.

Although our optimal clustering error rates depend only on $\mathcal{D}_\pi$, the success of ℓ_2-based spectral clustering in Stage I requires $H \gtrsim \frac{1}{\mathcal{D}_\pi \Delta_{\boldsymbol{W}}^2}$. For comparison, Kausik et al. (2023) requires $H \gtrsim \frac{1}{(\alpha\Delta^2)^3}$, reflecting a heavier dependence on more suboptimal separability gap. To make the contrast between $\Delta_{\boldsymbol{W}}^2$ and $\alpha\Delta^2$ explicit, we construct an instance of MMC – which may be of independent interest – where the three gaps exhibit a strict hierarchy in terms of S-factor (see Appendix E for the proof):

Proposition 5.2. *There exists an instance of* ***clustering in MMC*** *such that* $\mathcal{D}_\pi \asymp S\Delta_{\boldsymbol{W}}^2 \asymp S^2\alpha\Delta^2$*.*

[4] Their theorem is stated for $\|\cdot\|_{1\to\infty}$, but the same argument applies to $\|\cdot\|_{2\to\infty}$ with minimal modification.

[5] Kausik et al. (2023) wrote their guarantees in terms of the mixing time t_{mix}. Here, for a clear comparison, we write them in terms of $\gamma_{\mathrm{ps}}^{-1}$. Indeed, both are of same order, up to $\log \pi_{\min}^{-1}$ (Paulin, 2015, Proposition 3.4).

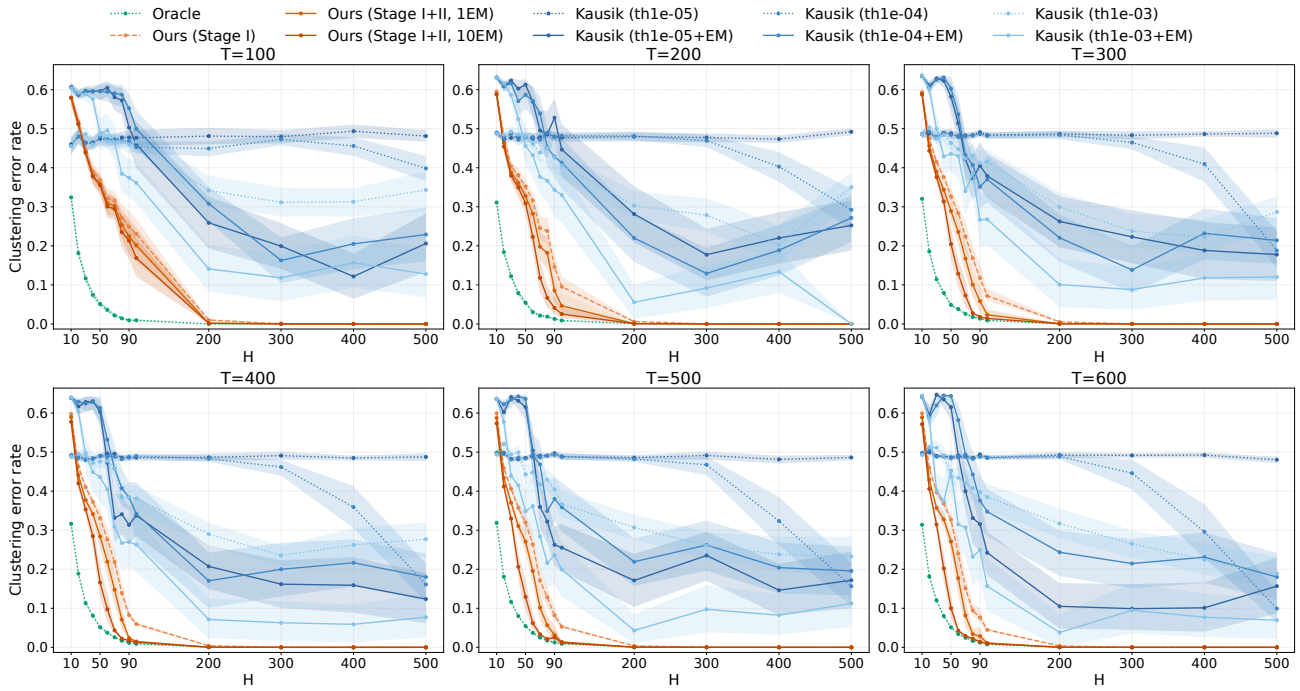

Figure 1: Clustering error rates of our and Kausik et al. (2023)'s algorithms on **Cyclic-Bump MMC**.

(4) Computational Scalability. While our algorithm achieves near-optimal statistical guarantees, it is not currently optimized for computational efficiency. Specifically, the SVD step in Stage I (line 1 of Algorithm 1) requires $\mathcal{O}(TS^2)$ space and $\mathcal{O}\big(TS^2 \min(T, S^2)\big)$ time, which poses a scalability bottleneck for large state spaces S or trajectory counts T. We leave the development of scalable variants to future work; see Appendix J**(1)** for some potential directions.

6 EXPERIMENTS

Setting. We evaluate our algorithm on the **Cyclic-Bump MMC**, a synthetic instance from Proposition 5.2 (Appendix E.1), with $S = 10$, $\zeta = 0.9/S$, and $K = 3$. We compare three variants of our method (known K): `Stage I`, `Stage I+II (EM1)`, and `Stage I+II (EM10)`, where the latter tests the benefit of multiple EM iterations. We benchmark against Kausik et al. (2023), denoted `Kausik (th1e-05)`, `Kausik (th1e-04)`, and `Kausik (th1e-03)`, where 'th' refers to the threshold level, a hyperparameter that we sweep over $\{10^{-5}, 10^{-4}, 10^{-3}\}$. '+EM' denotes running 10 EM iterations on top of the subspace-based clustering of Kausik et al. (2023). We also include the `Oracle` likelihood test, which is statistically optimal by the Neyman-Pearson lemma (Neyman and Pearson, 1933). Further details are relegated to Appendix F.1.

Results. As shown in Figure 1, `Stage I` alone already outperforms Kausik et al. (2023). We also note that their performance is highly sensitive to the tuning of the threshold hyperparameter. Adding Stage II refinement (`EM1`, `EM10`) further boosts accuracy—especially in the low-H, high-T regime—approaching the `Oracle` curve as theoretically predicted.

Additional Experiments. In Appendix F, we provide four ablation studies: the impact of Stage II iterations, the impact of S and T on performance and runtime, unknown K, and varying pseudo-spectral gaps γ_{ps} (Appendices F.2–F.5). We also evaluate on the Last.fm 1K (Celma, 2010) dataset in Appendix F.6, where we *significantly outperform* Kausik et al. (2023).

7 CONCLUSION

This paper studies trajectory clustering in MMCs. We established an instance-specific lower bound on the achievable clustering error and developed a computationally tractable two-stage algorithm that attains near-optimal rates without requiring any prior knowledge of the underlying MMC model. The key ingredient is Stage I, which introduces a new injective Euclidean embedding for ergodic Markov chains, thereby enabling sharp concentration guarantees for spectral clustering. We numerically validate the efficacy of our algorithm. Further future directions are detailed in Appendix J.

Acknowledgements

We thank all the anonymous reviewers for their helpful comments that helped us improve the manuscript. A. Proutière is supported by Vetenskaprådet, Digital Futures, and the Wallenberg AI, Autonomous Systems and Software program.

References

Emmanuel Abbe. Community Detection and Stochastic Block Models. *Foundations and Trends® in Communications and Information Theory*, 14(1-2):1–162, 2018. ISSN 1567-2190. doi:10.1561/0100000067.

Emmanuel Abbe and Colin Sandon. Community Detection in General Stochastic Block models: Fundamental Limits and Efficient Algorithms for Recovery. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 670–688, 2015a. doi:10.1109/FOCS.2015.47.

Emmanuel Abbe and Colin Sandon. Recovering Communities in the General Stochastic Block Model Without Knowing the Parameters. In *Advances in Neural Information Processing Systems*, volume 28, pages 676–684. Curran Associates, Inc., 2015b. URL https://arxiv.org/abs/1506.03729.

Shun'ichi Amari. *Information Geometry and Its Applications*. Applied Mathematical Sciences. Springer Tokyo, 2016. doi:10.1007/978-4-431-55978-8.

Arash A. Amini, Aiyou Chen, Peter J. Bickel, and Elizaveta Levina. Pseudo-likelihood methods for community detection in large sparse networks. *The Annals of Statistics*, 41(4):2097 – 2122, 2013. doi:10.1214/13-AOS1138.

Animashree Anandkumar, Daniel Hsu, and Sham M. Kakade. A Method of Moments for Mixture Models and Hidden Markov Models. In *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 33.1–33.34, Edinburgh, Scotland, 25–27 Jun 2012. PMLR. URL https://proceedings.mlr.press/v23/anandkumar12.html.

Kamyar Azizzadenesheli, Alessandro Lazaric, and Animashree Anandkumar. Reinforcement Learning of POMDPs using Spectral Methods. In *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 193–256, Columbia University, New York, New York, USA, 23–26 Jun 2016. PMLR. URL https://proceedings.mlr.press/v49/azizzadenesheli16a.html.

Kamyar Azizzadenesheli, Alessandro Lazaric, and Animashree Anandkumar. Reinforcement Learning in Rich-Observation MDPs using Spectral Methods. In *Multi-disciplinary Conference on Reinforcement Learning and Decision Making (RLDM)*, 2017. URL https://arxiv.org/abs/1611.03907.

Imon Banerjee, Harsha Honnappa, and Vinayak Rao. Off-line Estimation of Controlled Markov Chains: Minimaxity and Sample Complexity. *Operations Research*, 73(4):2281–2295, 2025. doi:10.1287/opre.2023.0046.

Patrick Billingsley. Statistical Methods in Markov Chains. *The Annals of Mathematical Statistics*, 32 (1):12 – 40, 1961. doi:10.1214/aoms/1177705136.

G. Birkhoff. Tres observaciones sobre el álgebra lineal. *Revista, Universitad Nacional de Tucumán, Serie A*, 5:147–154, 1946.

Isadore Blumen, Marvin Kogan, and Philip J. McCarthy. *The Industrial Mobility of Labor as a Probability Process*, volume 6 of *Cornell Studies in Industrial and Labor Relations*. Cornell University, Ithaca, NY, 1955.

Igor Cadez, David Heckerman, Christopher Meek, Padhraic Smyth, and Steven White. Model-Based Clustering and Visualization of Navigation Patterns on a Web Site. *Data Mining and Knowledge Discovery*, 7(4):399–424, Oct 2003. ISSN 1573-756X. doi:10.1023/A:1024992613384.

Sergio Calo, Anders Jonsson, Gergely Neu, Ludovic Schwartz, and Javier Segovia-Aguas. Bisimulation Metrics are Optimal Transport Distances, and Can be Computed Efficiently. In *Advances in Neural Information Processing Systems*, volume 37, pages 134345–134386. Curran Associates, Inc., 2024. URL https://openreview.net/forum?id=CSjVSnvTbG.

Sergio Calo, Anders Jonsson, Gergely Neu, Ludovic Schwartz, and Javier Segovia-Aguas. Distances for Markov chains from sample streams. *arXiv preprint arXiv:2505.18005*, 2025. URL https://arxiv.org/abs/2505.18005.

Clément L. Canonne. A short note on learning discrete distributions. *arXiv preprint arXiv:2002.11457*, 2020. URL https://arxiv.org/abs/2002.11457.

Joshua Cape, Minh Tang, and Carey E. Priebe. The two-to-infinity norm and singular subspace geometry with applications to high-dimensional statistics. *The Annals of Statistics*, 47(5):2405 – 2439, 2019. doi:10.1214/18-AOS1752.

Gilles Celeux and Gérard Govaert. A classification EM algorithm for clustering and two stochastic versions. *Computational Statistics & Data Analysis*, 14 (3):315–332, 1992. ISSN 0167-9473. doi:10.1016/0167-9473(92)90042-E.

Óscar Celma. *Music Recommendation and Discovery – The Long Tail, Long Fail, and Long Play in the*

Digital Music Space. Springer Berlin, Heidelberg, 2010. doi:10.1007/978-3-642-13287-2.

Fang Chen, László Lovász, and Igor Pak. Lifting Markov Chains to Speed up Mixing. In *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing*, STOC '99, page 275–281, New York, NY, USA, 1999. Association for Computing Machinery. ISBN 1581130678. doi:10.1145/301250.301315.

Xin Chen and Anderson Y. Zhang. Achieving Optimal Clustering in Gaussian Mixture Models with Anisotropic Covariance Structures. In *Advances in Neural Information Processing Systems*, volume 37, pages 113698–113741. Curran Associates, Inc., 2024. URL https://openreview.net/forum?id=ge8GZn8Gtu.

Fan R. K. Chung. *Spectral Graph Theory*, volume 92 of *CBMS Regional Conference Series in Mathematics*. American Mathematical Society, Providence, RI, 1997. ISBN 978-0-8218-0315-8.

Ronald R. Coifman and Stéphane Lafon. Diffusion maps. *Applied and Computational Harmonic Analysis*, 21(1):5–30, 2006. ISSN 1063-5203. doi:10.1016/j.acha.2006.04.006. Special Issue: Diffusion Maps and Wavelets.

Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley-Interscience, Hoboken, NJ, 2nd edition, 2006. ISBN 978-0-471-24195-9.

I. Csiszar and Z. Talata. Context tree estimation for not necessarily finite memory processes, via BIC and MDL. *IEEE Transactions on Information Theory*, 52 (3):1007–1016, 2006. doi:10.1109/TIT.2005.864431.

Persi Diaconis. *Group Representations in Probability and Statistics*, volume 11 of *Institute of Mathematical Statistics Lecture Notes—Monograph Series*. Institute of Mathematical Statistics, Hayward, CA, 1988. ISBN 0-940600-14-5. doi:10.1214/lnms/1215467407.

Thomas J. DiCiccio and Bradley Efron. Bootstrap confidence intervals. *Statistical Science*, 11(3):189 – 228, 1996. doi:10.1214/ss/1032280214.

Maximilien Dreveton, Felipe Fernandes, and Daniel Figueiredo. Exact recovery and Bregman hard clustering of node-attributed Stochastic Block Model. In *Advances in Neural Information Processing Systems*, volume 36, pages 37827–37848. Curran Associates, Inc., 2023. URL https://openreview.net/forum?id=TjJJmcHw9p.

Maximilien Dreveton, Alperen Gözeten, Matthias Grossglauser, and Patrick Thiran. Universal Lower Bounds and Optimal Rates: Achieving Minimax Clustering Error in Sub-Exponential Mixture Models. In *Proceedings of Thirty Seventh Conference on Learning Theory*, volume 247 of *Proceedings of Machine Learning Research*, pages 1451–1485. PMLR, 30 Jun–03 Jul 2024. URL https://proceedings.mlr.press/v247/dreveton24a.html.

B. Efron. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1):1 – 26, 1979. doi:10.1214/aos/1176344552.

Absalom E. Ezugwu, Abiodun M. Ikotun, Olaide O. Oyelade, Laith Abualigah, Jeffery O. Agushaka, Christopher I. Eke, and Andronicus A. Akinyelu. A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects. *Engineering Applications of Artificial Intelligence*, 110:104743, 2022. ISSN 0952-1976. doi:10.1016/j.engappai.2022.104743.

Jianqing Fan, Bai Jiang, and Qiang Sun. Hoeffding's Inequality for General Markov Chains and Its Applications to Statistical Learning. *Journal of Machine Learning Research*, 22(139):1–35, 2021. URL http://jmlr.org/papers/v22/19-479.html.

Santo Fortunato. Community detection in graphs. *Physics Reports*, 486(3):75–174, 2010. ISSN 0370-1573. doi:10.1016/j.physrep.2009.11.002.

Chao Gao, Zongming Ma, Anderson Y. Zhang, and Harrison H. Zhou. Achieving Optimal Misclassification Proportion in Stochastic Block Models. *Journal of Machine Learning Research*, 18(60):1–45, 2017. URL http://jmlr.org/papers/v18/16-245.html.

Aurélien Garivier, Pierre Ménard, and Gilles Stoltz. Explore First, Exploit Next: The True Shape of Regret in Bandit Problems. *Mathematics of Operations Research*, 44(2):377–399, 2019. doi:10.1287/moor.2017.0928.

Mark Girolami and Ata Kabán. Simplicial Mixtures of Markov Chains: Distributed Modelling of Dynamic User Profiles. In *Advances in Neural Information Processing Systems*, volume 16, pages 9–16. MIT Press, 2003. URL https://proceedings.neurips.cc/paper_files/paper/2003/file/4a8423d5e91fda00bb7e46540e2b0cf1-Paper.pdf.

Rishi Gupta, Ravi Kumar, and Sergei Vassilvitskii. On Mixtures of Markov Chains. In *Advances in Neural Information Processing Systems*, volume 29, pages 3449–3457. Curran Associates, Inc., 2016. URL https://proceedings.neurips.cc/paper/2016/hash/8b5700012be65c9da25f49408d959ca0-Abstract.html.

Wassily Hoeffding. Probability Inequalities for Sums of Bounded Random Variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963. doi:10.1080/01621459.1963.10500830.

Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, 2 edition, 2012.

Daniel Hsu, Sham M. Kakade, and Tong Zhang. A spectral algorithm for learning Hidden Markov Models. *Journal of Computer and System Sciences*, 78(5):1460–1480, 2012. ISSN 0022-0000. doi:10.1016/j.jcss.2011.12.025. JCSS Special Issue: Cloud Computing 2011.

Svante Janson. Large deviations for sums of partly dependent random variables. *Random Structures & Algorithms*, 24(3):234–248, 2004. doi:10.1002/rsa.20008.

Yassir Jedra and Alexandre Proutière. Finite-Time Identification of Linear Systems: Fundamental Limits and Optimal Algorithms. *IEEE Transactions on Automatic Control*, 68(5):2805–2820, 2023. doi:10.1109/TAC.2022.3221705.

Yassir Jedra, Junghyun Lee, Alexandre Proutière, and Se-Young Yun. Nearly Optimal Latent State Decoding in Block MDPs. In *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*, volume 206 of *Proceedings of Machine Learning Research*, pages 2805–2904. PMLR, 25–27 Apr 2023. URL https://proceedings.mlr.press/v206/jedra23a.html.

Bai Jiang, Qiang Sun, and Jianqing Fan. Bernstein's inequalities for general Markov chains. *arXiv preprint arXiv:1805.10721*, 2018. URL https://arxiv.org/abs/1805.10721.

Ravindran Kannan and Santosh Vempala. Spectral Algorithms. *Foundations and Trends® in Theoretical Computer Science*, 4(3–4):157–288, 2009. ISSN 1551-305X. doi:10.1561/0400000025.

Chinmaya Kausik, Kevin Tan, and Ambuj Tewari. Learning Mixtures of Markov Chains and MDPs. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 15970–16017. PMLR, 23–29 Jul 2023. URL https://proceedings.mlr.press/v202/kausik23a.html.

Michael Kearns, Yishay Mansour, and Andrew Y. Ng. An Information-Theoretic Analysis of Hard and Soft Assignment Methods for Clustering. In *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence*, UAI'97, page 282–293, San Francisco, CA, USA, 1997. Morgan Kaufmann Publishers Inc. ISBN 1558604855.

Vladimir Yu Kiselev, Tallulah S. Andrews, and Martin Hemberg. Challenges in unsupervised clustering of single-cell RNA-seq data. *Nature Reviews Genetics*, 20(5):273–282, May 2019. ISSN 1471-0064. doi:10.1038/s41576-018-0088-9.

Jeongyeol Kwon, Yonathan Efroni, Constantine Caramanis, and Shie Mannor. RL for Latent MDPs: Regret Guarantees and a Lower Bound. In *Advances in Neural Information Processing Systems*, volume 34, pages 24523–24534. Curran Associates, Inc., 2021. URL https://openreview.net/forum?id=CLCVcl1rSPP.

Jeongyeol Kwon, Constantine Caramanis, Shie Mannor, and Yonathan Efroni. RL in Latent MDPs is Tractable: Online Guarantees via Off-Policy Evaluation. In *Advances in Neural Information Processing Systems*, volume 37, pages 82726–82756. Curran Associates, Inc., 2024a. URL https://openreview.net/forum?id=juJl2uSq4D.

Jeongyeol Kwon, Yonathan Efroni, Shie Mannor, and Constantine Caramanis. Prospective Side Information for Latent MDPs. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pages 25755–25783. PMLR, 21–27 Jul 2024b. URL https://proceedings.mlr.press/v235/kwon24a.html.

Tse L. Lai and Herbert Robbins. Asymptotically Efficient Adaptive Allocation Rules. *Advances in Applied Mathematics*, 6(1):4–22, 1985. doi:10.1016/0196-8858(85)90002-8.

Pierre-Simon Laplace. *Théorie analytique des probabilités*. Courcier, Paris, 1812.

David A. Levin and Yuval Peres. *Markov Chains and Mixing Times*. American Mathematical Society, 2 edition, 2017. ISBN 978-1-4704-2962-1.

Pascal Lezaud. Chernoff-type bound for finite Markov chains. *The Annals of Applied Probability*, 8(3):849 – 867, 1998. doi:10.1214/aoap/1028903453.

Matthias Löffler, Anderson Y. Zhang, and Harrison H. Zhou. Optimality of spectral clustering in the Gaussian mixture model. *The Annals of Statistics*, 49(5): 2506 – 2530, 2021. doi:10.1214/20-AOS2044.

Yu Lu and Harrison H. Zhou. Statistical and Computational Guarantees of Lloyd's Algorithm and its Variants. *arXiv preprint arXiv:1612.02099*, 2016. URL https://arxiv.org/abs/1612.02099.

Lucas Maystre, Tiffany Wu, Roberto Sanchis-Ojeda, and Tony Jebara. Multistate analysis with infinite mixtures of Markov chains. In *Proceedings of the Thirty-Eighth Conference on Uncertainty in Artificial Intelligence*, volume 180 of *Proceedings of Machine Learning Research*, pages 1350–1359. PMLR, 01–05 Aug 2022. URL https://proceedings.mlr.press/v180/maystre22a.html.

Colin McDiarmid. *On the method of bounded differences*, page 148–188. London Mathematical Society Lecture Note Series. Cambridge University Press, 1989.

Geoffrey J. McLachlan, Sharon X. Lee, and Suren I. Rathnayake. Finite Mixture Models. *Annual Review of Statistics and Its Application*, 6:355–378, 2019. doi:10.1146/annurev-statistics-031017-100325.

A. Yu. Mitrophanov. Sensitivity and convergence of uniformly ergodic Markov chains. *Journal of Applied Probability*, 42(4):1003–1014, 2005. doi:10.1239/jap/1134587812.

Jerzy Neyman and Egon Sharpe Pearson. IX. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 231(694-706): 289–337, 1933. doi:10.1098/rsta.1933.0009.

Frank Nielsen. An Elementary Introduction to Information Geometry. *Entropy*, 22(10), Sep 2020. ISSN 1099-4300. doi:10.3390/e22101100.

Daniel Paulin. Concentration inequalities for Markov chains by Marton couplings and spectral methods. *Electronic Journal of Probability*, 20(none):1 – 32, 2015. doi:10.1214/EJP.v20-4039.

Carsten Stig Poulsen. Mixed Markov and latent Markov modelling applied to brand choice behaviour. *International Journal of Research in Marketing*, 7(1): 5–19, 1990. ISSN 0167-8116. doi:10.1016/0167-8116(90)90028-L.

Maxim Rabinovich, Aaditya Ramdas, Michael I. Jordan, and Martin J. Wainwright. Function-Specific Mixing Times and Concentration Away from Equilibrium. *Bayesian Analysis*, 15(2):505 – 532, 2020. doi:10.1214/19-BA1151.

Christian P Robert and George Casella. *Monte Carlo Statistical Methods*. Springer Texts in Statistics. Springer New York, 2nd edition, 2004.

Jaron Sanders, Alexandre Proutière, and Se-Young Yun. Clustering in Block Markov Chains. *The Annals of Statistics*, 48(6):3488 – 3512, 2020. doi:10.1214/19-AOS1939.

Fabian Spaeh and Charalampos Tsourakakis. Learning Mixtures of Markov Chains with Quality Guarantees. In *Proceedings of the ACM Web Conference 2023*, WWW '23, page 662–672, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450394161. doi:10.1145/3543507.3583524.

Fabian Spaeh and Charalampos E. Tsourakakis. Markovletics: Methods and A Novel Application for Learning Continuous-Time Markov Chain Mixtures. In *Proceedings of the ACM Web Conference 2024*, WWW '24, page 4160–4171, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400701719. doi:10.1145/3589334.3645491.

Fabian Spaeh, Konstantinos Sotiropoulos, and Charalampos E. Tsourakakis. ULTRA-MC: A Unified Approach to Learning Mixtures of Markov Chains via Hitting Times. *arXiv preprint arXiv:2405.15094*, 2024. URL https://arxiv.org/abs/2405.15094.

Ji Tang, Linfeng Liu, Jiagao Wu, Jian Zhou, and Yang Xiang. Trajectory clustering method based on spatial-temporal properties for mobile social networks. *Journal of Intelligent Information Systems*, 56(1):73–95, Feb 2021. ISSN 1573-7675. doi:10.1007/s10844-020-00607-8.

Jean Tarbouriech and Alessandro Lazaric. Active Exploration in Markov Decision Processes. In *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, volume 89 of *Proceedings of Machine Learning Research*, pages 974–982. PMLR, 16–18 Apr 2019. URL https://proceedings.mlr.press/v89/tarbouriech19a.html.

Jean Tarbouriech, Shubhanshu Shekhar, Matteo Pirotta, Mohammad Ghavamzadeh, and Alessandro Lazaric. Active Model Estimation in Markov Decision Processes. In *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence (UAI)*, volume 124 of *Proceedings of Machine Learning Research*, pages 1019–1028. PMLR, 03–06 Aug 2020. URL https://proceedings.mlr.press/v124/tarbouriech20a.html.

Alexandre B. Tsybakov. *Introduction to Nonparametric Estimation*. Springer Series in Statistics. Springer New York, 2009.

Tim van Erven and Peter Harremos. Rényi Divergence and Kullback-Leibler Divergence. *IEEE Transactions on Information Theory*, 60(7):3797–3820, 2014. doi:10.1109/TIT.2014.2320500.

Thomas van Vuren, Thomas Cronk, and Jaron Sanders. Estimating the number of clusters of a Block Markov Chain. *arXiv preprint arXiv:2407.18287*, 2024. URL https://arxiv.org/abs/2407.18287.

Mathukumalli Vidyasagar. An Elementary Derivation of the Large Deviation Rate Function for Finite State Markov Chains. *Asian Journal of Control*, 16(1):1–19, 2014. doi:10.1002/asjc.806.

Ulrike von Luxburg. A tutorial on spectral clustering. *Statistics and Computing*, 17(4):395–416, 2007. doi:10.1007/s11222-007-9033-z.

Youjia Wang and Michael C.H. Choi. Information divergences of Markov chains and their applications. *arXiv preprint arXiv:2312.04863*, 2023. URL https://arxiv.org/abs/2312.04863.

Geoffrey Wolfer and Aryeh Kontorovich. Estimating the Mixing Time of Ergodic Markov Chains. In *Pro-*

ceedings of the Thirty-Second Conference on Learning Theory, volume 99 of *Proceedings of Machine Learning Research*, pages 3120–3159. PMLR, 25–28 Jun 2019. URL https://proceedings.mlr.press/v99/wolfer19a.html.

Geoffrey Wolfer and Aryeh Kontorovich. Statistical estimation of ergodic Markov chain kernel over discrete state space. *Bernoulli*, 27(1):532 – 553, 2021. doi:10.3150/20-BEJ1248.

Geoffrey Wolfer and Aryeh Kontorovich. Improved estimation of relaxation time in nonreversible Markov chains. *The Annals of Applied Probability*, 34(1A): 249 – 276, 2024. doi:10.1214/23-AAP1963.

David P. Woodruff. Sketching as a Tool for Numerical Linear Algebra. *Foundations and Trends® in Theoretical Computer Science*, 10(1–2):1–157, 2014. ISSN 1551-305X. doi:10.1561/0400000060.

Yuqing Yang, Jianghui Cai, Haifeng Yang, Jifu Zhang, and Xujun Zhao. TAD: A trajectory clustering algorithm based on spatial-temporal density analysis. *Expert Systems with Applications*, 139:112846, 2020. ISSN 0957-4174. doi:10.1016/j.eswa.2019.112846.

Bin Yu. Rates of Convergence for Empirical Processes of Stationary Mixing Sequences. *The Annals of Probability*, 22(1):94 – 116, 1994. doi:10.1214/aop/1176988849.

Se-Young Yun and Alexandre Proutière. Community Detection via Random and Adaptive Sampling. In *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 138–175, Barcelona, Spain, 13–15 Jun 2014a. PMLR. URL https://proceedings.mlr.press/v35/yun14.html.

Se-Young Yun and Alexandre Proutière. Accurate Community Detection in the Stochastic Block Model via Spectral Algorithms. *arXiv preprint arXiv:1412.7335*, 2014b. URL https://arxiv.org/abs/1412.7335.

Se-Young Yun and Alexandre Proutière. Optimal Cluster Recovery in the Labeled Stochastic Block Model. In *Advances in Neural Information Processing Systems*, volume 29, pages 973–981. Curran Associates, Inc., 2016. URL https://papers.nips.cc/paper_files/paper/2016/hash/a8849b052492b5106526b2331e526138-Abstract.html.

Se-Young Yun and Alexandre Proutière. Optimal Sampling and Clustering in the Stochastic Block Model. In *Advances in Neural Information Processing Systems*, volume 32, pages 13444–13452. Curran Associates, Inc., 2019. URL https://papers.nips.cc/paper_files/paper/2019/hash/8fe04df45a22b63156ebabbb064fcd5e-Abstract.html.

Se-Young Yun, Marc Lelarge, and Alexandre Proutière. Streaming, Memory Limited Algorithms for Community Detection. In *Advances in Neural Information Processing Systems*, volume 27, pages 3167–3175. Curran Associates, Inc., 2014. URL https://arxiv.org/abs/1411.1279.

Yang Zhou, Quan Yuan, Chao Yang, and Yinhai Wang. Who you are determines how you travel: Clustering human activity patterns with a Markov-chain-based mixture model. *Travel Behaviour and Society*, 24:102–112, 2021. ISSN 2214-367X. doi:10.1016/j.tbs.2021.03.005.

Checklist

1. For all models and algorithms presented, check if you include:
   (a) A clear description of the mathematical setting, assumptions, algorithm, and/or model. [Yes] Section 3,4
   (b) An analysis of the properties and complexity (time, space, sample size) of any algorithm. [Yes] Section 3,4
   (c) (Optional) Anonymized source code, with specification of all dependencies, including external libraries. [Yes] Appendix F.

2. For any theoretical claim, check if you include:
   (a) Statements of the full set of assumptions of all theoretical results. [Yes] Section 3,4
   (b) Complete proofs of all theoretical results. [Yes] The whole Appendix
   (c) Clear explanations of any assumptions. [Yes] Section 2, 4

3. For all figures and tables that present empirical results, check if you include:
   (a) The code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL). [Yes] Appendix F.
   (b) All the training details (e.g., data splits, hyperparameters, how they were chosen). [Yes] Appendix F.
   (c) A clear definition of the specific measure or statistics and error bars (e.g., with respect to the random seed after running experiments multiple times). [Yes] Appendix F.
   (d) A description of the computing infrastructure used. (e.g., type of GPUs, internal cluster, or cloud provider). [No] Simple CPU experiments.

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets, check if you include:
   (a) Citations of the creator If your work uses existing assets. [Yes] Appendix F
   (b) The license information of the assets, if applicable. [Not Applicable]
   (c) New assets either in the supplemental material or as a URL, if applicable. [Yes] Appendix F
   (d) Information about consent from data providers/curators. [Not Applicable]
   (e) Discussion of sensible content if applicable, e.g., personally identifiable information or offensive content. [Not Applicable]

5. If you used crowdsourcing or conducted research with human subjects, check if you include:
   (a) The full text of instructions given to participants and screenshots. [Not Applicable]
   (b) Descriptions of potential participant risks, with links to Institutional Review Board (IRB) approvals if applicable. [Not Applicable]
   (c) The estimated hourly wage paid to participants and the total amount spent on participant compensation. [Not Applicable]

Contents

1 INTRODUCTION **1**

2 PROBLEM SETTING **2**

3 FUNDAMENTAL LOWER BOUND **3**

4 TWO-STAGE ALGORITHM **4**

4.1 Stage I. Initial Spectral Clustering 4

4.2 Stage II. One-Shot Trajectory Likelihood Improvement 6

5 DISCUSSIONS **7**

6 EXPERIMENTS **9**

7 CONCLUSION **9**

A PROOF OF THEOREM 3.1 – FUNDAMENTAL LOWER BOUND **18**

B PROOF OF THEOREM 4.1 – INITIAL SPECTRAL CLUSTERING **22**

B.1 Proof of Lemma B.1 22

B.2 Proof of Lemma B.2 25

B.3 Proof of Lemma B.3 27

C PROOF OF THEOREM 4.2 – LIKELIHOOD IMPROVEMENT **28**

C.1 Proof of Proposition C.1 28

C.2 Proof of Proposition C.2 30

C.3 Proof of Lemma C.2 32

D PROOF OF PROPOSITION 5.1 – RELATIONSHIP BETWEEN DIFFERENT GAPS **35**

D.1 (a) Relationship between $\mathcal{D}_\pi$ and Kausik et al. (2023)'s α, Δ 35

D.2 (b) Relationship between $\mathcal{D}_\pi$ and $\Delta_{\mathbf{W}}$ 36

D.3 (c) Relationship between $\Delta_{\mathbf{W}}$ and Kausik et al. (2023)'s α, Δ 37

E PROOF OF PROPOSITION 5.2 – HARD INSTANCE WITH S-DEPENDENT SEPARATION ACROSS GAPS **38**

E.1 Construction of Cyclic-Bump MMC 38

E.2 Explicit Computations of the Gaps 38

E.3 Discussions 40

F DEFERRED EXPERIMENTAL DETAILS & ADDITIONAL ABLATIONS **41**

F.1 Deferred Experimental Details from Section 6 41

F.2 Ablation #1. Impact of Number of Iterations for Stage II 41
F.3 Ablation #2. Impact of S and T on Performance and Runtime 41
F.4 Ablation #3. Our Algorithm with Unknown K 43
F.5 Ablation #4. Impact of γ_{ps} 43
F.6 Additional Experiment with Last.fm 1K 44

G RELATIONSHIPS BETWEEN OUR L-EMBEDDING AND OTHER SPECTRAL EMBEDDINGS 47

H RANK OF W MAY BE MUCH SMALLER THAN THE NUMBER OF CLUSTERS K 48

I RELAXING THE KNOWLEDGE OF γ_{ps} 49

J ADDITIONAL FUTURE DIRECTIONS 50

A PROOF OF THEOREM 3.1 – FUNDAMENTAL LOWER BOUND

Let $(\varepsilon, \delta) \in [0,1] \times (0, 1/2]$, $\boldsymbol{\alpha} \in \Delta([K])$, and $T \in \mathbb{N}$ be fixed. Let $c = \varepsilon T$ be the "target" number of misclassifications. Let $f^{(0)} \triangleq f : [T] \to [K]$ be the ground-truth clustering satisfying $\boldsymbol{\alpha}^{(0)} := \boldsymbol{\alpha}(f^{(0)}) = \boldsymbol{\alpha}$, and let us denote $\mathcal{C}_k^{(0)} := (f^{(0)})^{-1}(k) = \{t \in [T] : f^{(0)}(t) = k\}$ be the set of trajectories in the k-th cluster.

We recall the following notion of distance between two clustering functions $f, g : [T] \to [K]$:

$$E_T(f, g) := \min_{\sigma \in \mathrm{Sym}(K)} \sum_{t \in [T]} \mathbb{1}[f(t) \neq \sigma(g(t))]. \tag{8}$$

This is a proper metric on a suitable space of allocations, as shown in the following proposition (its proof deferred to the end of this section):

Proposition A.1. *Let $\mathcal{F}(T, K) := [K]^T = \{f \mid f : [T] \to [K]\}$ and $\widetilde{\mathcal{F}}(T, K) := \mathcal{F}(T, K)/\sim$, where $f \sim g$ iff $\exists \sigma \in \mathrm{Sym}(K)$ s.t. $f = \sigma \circ g$. Then, $E_T(\cdot, \cdot)$ is a proper metric on $\widetilde{\mathcal{F}}(T, K)$.*

Let us arbitrarily fix $a \neq b \in [K]$ that we will optimize later. We construct M_T alternate hypotheses $\{f^{(m)} : [T] \to [K]\}_{m \in [M_n]}$ by re-allocating some trajectories from cluster b to a. We denote the probability measure and expectation w.r.t. each m-th allocation as $\mathbb{P}_m$ and $\mathbb{E}_m$, respectively.

The construction will satisfy the following conditions for all $m \in [M_T]$: denoting $\mathcal{C}_k^{(m)} := (f^{(m)})^{-1}(k)$ for $k \in [K]$,

(C1) ${\mathcal{C}_a}^{(0)} \subset \bigcap_{m \in [M_T]} {\mathcal{C}_a}^{(m)}$, $|{\mathcal{C}_a}^{(m)}| = \alpha_a n + 2c$, and $|{\mathcal{C}_b}^{(m)}| = \alpha_b n - 2c$,

(C2) $E_T(f^{(m)}, f^{(\ell)}) > 2c, \quad \forall m, \ell \in \{0\} \cup [M_T]$,

(C3) $\mathcal{C}_k^{(m)} = \mathcal{C}_k^{(0)}, \quad \forall k \in [K] \setminus \{a, b\}$.

The following lemma quantifies the minimum number of such hypotheses:

Lemma A.1. $M_T \geq 2 \left(\frac{\alpha_b T}{16ec} \right)^c$.

Proof. The combinatorial argument here is due to Yun and Proutière (2019, Appendix B), which we reproduce here for completeness.

First, note that there are total of $\binom{\alpha_b T}{2c}$ partitions satisfying **(C1)** and **(C3)** by moving $2c$ trajectories from ${\mathcal{C}_b}^{(0)}$ to ${\mathcal{C}_a}^{(m)}$. For each such partition, there are at most $\sum_{\ell=0}^{c} \binom{2c}{\ell} \binom{\alpha_b T - 2c}{\ell}$ partitions that still satisfy **(C1)** and **(C3)** but violate **(C2)**, as such partitions must be created via swapping ℓ elements from ${\mathcal{C}_b}^{(m)}$ and ℓ elements from ${\mathcal{C}_a}^{(m)} \setminus {\mathcal{C}_a}^{(0)}$ with $\ell \leq c$. Thus, we have that

$$M_T \geq \frac{\binom{\alpha_b T}{2c}}{\sum_{\ell=0}^{c} \binom{2c}{\ell} \binom{\alpha_b T - 2c}{\ell}} \overset{(*)}{\geq} \frac{\left(\frac{\alpha_b T}{2c} \right)^{2c}}{2^{2c-1} \left(\frac{e \alpha_b T}{c} \right)^c} = 2 \left(\frac{\alpha_b T}{16ec} \right)^c,$$

where at $(*)$, we use the elementary inequalities $\left(\frac{n}{m}\right)^m \leq \binom{n}{m} \leq \left(\frac{en}{m}\right)^m$ for $n \geq m \geq 1$. □

Then, for a clustering algorithm $\mathcal{A}$ that outputs a $\widehat{f}_{\mathcal{A}} : [T] \to [K]$, consider the following hypothesis testing procedure:

$$R_a(\mathcal{A}) = \operatorname*{arg\,min}_{m \in \{0\} \cup [M_T]} \min_{\sigma \in \mathrm{Sym}(K)} \left| {\mathcal{C}_a}^{(m)} \triangle \widehat{f}_{\mathcal{A}}^{-1}(\sigma(a)) \right|, \tag{9}$$

where for two sets A, B, $A \triangle B := (A \setminus B) \cup (B \setminus A)$ is their symmetric difference.

Intuitively, R_a outputs the most likely hypothesis out of $\{0\} \cup [M_T]$, given some estimated cluster for a. Let $\mathcal{E}^{(m)} := \{R_a(\mathcal{A}) = m\}$ be disjoint events across $m \in \{0\} \cup [M_T]$. Note that for each hypothesis $m \in [M_T]$, the

corresponding $\boldsymbol{\alpha}^{(m)} := \boldsymbol{\alpha}(f^{(m)})$ satisfies $\|\boldsymbol{\alpha} - \boldsymbol{\alpha}^{(m)}\|_2 = \frac{2\sqrt{2}c}{T} = 2\sqrt{2}\varepsilon \le \beta$. As $\mathcal{A}$ is $(\varepsilon, \beta, \delta)$-locally stable and any Type I/II error w.r.t. each model $\Psi_n^{(m)}$ results in at least $s+1$ misclassifications (see (**C2**)), we must have that

$$\min\left\{\mathbb{P}_0\left(\mathcal{E}^{(0)}\right), \min_{m\in[M_T]} \mathbb{P}_m\left(\mathcal{E}^{(m)}\right)\right\} \ge 1-\delta. \tag{10}$$

Furthermore, we note that

$$\mathbb{P}_0\left(\bigcup_{m\in[M_T]} \mathcal{E}^{(m)}\right) \le \delta, \tag{11}$$

as if $\mathcal{E}^{(m)}$ is true for some $m \in [M_T]$, then

$$2c < E_T(f, f^{(m)}) \le E_T(\widehat{f}_{\mathcal{A}}, f) + E_T(\widehat{f}_{\mathcal{A}}, f^{(m)}) \le E_T(\widehat{f}_{\mathcal{A}}, f) + c \Longrightarrow E_T(\widehat{f}_{\mathcal{A}}, f) > c,$$

which holds with probability at most δ due to $\mathcal{A}$ being $(\varepsilon, \beta, \delta)$-stable. Note that due to the disjointness of $\mathcal{E}^{(m)}$'s, we additionally have that

$$\frac{1}{M_T}\sum_{m\in[M_T]} \mathbb{P}_0\left(\mathcal{E}^{(m)}\right) \le \delta \Longrightarrow \min_{m\in[M_T]} \mathbb{P}_0\left(\mathcal{E}^{(m)}\right) \le \frac{\delta}{M_T}. \tag{12}$$

We now define the log-likelihood ratio of the given trajectories $\left\{\mathcal{T}_t = (s_{t,1}, \cdots, s_{t,H})\right\}_{t\in[T]}$ under the true model $(m=0)$ and the m-th alternate model:

$$\mathcal{Q}_T^{(m)} := \log \frac{\mathbb{P}_m(\{\mathcal{T}_t\}_{t\in[T]})}{\mathbb{P}_0(\{\mathcal{T}_t\}_{t\in[T]})} = \sum_{t\in\mathcal{C}_a{}^{(m)}\setminus\mathcal{C}_a{}^{(0)}} \log \frac{\mathbb{P}^{(a)}(s_{t,1}, \cdots, s_{t,H})}{\mathbb{P}^{(b)}(s_{t,1}, \cdots, s_{t,H})}. \tag{13}$$

Taking the expectation w.r.t. the m-th alternate model, we have

$$\begin{aligned}
\mathbb{E}_m\left[\mathcal{Q}_T^{(m)}\right] &= \sum_{t\in\mathcal{C}_a{}^{(m)}\setminus\mathcal{C}_a{}^{(0)}} \mathbb{E}_m\left[\log \frac{\mathbb{P}^{(a)}(s_{t,1}, \cdots, s_{t,H})}{\mathbb{P}^{(b)}(s_{t,1}, \cdots, s_{t,H})}\right] \\
&= \sum_{t\in\mathcal{C}_a{}^{(m)}\setminus\mathcal{C}_a{}^{(0)}} \mathbb{E}_m\left[\log \frac{\mu^{(a)}(s_{t,1})}{\mu^{(b)}(s_{t,1})} + \sum_{h=1}^{H-1} \log \frac{p^{(a)}(s_{t,h+1}|s_{t,h})}{p^{(b)}(s_{t,h+1}|s_{t,h})}\right] \\
&= 2c\left[\mathsf{KL}\left(\mu^{(a)}, \mu^{(b)}\right) + \sum_{h=1}^{H-1}\sum_{s\in\mathcal{S}} \mathbb{P}^{(a)}(s_{t,h}=s)\,\mathsf{KL}\left(p^{(a)}(\cdot|s), p^{(b)}(\cdot|s)\right)\right] \\
&\qquad\qquad \text{(Law of total expectation w.r.t. } s_{t,h}) \\
&= 2(H-1)c\underbrace{\left[\frac{1}{H-1}\mathsf{KL}\left(\mu^{(a)}, \mu^{(b)}\right) + \sum_{s\in\mathcal{S}} \mathbb{P}_H^{(a)}(s)\,\mathsf{KL}\left(p^{(a)}(\cdot|s), p^{(b)}(\cdot|s)\right)\right]}_{=:\mathcal{D}^{(a,b)}}. \\
&\qquad\qquad \left(\mathbb{P}_H^{(a)}(s) := \tfrac{1}{H-1}\textstyle\sum_{h=1}^{H-1} \mathbb{P}^{(a)}(s_{t,h}=s)\right)
\end{aligned}$$

There are two paths from here, one inspired by the arguments of Jedra and Proutière (2023) that utilizes data processing inequality, and another inspired by Yun and Proutière (2019) that utilizes Markov's inequality. The former leads to the non-asymptotic guarantee, while the latter leads to the asymptotic guarantee.

Non-asymptotic guarantee. We recall the KL-version of the data processing inequality[6]:

[6]This has been referred to as the "Fundamental Inequality" in Garivier et al. (2019), as using this allows for a strikingly simple proof of distribution-dependent bandit regret lower bounds.

Lemma A.2 (Data Processing Inequality; Lemma 1 of Garivier et al. (2019))**.** *Consider a measurable space* $(\Gamma, \mathcal{G})$ *equipped with two probability measure* $\mathbb{P}_1$ *and* $\mathbb{P}_2$*. Then, we have that*

$$\mathsf{KL}(\mathbb{P}_1, \mathbb{P}_2) \geq \sup_Z \mathsf{kl}(\mathbb{E}_1[Z], \mathbb{E}_2[Z]), \tag{14}$$

where $\sup_Z$ *is over all possible* $\mathcal{G}$*-measurable random variable* $Z : \Omega \to [0,1]$ *and* kl *is the Bernoulli KL divergence, i.e.,* $\mathsf{kl}(p,q) := p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q}$ *for* $p, q \in (0,1)$*.*

In the above lemma, let $(\Gamma, \mathcal{G})$ be our observation space of T trajectories, $\mathbb{P}_1 = \mathbb{P}_m$, $\mathbb{P}_2 = \mathbb{P}_0$, and $Z = \mathbb{1}\{\mathcal{E}^{(m)}\}$. Then, we have that

$$\begin{aligned}
2(H-1)c\mathcal{D}^{(a,b)} &= \frac{1}{M_T} \sum_{m \in [M_T]} \mathsf{KL}\left(\mathbb{P}_m, \mathbb{P}_0\right) \\
&\geq \frac{1}{M_T} \sum_{m \in [M_T]} \mathsf{kl}\left(\mathbb{P}_m(\mathcal{E}^{(m)}), \mathbb{P}_0(\mathcal{E}^{(m)})\right) && \text{(Lemma A.2)} \\
&\geq \mathsf{kl}\left(\frac{1}{M_T} \sum_{m \in [M_T]} \mathbb{P}_m(\mathcal{E}^{(m)}), \frac{1}{M_T} \sum_{m \in [M_T]} \mathbb{P}_0\left(\mathcal{E}^{(m)}\right)\right) && \text{(joint convexity of } \mathsf{kl}) \\
&= \mathsf{kl}\left(\frac{1}{M_T} \sum_{m \in [M_T]} \mathbb{P}_m(\mathcal{E}^{(m)}), \frac{1}{M_T} \mathbb{P}_0\left(\bigcup_{m \in [M_T]} \mathcal{E}^{(m)}\right)\right) && (\mathcal{E}^{(m)}\text{'s are disjoint}) \\
&\overset{(*)}{\geq} \mathsf{kl}\left(1-\delta, \frac{\delta}{M_T}\right) \\
&\overset{(**)}{\geq} (1-\delta) \log \frac{M_T}{\delta} - \log 2 \\
&\geq \frac{1}{2} \log \frac{1}{\delta} + \frac{c}{2} \log \frac{\alpha_b T}{16ec} - \frac{1}{2} \log 2 && \text{(Lemma A.1, } \delta \in (0, 1/2]) \\
&\geq \frac{1}{2} \log \frac{1}{2\delta} + \frac{c}{2} \log \frac{\alpha_{\min} T}{16ec}.
\end{aligned}$$

$(*)$ follows from $\mathbb{P}_m(\mathcal{E}^{(m)}) \geq 1 - \delta$ (as the algorithm is $(\varepsilon, \beta, \delta)$-stable), Eqn. (11), and the following well-known properties of kl:

- $q \mapsto \mathsf{kl}(p,q)$ is increasing in $[p, 1]$, as

$$\partial_q \mathsf{kl}(p,q) = -\frac{p}{q} + \frac{1-p}{1-q} = \frac{q-p}{q(1-q)} \geq 0,$$

- $p \mapsto \mathsf{kl}(p,q)$ is increasing on $p \in [1/2, 1]$ when $q < 1/2$, as

$$\partial_p \mathsf{kl}(p,q) = \log \frac{p}{q} - \log \frac{1-p}{1-q} = \log \frac{p}{1-p} + \frac{1-q}{q} \geq 0.$$

$(**)$ follows from $\mathsf{kl}(p,q) \geq p \log \frac{1}{q} - \log 2$ (Eqn. (11) of Garivier et al. (2019)).

We conclude by rearranging the resulting inequality in terms of δ and optimizing over $a \neq b \in [K]$.

□(for non-asymptotic)

Asymptotic guarantee. Let $m^* = \arg\min_{m \in [M_T]} \mathbb{P}_0[\mathcal{E}^{(m)}]$. Then,

$$\mathbb{P}_{m^*}[\mathcal{Q}_T^{(m^*)} \leq \log M_T] = \mathbb{P}_{m^*}\left[\mathcal{Q}_T^{(m^*)} \leq \log M_T, \mathcal{E}^{(m^*)}\right] + \mathbb{P}_{m^*}\left[\mathcal{Q}_T^{(m^*)} \leq \log M_T, (\mathcal{E}^{(m^*)})^{\complement}\right]$$

$$
\begin{aligned}
&\leq \exp(\log M_T)\mathbb{P}_0\left[\mathcal{Q}_T^{(m^*)} \leq \log M_T, \mathcal{E}^{(m^*)}\right] + \mathbb{P}_{m^*}\left[(\mathcal{E}^{(m^*)})^{\complement}\right] \\
&\qquad\qquad \text{(change of measure from the alternate model } m^* \text{ to the true model)} \\
&\leq M_T\mathbb{P}_0[\mathcal{E}^{(m^*)}] + \delta_T \\
&\leq 2\delta_T, \qquad\qquad \text{(Eqn. (12))}
\end{aligned}
$$

where $(\delta_T)_{T\in\mathbb{N}}$ is such that $\delta_T \to 0$ as $T \to \infty$.

Thus, by Markov's inequality,

$$
\begin{aligned}
1 - 2\delta_T &\leq \mathbb{P}_{m^*}\left[\mathcal{Q}_T^{(m^*)} > \log M_T\right] \\
&\leq \frac{2(H-1)c\mathcal{D}^{(a,b)}}{\log M_T} \\
&\leq \frac{2(H-1)c\mathcal{D}^{(a,b)}}{c\log\frac{\alpha_b T}{16ec} + \log 2}. \qquad \text{(Lemma A.1)} \\
&\leq \frac{2(H-1)c\mathcal{D}^{(a,b)}}{c\log\frac{\alpha_{\min} T}{16ec} + \log 2}.
\end{aligned}
$$

Taking the min over $a \neq b \in [K]$ and the limit $n \to \infty$, we have that

$$
\liminf_{T\to\infty} \frac{2(H-1)\mathcal{D}}{\log\frac{\alpha_{\min}}{16e\varepsilon}} \geq 1. \tag{15}
$$

□(for asymptotic)

Proof of Proposition A.1. Positivity and symmetricity are trivial. Thus, it remains to show the triangle inequality. Let $f, g, h \in \mathcal{F}(T, K)$. Then,

$$
\begin{aligned}
E_T(f,h) &= \min_{\sigma\in\mathrm{Sym}(K)} \sum_{t\in[T]} \mathbb{1}[f(t) \neq \sigma(h(t))] \\
&\leq \min_{\nu\in\mathrm{Sym}(K)} \min_{\sigma\in\mathrm{Sym}(K)} \sum_{t\in[T]} \mathbb{1}[f(t) \neq \nu(g(t)) \vee \nu(g(t)) \neq \sigma(h(t))] \\
&\leq \min_{\nu\in\mathrm{Sym}(K)} \min_{\sigma\in\mathrm{Sym}(K)} \sum_{t\in[T]} \left(\mathbb{1}[f(t) \neq \nu(g(t))] + \mathbb{1}[\nu(g(t)) \neq \sigma(h(t))]\right) \\
&\leq \min_{\nu\in\mathrm{Sym}(K)} \sum_{t\in[T]} \mathbb{1}[f(t) \neq \nu(g(t))] + \min_{\nu\in\mathrm{Sym}(K)} \min_{\sigma\in\mathrm{Sym}(K)} \sum_{t\in[T]} \mathbb{1}[\nu(g(t)) \neq \sigma(h(t))] \\
&= E_T(f,g) + \min_{\nu'\in\mathrm{Sym}(K)} \mathbb{1}[g(t) \neq \nu'(h(t))] \qquad (\nu' = \nu^{-1}\circ\sigma) \\
&= E_T(f,g) + E_T(g,h).
\end{aligned}
$$

□

B PROOF OF THEOREM 4.1 – INITIAL SPECTRAL CLUSTERING

We remark in advance that the proofs of all supporting lemmas and propositions are deferred to the end of this section.

We begin by establishing the following concentration result for $\widehat{\boldsymbol{W}}$:

Lemma B.1 (Concentration of $\widehat{\boldsymbol{W}}$ to $\boldsymbol{W}$)**.** *Let $\delta \in (0,1)$. For each $s \in \mathcal{S}$ and $t \in [T]$, the following holds with probability at least $1-\delta$, given that $H \geq \left(\frac{25}{1-\pi(s)} \vee 52\right) \frac{2}{\pi(s)\gamma_{\text{ps}}} \log \frac{8}{\pi_{\min}\delta^2}$:*

$$\left\| \sqrt{\pi^{(f(t))}(s)} p^{(f(t))}(\cdot|s) - \frac{1}{\sqrt{H\widehat{N}_t(s)}} \widehat{N}_t(s,\cdot) \right\|_2 \leq 4\sqrt{\frac{2}{H\gamma_{\text{ps}}} \log \frac{8(H\pi(s)+1)}{\pi_{\min}\delta^2}}. \tag{16}$$

If we additionally assume that $H \geq \frac{9\pi(s)}{\pi_{\min}}$, the following holds with probability at least $1-\delta$:

$$\left|(\boldsymbol{W})_{t,:} - (\widehat{\boldsymbol{W}})_{t,:}\right| \leq \widehat{r} \triangleq 8\sqrt{\frac{S}{H\gamma_{\text{ps}}} \log \frac{H}{\delta}} \implies \left\|\boldsymbol{W} - \widehat{\boldsymbol{W}}\right\|_{2\to\infty} \leq 8\sqrt{\frac{S}{H\gamma_{\text{ps}}} \log \frac{TH}{\delta}}. \tag{17}$$

First, as $\left\|\boldsymbol{W} - \widehat{\boldsymbol{W}}\right\|_2 \leq \sqrt{T}\left\|\boldsymbol{W} - \widehat{\boldsymbol{W}}\right\|_{2\to\infty}$ (Cape et al., 2019, Proposition 6.3), with our choice of $\widehat{\sigma}_{\text{thres}} := 8\sqrt{\frac{TS}{H\gamma_{\text{ps}}} \log \frac{TH}{\delta}}$ (line 3), we have from Lemma B.1 and Weyl's inequality for singular values (Horn and Johnson, 2012, Problem 7.3.P16) that $\mathbb{P}\left(\widehat{R} \leq R = \text{rank}(\boldsymbol{W})\right) \geq 1-\delta$. We will condition on this event throughout the remainder of the proof.

Then, we have that

$$\begin{aligned}
\left\|\widehat{\boldsymbol{X}}\widehat{\boldsymbol{V}}_{1:R}^\top - \boldsymbol{W}\right\|_F^2 &\leq 2R\left\|\widehat{\boldsymbol{X}}\widehat{\boldsymbol{V}}_{1:R}^\top - \boldsymbol{W}\right\|_2^2 \\
&\leq 4R\left(\left\|\widehat{\boldsymbol{X}}\widehat{\boldsymbol{V}}_{1:R}^\top - \widehat{\boldsymbol{W}}\right\|_2^2 + \left\|\widehat{\boldsymbol{W}} - \boldsymbol{W}\right\|_2^2\right) \\
&\leq 8R(\widehat{\sigma}_{\text{thres}})^2, && \text{(with probability at least } 1-\delta) \\
&= 2^9 R\frac{TS}{H\gamma_{\text{ps}}} \log \frac{TH}{\delta}.
\end{aligned}$$

Let us denote $\widehat{\boldsymbol{Y}} := \widehat{\boldsymbol{X}}\widehat{\boldsymbol{V}}_{1:R}$. We now have the following lemma:

Lemma B.2. *Additionally assume that $H \gtrsim \frac{S}{\Delta_{\boldsymbol{W}}^2\gamma_{\text{ps}}} \log \frac{H}{\delta} \log \frac{TH}{\delta}$. Then, if $t \in [T]$ is misclassified in Algorithm 1, then $\left\|\widehat{\boldsymbol{Y}}_{t,:} - \boldsymbol{W}_{t,:}\right\|_2 > \frac{\Delta_{\boldsymbol{W}}}{4}$.*

Then, we have that with probability at least $1-\delta$,

$$E_T(\widehat{f}_0, f) \leq \frac{\left\|\widehat{\boldsymbol{Y}} - \boldsymbol{W}\right\|_F^2}{\frac{\Delta_{\boldsymbol{W}}^2}{2^4}} \leq \frac{2^{13}TRS}{H\gamma_{\text{ps}}\Delta_{\boldsymbol{W}}^2} \log \frac{TH}{\delta}.$$

□

B.1 Proof of Lemma B.1

Note that

$$\left\| \sqrt{\pi^{(f(t))}(s)} p^{(f(t))}(\cdot|s) - \frac{1}{\sqrt{H\widehat{N}_t(s)}} \widehat{N}_t(s,\cdot) \right\|_2$$

$$\leq \sqrt{\pi^{(f(t))}(s)}\left\|p^{(f(t))}(\cdot|s) - \frac{\widehat{N}_t(s,\cdot)}{\widehat{N}_t(s)}\right\|_2 + \left|\sqrt{\pi^{(f(t))}(s)} - \sqrt{\frac{\widehat{N}_t(s)}{H}}\right|\left\|\frac{\widehat{N}_t(s,\cdot)}{\widehat{N}_t(s)}\right\|_2$$

$$\leq \sqrt{\pi^{(f(t))}(s)}\underbrace{\left\|p^{(f(t))}(\cdot|s) - \frac{\widehat{N}_t(s,\cdot)}{\widehat{N}_t(s)}\right\|_2}_{(i)} + \underbrace{\left|\sqrt{\pi^{(f(t))}(s)} - \sqrt{\frac{\widehat{N}_t(s)}{H}}\right|}_{(ii)}\underbrace{\left\|\frac{\widehat{N}_t(s,\cdot)}{\widehat{N}_t(s)}\right\|_1}_{=1}$$

Let us bound (i) and (ii) separately. For simplicity, we omit the dependency on $t \in [T]$.

Bounding (i)**.** Here, we follow the proof strategy of Wolfer and Kontorovich (2021, Theorem 3.1), with two crucial modifications: the quantity being bounded is the ℓ_2-norm (not ℓ_1-norm) error, and we only need to bound the error stemming from a single state s.

Let $n_s \in \mathbb{N}$ to be determined later, and let $\mathbb{P}_\pi$ be the probability measure induced when the Markov chain starts from its stationary distribution. Then,

$$\mathbb{P}_\pi\left(\left\|p(\cdot|s) - \frac{\widehat{N}(s,\cdot)}{\widehat{N}(s)}\right\|_2 > \varepsilon\right) \leq \sum_{n=n_s}^{3n_s}\mathbb{P}_\pi\left(\left\|p(\cdot|s) - \frac{\widehat{N}(s,\cdot)}{\widehat{N}(s)}\right\|_2 > \varepsilon,\ \widehat{N}(s) = n_s\right) + \mathbb{P}_\pi\left(\widehat{N}(s) \notin [\![n_s, 3n_s]\!]\right).$$

For the first term, we utilize the scheme of reducing learning the Markovian transition kernel to learning a discrete distribution, as described in the proof of Billingsley (1961, Theorem 3.1). Then, denoting $\widehat{p}_n(\cdot|s)$ to be the estimate of $p(\cdot|s)$ from n i.i.d. samples of $p(\cdot|s)$, we have that

$$\begin{aligned}\sum_{n=n_s}^{3n_s}\mathbb{P}_\pi\left(\left\|p(\cdot|s) - \frac{\widehat{N}(s,\cdot)}{\widehat{N}(s)}\right\|_2 > \varepsilon,\ \widehat{N}(s) = n\right) &\leq \sum_{n=n_s}^{3n_s}\mathbb{P}_\pi\left(\left\|p(\cdot|s) - \widehat{p}_n(\cdot|s)\right\|_2 > \varepsilon\right)\\ &\overset{(*)}{\leq} \sum_{n=n_s}^{3n_s}\exp\left(-\frac{\varepsilon^2 n}{2}\right)\\ &\leq (2n_s+1)\exp\left(-\frac{\varepsilon^2 n_s}{2}\right),\end{aligned}$$

where $(*)$ follows from the following ℓ_2-concentration of learning discrete distribution:

Lemma B.3 (Theorem 9 of Canonne (2020))**.** *Let $\mathcal{S}$ be a finite state space, $p \in \mathcal{P}(\mathcal{S})$, and $\delta \in (0,1)$. We are given $\{s_i\}_{i\in[N]}$ with $s_i \overset{i.i.d.}{\sim} p$. Let $\hat{p}_N \in \mathcal{P}(\mathcal{S})$ be defined as $\hat{p}_N(s) := \frac{1}{N}\sum_{i\in[N]}\mathbb{1}[s_i = s]$ for each $s \in \mathcal{S}$. Then, we have the following:*

$$\mathbb{P}\left(\|p - \hat{p}_N\|_2 \geq \varepsilon\right) \leq \exp\left(-\frac{\varepsilon^2 N}{2}\right), \tag{18}$$

We now choose $n_s = \frac{H\pi(s)}{2}$ and utilize the Markovian Bernstein concentration, which we recall here:

Lemma B.4 (Theorem 3.4 of Paulin (2015))**.** *Suppose of $(X_h)_{h\geq 1}$ is an ergodic Markov chain over $\mathcal{S}$ with transition probability p, initial distribution μ, and pseudo spectral gap γ_{ps}. Let $\phi \in L^2(\pi)$ with $\sup_{s\in\mathcal{S}}|\phi(s) - \mathbb{E}_\pi[\phi]| \leq C$ for some $C > 0$, and $V_\phi := \text{Var}_\pi[\phi]$. Then, we have the following: for any $u \geq 0$,*

$$\mathbb{P}\left(\sum_{h=1}^{H}(\phi(X_h) - \mathbb{E}_\pi[\phi(X_h)]) \geq u\right) \leq \exp\left(-\frac{u^2\gamma_{\text{ps}}}{8\left(H + \frac{1}{\gamma_{\text{ps}}}\right)V_\phi + 20Cu}\right). \tag{19}$$

For two integers $a \geq b$, denote $[\![a,b]\!] := \{a, a+1, \cdots, b\}$. We then have that

$$\mathbb{P}_\pi\left(\widehat{N}(s) \notin \left[\!\!\left[\frac{H\pi(s)}{2}, \frac{3H\pi(s)}{2}\right]\!\!\right]\right) = \mathbb{P}_\pi\left(\widehat{N}(s) - H\pi(s) \notin \left[\!\!\left[-\frac{H\pi(s)}{2}, \frac{H\pi(s)}{2}\right]\!\!\right]\right)$$

$$
\begin{aligned}
&\leq \mathbb{P}_\pi\left(\left|\widehat{N}(s) - H\pi(s)\right| > \frac{H\pi(s)}{2}\right) \\
&\leq 2\exp\left(-\frac{(H\pi(s)/2)^2\gamma_{\mathrm{ps}}}{8\left(H + \frac{1}{\gamma_{\mathrm{ps}}}\right)\pi(s)(1-\pi(s)) + 20(H\pi(s)/2)}\right) \\
&\leq 2\exp\left(-\frac{H\pi(s)\gamma_{\mathrm{ps}}}{104}\right). \qquad \text{(Assume that } H > \tfrac{1}{\gamma_{\mathrm{ps}}}\text{)}
\end{aligned}
$$

Combining everything, we have that

$$
\mathbb{P}_\pi\left(\left\|p(\cdot|s) - \frac{\widehat{N}(s,\cdot)}{\widehat{N}(s)}\right\|_2 > \varepsilon\right) \leq (H\pi(s)+1)\exp\left(-\frac{\varepsilon^2 H\pi(s)}{2}\right) + 2\exp\left(-\frac{H\pi(s)\gamma_{\mathrm{ps}}}{104}\right),
$$

and thus, with $H \geq \frac{1}{\gamma_{\mathrm{ps}}} \vee \frac{104}{\pi(s)\gamma_{\mathrm{ps}}}\log\frac{2}{\delta}$, we have that

$$
\mathbb{P}_\pi\left(\left\|p(\cdot|s) - \frac{\widehat{N}(s,\cdot)}{\widehat{N}(s)}\right\|_2 > \sqrt{\frac{2}{H\pi(s)}\log\frac{2(H\pi(s)+1)}{\delta}}\right) \leq \delta. \tag{20}
$$

Here, we recall the following result that quantifies the price of non-stationarity due to the initial distribution μ not necessarily being the stationary distribution π:

Lemma B.5 (Proposition 3.10 of Paulin (2015))**.** *Let $(X_1, \cdots, X_H)$ be a (time-homogeneous) Markov chain with stationary distribution π. For any measurable $g : \mathcal{S}^H \to \mathbb{R}$ and initial distribution $\mu \in \Delta(\mathcal{S})$, we have that for any $u \geq 0$,*

$$
\mathbb{P}_\mu\left(g(X_1, \cdots, X_H) \geq u\right) \leq \sqrt{\left[\sum_{s\in\mathcal{S}} \frac{\mu(s)^2}{\pi(s)}\right]\mathbb{P}_\pi\left(g(X_1, \cdots, X_H) \geq u\right)}. \tag{21}
$$

Especially as $|\mathcal{S}| < \infty$, we can bound $\sum_{s\in\mathcal{S}} \frac{\mu(s)^2}{\pi(s)} \leq \frac{1}{\pi_{\min}}$.

Thus,

$$
\mathbb{P}\left(\left\|p(\cdot|s) - \frac{\widehat{N}(s,\cdot)}{\widehat{N}(s)}\right\|_2 > \sqrt{\frac{2}{H\pi(s)}\log\frac{2(H\pi(s)+1)}{\delta}}\right) \leq \sqrt{\frac{\delta}{\pi_{\min}}}.
$$

Reparametrizing finally gives

$$
\mathbb{P}\left(\left\|p(\cdot|s) - \frac{\widehat{N}(s,\cdot)}{\widehat{N}(s)}\right\|_2 > \sqrt{\frac{2}{H\pi(s)}\log\frac{8(H\pi(s)+1)}{\pi_{\min}\delta^2}}\right) \leq \frac{\delta}{2}, \tag{22}
$$

given that $H \geq \frac{1}{\gamma_{\mathrm{ps}}} \vee \frac{104}{\pi(s)\gamma_{\mathrm{ps}}}\log\frac{8}{\pi_{\min}\delta^2}$.

Bounding *(ii)***.** Note that

$$
\left|\sqrt{\pi(s)} - \sqrt{\frac{\widehat{N}(s)}{H}}\right| = \frac{\left|\pi(s) - \frac{\widehat{N}(s)}{H}\right|}{\sqrt{\pi(s)} + \sqrt{\frac{\widehat{N}(s)}{H}}} \leq \frac{1}{\sqrt{\pi(s)}}\left|\pi(s) - \frac{\widehat{N}(s)}{H}\right|.
$$

Again, we invoke Lemma B.4:

$$
\mathbb{P}_\pi\left(\left|\pi(s) - \frac{\widehat{N}(s)}{H}\right| > \sqrt{\frac{32\pi(s)(1-\pi(s))}{H\gamma_{\mathrm{ps}}}\log\frac{8}{\pi_{\min}\delta^2}}\right)
$$

$$
\begin{aligned}
&\le 2\exp\left(-\frac{32\pi(s)(1-\pi(s))H\log\frac{8}{\pi_{\min}\delta^2}}{8\left(H+\frac{1}{\gamma_{\mathrm{ps}}}\right)\pi(s)(1-\pi(s))+20\sqrt{\frac{32\pi(s)(1-\pi(s))H}{\gamma_{\mathrm{ps}}}\log\frac{8}{\pi_{\min}\gamma_{\mathrm{ps}}}}}\right)\\
&\le 2\exp\left(-\frac{2\sqrt{H\pi(s)(1-\pi(s))}\log\frac{8}{\pi_{\min}\delta^2}}{\sqrt{H\pi(s)(1-\pi(s))}+5\sqrt{\frac{2}{\gamma_{\mathrm{ps}}}\log\frac{8}{\pi_{\min}\delta^2}}}\right) && (H\ge\tfrac{1}{\gamma_{\mathrm{ps}}})\\
&\le 2\exp\left(-\frac{2\sqrt{H\pi(s)(1-\pi(s))}\log\frac{8}{\pi_{\min}\delta^2}}{2\sqrt{H\pi(s)(1-\pi(s))}}\right) && (H\ge\tfrac{50}{\pi(s)(1-\pi(s))\gamma_{\mathrm{ps}}}\log\tfrac{8}{\pi_{\min}\delta^2})\\
&=\frac{\pi_{\min}\delta^2}{4}.
\end{aligned}
$$

Using Lemma B.5, we then have

$$
\mathbb{P}\left(\left|\pi(s)-\frac{\widehat{N}(s)}{H}\right|>\sqrt{\frac{32\pi(s)(1-\pi(s))}{H\gamma_{\mathrm{ps}}}\log\frac{8}{\pi_{\min}\delta^2}}\right)\le\frac{\delta}{2}. \tag{23}
$$

Combining everything. By union bound, we have that with probability at least $1-\delta$,

$$
\begin{aligned}
\left\|\sqrt{\pi^{(f(t))}(s)}p^{(f(t))}(\cdot|s)-\frac{1}{\sqrt{H\widehat{N}_t(s)}}\widehat{N}_t(s,\cdot)\right\|_2 &\le\sqrt{\frac{2}{H}\log\frac{8(H\pi(s)+1)}{\pi_{\min}\delta^2}}+\sqrt{\frac{32(1-\pi(s))}{H\gamma_{\mathrm{ps}}}\log\frac{8}{\pi_{\min}\delta^2}}\\
&\le 4\sqrt{\frac{2}{H\gamma_{\mathrm{ps}}}\log\frac{8(H\pi(s)+1)}{\pi_{\min}\delta^2}}.
\end{aligned} \tag{24}
$$

□

B.2 Proof of Lemma B.2

Here, we are largely inspired by the proof strategies of Yun and Proutière (2014a, Appendix C) and Yun and Proutière (2016, Appendix C.2).

We start by defining the following sets:

$$
\mathcal{I}_k:=\left\{t\in f^{-1}(k):\left\|\widehat{\boldsymbol{Y}}_{t,:}-\boldsymbol{W}_{t,:}\right\|_2^2\le\frac{1}{4}(\widehat{r})^2\log\frac{TH}{\delta}\right\},\quad k\in[K] \tag{25}
$$

$$
\mathcal{O}:=\left\{t\in[T]:\min_{t'\in[T]}\left\|\widehat{\boldsymbol{Y}}_{t,:}-\boldsymbol{W}_{t',:}\right\|_2^2\ge 4(\widehat{r})^2\log\frac{TH}{\delta}\right\}. \tag{26}
$$

Intuitively, $\mathcal{I}_k$ is the subset of trajectories of $f^{-1}(k)$ that will be classified accurately (with high probability), and $\mathcal{O}$ is the set of trajectories that are clustered at the end (i.e., lines 12-13).

Then the following properties hold:

(i) $\left(\bigcup_{k=1}^K\mathcal{I}_k\right)\cap\mathcal{Q}_t=\emptyset$ for any $t\in\mathcal{O}$. This is because for any $t'\in\mathcal{I}_k$,

$$
\begin{aligned}
\left\|\widehat{\boldsymbol{Y}}_{t,:}-(\widehat{\boldsymbol{Y}})_{t',:}\right\|_2^2 &\ge\frac{1}{2}\left\|\widehat{\boldsymbol{Y}}_{t,:}-\boldsymbol{W}_{t,:}\right\|_2^2-\left\|(\widehat{\boldsymbol{Y}})_{t',:}-\boldsymbol{W}_{t,:}\right\|_2^2 && ((a+b)^2\le 2(a^2+b^2))\\
&\ge 2(\widehat{r})^2\log\frac{TH}{\delta}-\frac{1}{4}(\widehat{r})^2\log\frac{TH}{\delta} && (\text{Definition of }\mathcal{O}\text{ and }\mathcal{I}_k)\\
&=\frac{7}{4}(\widehat{r})^2\log\frac{TH}{\delta}\ge(\widehat{r})^2\log\frac{TH}{\delta},
\end{aligned}
$$

i.e., $t'\in\mathcal{Q}_t\Rightarrow t'\notin\bigcup_{k=1}^{K_s}\mathcal{I}_k$.

(ii) $\underline{\left|[T] \setminus \cup_{k=1}^{K} \mathcal{I}_k\right| \leq \frac{RT}{\log \frac{TH}{\delta}}}$. This is because

$$\left|[T] \setminus \cup_{k=1}^{K} \mathcal{I}_k\right| \leq \frac{\left\|\widehat{\boldsymbol{Y}} - \boldsymbol{W}\right\|_F^2}{\min_{t \in [T] \setminus \cup_{k=1}^{K} \mathcal{I}_k} \left\|\widehat{\boldsymbol{Y}}_{t,:} - \boldsymbol{W}_{t,:}\right\|_2^2} \leq \frac{8RT(\widehat{r})^2}{\frac{1}{4}(\widehat{r})^2 \log \frac{TH}{\delta}} = \frac{32RT}{\log \frac{TH}{\delta}}$$

(iii) $\underline{\mathcal{I}_k \subseteq \mathcal{Q}_t \text{ for any } t \in \mathcal{I}_k.}$ This is because for any $t, t' \in \mathcal{I}_k$,

$$\left\|\widehat{\boldsymbol{Y}}_{t',:} - \widehat{\boldsymbol{Y}}_{t,:}\right\|_2^2 \leq 2\left\|\widehat{\boldsymbol{Y}}_{t',:} - \boldsymbol{W}_{t',:}\right\|_2^2 + 2\left\|\widehat{\boldsymbol{Y}}_{t,:} - \boldsymbol{W}_{t,:}\right\|_2^2 \leq (\widehat{r})^2 \log \frac{TH}{\delta}$$

(iv) $\underline{\text{If } \mathcal{Q}_t \cap \mathcal{I}_k \neq \emptyset, \text{ then } \mathcal{Q}_t \cap \bigcup_{k' \neq k} \mathcal{I}_k = \emptyset.}$ We show this via *reductio ad absurdum*. Suppose there exists $k \neq k'$ such that $\mathcal{Q}_t \cap \mathcal{I}_k \neq \emptyset$ and $\mathcal{Q}_t \cap \mathcal{I}_{k'} \neq \emptyset$. Let $t_k \in \mathcal{Q}_t \cap \mathcal{I}_k$ and $t_{k'} \in \mathcal{Q}_t \cap \mathcal{I}_{k'}$. Then, by definition of $\mathcal{Q}_t$,

$$\begin{aligned}
\widehat{r}\sqrt{\log \frac{TH}{\delta}} &\geq \left\|\widehat{\boldsymbol{Y}}_{t_k,:} - \widehat{\boldsymbol{Y}}_{t_{k'},:}\right\|_2 \\
&\geq \left\|\boldsymbol{W}_{t_k,:} - \boldsymbol{W}_{t_{k'},:}\right\|_2 - \left\|\widehat{\boldsymbol{Y}}_{t_k,:} - \boldsymbol{W}_{t_k,:}\right\|_2 - \left\|\widehat{\boldsymbol{Y}}_{t_{k'},:} - \boldsymbol{W}_{t_{k'},:}\right\|_2 \\
&\geq \Delta_{\boldsymbol{W}} - \widehat{r}\sqrt{\log \frac{TH}{\delta}}.
\end{aligned}$$

Recalling the definition of $\widehat{r}$, one can easily check that if $H \geq 2^8 \frac{S}{\Delta_{\boldsymbol{W}}^2 \gamma_{\mathrm{ps}}} \log \frac{H}{\delta} \log \frac{TH}{\delta}$, above cannot be true, a contradiction.

Claim B.1. $\widehat{K} = K$.

Proof. Without loss of generality, assume that $|\mathcal{I}_1| \geq |\mathcal{I}_2| \geq \cdots \geq |\mathcal{I}_K|$.

We first show that for each $k \in [K]$,

$$\exists t_k^\star \in \bigcup_{k' \in [K]} \mathcal{I}_{k'} \setminus \bigcup_{\ell=1}^{k-1} \mathcal{S}_\ell \quad \text{s.t.} \quad \left|\mathcal{Q}_{t_k^\star} \setminus \bigcup_{\ell=0}^{k-1} \mathcal{S}_\ell\right| \geq |\mathcal{I}_k|.$$

Due to the properties (iii) and (iv), and the greedy nature of lines 7-10, the above is indeed true.

Let $\{t_1^\star, \cdots, t_K^\star\}$ be the selected "centers" with $\mathcal{I}_k \subseteq \mathcal{Q}_{t_k^\star}$.

Now, we show that after k has reached $K+1$ in line 10, the **while** loop terminates. By property (ii), the number of remaining trajectories is

$$\left|[T] \setminus \bigcup_{k \in [K]} \mathcal{Q}_{t_k^\star}\right| \leq \left|[T] \setminus \bigcup_{k \in [K]} \mathcal{I}_k\right| \overset{(ii)}{\leq} \frac{32RT}{\log \frac{TH}{\delta}},$$

which is precisely the termination criterion at line 7. □

Thus, if a trajectory t is misclassified in the sense that it gets assigned to $t_{k'}^\star$ instead of $t_k^\star$, then it must be that $\left\|\widehat{\boldsymbol{Y}}_{t,:} - \widehat{\boldsymbol{Y}}_{t_k^\star}\right\|_2 > \left\|\widehat{\boldsymbol{Y}}_{t,:} - \widehat{\boldsymbol{Y}}_{t_{k'}^\star}\right\|_2$. By the triangle inequality, this then implies that

$$\begin{aligned}
\left\|\widehat{\boldsymbol{Y}}_{t,:} - \boldsymbol{W}_{t,:}\right\|_2 + \left\|\widehat{\boldsymbol{Y}}_{t_k^\star,:} - \boldsymbol{W}_{t_k^\star,:}\right\|_2 + \cancelto{0}{\left\|\boldsymbol{W}_{t,:} - \boldsymbol{W}_{t_k^\star,:}\right\|_2} &\geq \left\|\widehat{\boldsymbol{Y}}_{t,:} - \widehat{\boldsymbol{Y}}_{t_k^\star,:}\right\|_2 \\
&> \left\|\widehat{\boldsymbol{Y}}_{t,:} - \widehat{\boldsymbol{Y}}_{t_{k'}^\star,:}\right\|_2 \\
&> \left\|\boldsymbol{W}_{t,:} - \boldsymbol{W}_{t_{k'}^\star,:}\right\|_2 - \left\|\widehat{\boldsymbol{Y}}_{t,:} - \boldsymbol{W}_{t,:}\right\|_2 - \left\|\widehat{\boldsymbol{Y}}_{t_{k'}^\star,:} - \boldsymbol{W}_{t_{k'}^\star,:}\right\|_2,
\end{aligned}$$

i.e.,

$$\begin{aligned}\left\|\widehat{\boldsymbol{Y}}_{t,:}-\boldsymbol{W}_t\right\|_2 &> \frac{1}{2}\left(\left\|\boldsymbol{W}_{t,:}-\boldsymbol{W}_{t^\star_{k'},:}\right\|_2-\left\|\widehat{\boldsymbol{Y}}_{t^\star_{k'},:}-\boldsymbol{W}_{t^\star_k,:}\right\|_2-\left\|\widehat{\boldsymbol{Y}}_{t^\star_k,:}-\boldsymbol{W}_{t^\star_{k'},:}\right\|_2\right)\\ &\geq \frac{1}{2}\left(\Delta_{\boldsymbol{W}}-\frac{1}{2}\Delta_{\boldsymbol{W}}\right)=\frac{1}{4}\Delta_{\boldsymbol{W}}. \qquad (t^\star_k, t^\star_{k'}\in\textstyle\bigcup_{k=1}^K\mathcal{I}_k\text{, see the proof of property (iv)})\end{aligned}$$

□

B.3 Proof of Lemma B.3

As the exact constants for the ℓ_2-distance concentration have been left to exercise in Canonne (2020), we provide the complete proof here.

First, we have that

$$\begin{aligned}\mathbb{E}\left[\|\hat{p}_N-p\|_2^2\right] &= \sum_{s\in\mathcal{S}}\mathbb{E}\left[(\hat{p}_N(s)-p(s))^2\right]\\ &= \frac{1}{N^2}\sum_{s\in\mathcal{S}}\mathrm{Var}[\mathrm{Bin}(N,p(s))] \qquad (N\hat{p}_N(s)\sim\mathrm{Bin}(N,p(s)))\\ &= \frac{1}{N}\sum_{s\in\mathcal{S}}p(s)(1-p(s))\\ &= \frac{1}{N}\left(1-\sum_{s\in\mathcal{S}}p(s)^2\right)\leq\frac{1}{N}\left(1-\frac{1}{S}\right), \qquad \text{(Cauchy-Schwartz inequality)}\end{aligned}$$

i.e., with $N\geq\frac{4}{\varepsilon^2}\left(1-\frac{1}{S}\right)$,

$$\mathbb{E}\left[\|\hat{p}_N-p\|_2\right]\overset{(*)}{\leq}\sqrt{\mathbb{E}\left[\|\hat{p}_N-p\|_2^2\right]}\leq\frac{\varepsilon}{2},$$

where $(*)$ follows from Jensen's inequality.

We now utilize the McDiarmid's inequality (McDiarmid, 1989) to turn this into a *tight* high-probability guarantee. For $\boldsymbol{s}_N=(s_1,\cdots,s_N)$, let us define

$$f(\boldsymbol{s}_N):=\|\hat{p}_N-p\|_2=\sqrt{\sum_{i=1}^N(\hat{p}_N(s_i)-p(s_i))^2}.$$

Now, let $\boldsymbol{s}^{\neg i}=(s_1,\ldots,s_{i-1},s_i',s_{i+1},\ldots,s_N)$ for $s_i'\in\{0,1\}$ and $\hat{p}_N^{\neg i}$ be the empirical distribution using $\boldsymbol{x}^{\neg i}$. Observe that

$$\begin{aligned}|f(\boldsymbol{s})-f(\boldsymbol{s}^{\neg i})| &= \left|\|\hat{p}_N-p\|_2-\|\hat{p}_N^{\neg i}-p\|_2\right|\\ &\leq\|\hat{p}_N-\hat{p}_N^{\neg i}\|_2 \qquad \text{(Triangle inequality)}\\ &= \frac{1}{N}|s_i-s_i'|\leq\frac{1}{N}.\end{aligned}$$

Thus, f satisfies the bounded difference property with $c_i=1/N$.

Combining everything, we have that

$$\begin{aligned}\mathbb{P}\left(\|\hat{p}_N-p\|_2\geq\varepsilon\right) &\leq\mathbb{P}\left(\|\hat{p}_N-p\|_2-\mathbb{E}[\|\hat{p}_N-p\|_2]\geq\frac{\varepsilon}{2}\right) \qquad \left(\text{with } N\geq\tfrac{4}{\varepsilon^2}\left(1-\tfrac{1}{S}\right)\right)\\ &\leq\exp\left(-\frac{2\left(\frac{\varepsilon}{2}\right)^2}{N\left(\frac{1}{N}\right)^2}\right)\\ &=\exp\left(-\frac{\varepsilon^2N}{2}\right).\end{aligned}$$

□

C PROOF OF THEOREM 4.2 – LIKELIHOOD IMPROVEMENT

The proof largely follows the recipe of likelihood improvement for SBM clustering (Yun and Proutière, 2014b,a), which dates back to the seminal works of Abbe (2018); Amini et al. (2013) and recently extended to other variants such as labeled SBM (Yun and Proutière, 2016) and block Markov chains (Jedra et al., 2023; Sanders et al., 2020).

Let $\mathcal{H} \subseteq [T]$ be the maximum subset satisfying the following: whenever $t \in f^{-1}(k) \cap \mathcal{H}$,

$$\sum_{s,s' \in \mathcal{S}} \widehat{N}_t(s,s') \log \frac{p^{(k)}(s'|s)}{p^{(k')}(s'|s)} \geq C(H-1)\mathcal{D}_\pi, \quad \forall k' \neq k. \tag{27}$$

The first proposition bounds $|\mathcal{H}^{\complement}|$ in-expectation, which then leads to a *w.h.p.* guarantee via Markov's inequality:

Proposition C.1. *Suppose that* $H \gtrsim \frac{\eta_p^2}{\gamma_{\mathrm{ps}}\mathcal{D}_\pi} \log \frac{1}{\pi_{\min}}$. *With* $C = \frac{1}{2}$, *We have that*

$$\mathbb{E}[|\mathcal{H}^{\complement}|] \leq T(K-1)\exp\left(-\frac{\gamma_{\mathrm{ps}}(H-1)\mathcal{D}_\pi}{128(4\eta_p^2 + 5\log\eta_p)}\right). \tag{28}$$

Thus, by Markov's inequality, we have that

$$\mathbb{P}\left(|\mathcal{H}^{\complement}| \geq T\exp\left(-\frac{\gamma_{\mathrm{ps}}(H-1)\mathcal{D}_\pi}{256(4\eta_p^2 + 5\log\eta_p)}\right)\right) \leq K\exp\left(-\frac{\gamma_{\mathrm{ps}}(H-1)\mathcal{D}_\pi}{256(4\eta_p^2 + 5\log\eta_p)}\right). \tag{29}$$

We next show that all trajectories in $\mathcal{H}$ are correctly classified with high probability:

Proposition C.2. *Suppose that the following requirements on* H *and* T *hold:*

$$H \gtrsim \frac{1}{\gamma_{\mathrm{ps}}}\left(S^2\log T \vee \frac{\log T}{\pi_{\min}} \vee \frac{\eta_p\eta_\pi RS\log(TH)}{\alpha_{\min}H\Delta_{\boldsymbol{W}}^2\mathcal{D}_\pi}\right), \; TH \gtrsim \frac{(S\log T)^2}{\gamma_{\mathrm{ps}}\alpha_{\min}\mathcal{D}_\pi^2}. \tag{30}$$

Then, all trajectories in $\mathcal{H}$ *are correctly classified with probability at least* $1 - \frac{1}{\sqrt{\pi_{\min}}T} - \frac{1}{TH}$.

Note that the probability in Eqn. (29) goes to 0 as $T \to \infty$. from our requirement that $H \gtrsim \log T$ (ignoring other factors). Finally, the proof concludes by taking the worst-case, namely, that all the trajectories in $\mathcal{H}^{\complement}$ are misclassified, and applying a union bound.

□

C.1 Proof of Proposition C.1

First note that

$$\begin{aligned}
\mathbb{E}[|\mathcal{H}^{\complement}|] &= \sum_{t\in[T]} \mathbb{P}\left(\exists k' \neq f(t) \in [K] \text{ s.t. } \sum_{s,s'\in\mathcal{S}} \widehat{N}_t(s,s') \log\frac{p^{(f(t))}(s'|s)}{p^{(k')}(s'|s)} < C(H-1)\mathcal{D}_\pi\right) \\
&\leq (K-1)\sum_{t\in[T]} \max_{k'\neq f(t)\in[K]} \mathbb{P}\left(\sum_{s,s'\in\mathcal{S}} \widehat{N}_t(s,s') \log\frac{p^{(f(t))}(s'|s)}{p^{(k')}(s'|s)} < C(H-1)\mathcal{D}_\pi\right) \\
&\leq (K-1)\sum_{t\in[T]} \max_{k'\neq f(t)\in[K]} \mathbb{P}\left(\sum_{s,s'\in\mathcal{S}} \widehat{N}_t(s,s') \log\frac{p^{(f(t))}(s'|s)}{p^{(k')}(s'|s)} < C(H-1)\mathcal{D}_\pi(t;k')\right),
\end{aligned}$$

where we denote $\mathcal{D}_\pi(t;k') := \mathcal{D}_\pi^{(f(t),k')} = \sum_{s\in\mathcal{S}} \pi^{(f(t))}(s)\,\mathsf{KL}(p^{(f(t))}(\cdot|s), p^{(k')}(\cdot|s))$.

As done in Sanders et al. (2020); Jedra et al. (2023), define an "augmented" Markov chain, $X_{t,h} := (s_{t,h}, s_{t,h+1})$. For notational simplicity, let us denote $p := p^{(f(t))}$ and $\pi = \pi^{(f(t))}$.

This is a Markov chain on $\mathcal{S} \times \mathcal{S}$, with transition probability kernel $\tilde{p}(s_{t,h+1}, s_{t,h+2}|s_{t,h}, s_{t,h+1}) = p(s_{t,h+2}|s_{t,h+1})$ and stationary distribution $\tilde{\pi}(s, s') = \pi(s)p(s'|s)$. We have the following property that we provide the proof at the end:

Lemma C.1. *If the pseudo-spectral gap of $(s_{t,h})_h$ is γ, then so is that of $(X_{t,h})_h$.*

Now define a function $\phi_{k'} : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ as

$$\phi_{k'}(X_{t,h}) := \sum_{s,s' \in \mathcal{S}} \mathbb{1}[s_{t,h} = s, s_{t,h+1} = s'] \log \frac{p(s'|s)}{p^{(k')}(s'|s)} = \log \frac{p(s_{t,h+1}|s_{t,h})}{p^{(k')}(s_{t,h+1}|s_{t,h})}. \tag{31}$$

To apply the above concentration, we compute the necessary quantities. First, by Assumption 2, we have that for any $(s, s') \in \mathcal{S} \times \mathcal{S}$,

$$|\phi(s, s') - \mathbb{E}_\pi[\phi]| = \left| \log \frac{p(s'|s)}{p^{(k')}(s'|s)} - \sum_{s,s' \in \mathcal{S}} \pi(s)p(s'|s) \log \frac{p(s'|s)}{p^{(k')}(s'|s)} \right| \leq 2 \log \eta_p.$$

Also, note that

$$\begin{aligned}
\mathrm{Var}_\pi[\phi] &\leq \mathbb{E}_\pi[\phi^2] \\
&= \sum_{s,s'} \pi(s)p(s'|s) \left(\log \frac{p(s'|s)}{p^{(k')}(s'|s)} \right)^2 \\
&\leq 2 \left(\max_{s' \in \mathcal{S}} \frac{p(s'|s) \vee p^{(k')}(s'|s)}{p(s'|s) \wedge p^{(k')}(s'|s)} \right)^2 \sum_{s \in \mathcal{S}} \pi(s) \, \mathsf{KL} \left(p(\cdot|s), p^{(k')}(\cdot|s) \right) && \text{(Lemma 19 of SM6.3 of Sanders et al. (2020))} \\
&\leq 2\eta_p^2 \mathcal{D}_\pi(t; k'). && \text{(Assumption 2)}
\end{aligned}$$

With this, we have that

$$\begin{aligned}
&\mathbb{P}_\pi \left(\sum_{s,s' \in \mathcal{S}} \widehat{N}_t(s, s') \log \frac{p^{(k)}(s'|s)}{p^{(k')}(s'|s)} < C(H-1)\mathcal{D}_\pi(t; k') \right) \\
&= \mathbb{P}_\pi \left(\sum_{h=1}^{H-1} (\phi_{k'}(X_{t,h}) - \mathbb{E}[\phi_{k'}(X_{t,h})]) < -(1-C)(H-1)\mathcal{D}_\pi(t; k') \right) \\
&\leq \exp \left(-\frac{(1-C)^2 (H-1)^2 (\mathcal{D}_\pi(t; k'))^2 \gamma_{\mathrm{ps}}}{32\eta_p^2 (H-1)\mathcal{D}_\pi(t; k') + 40(\log \eta_p)(H-1)\mathcal{D}_\pi(t; k')} \right) && \text{(Lemma B.4, Lemma C.1, and } H > \tfrac{1}{\gamma_{\mathrm{ps}}}\text{)} \\
&\leq \exp \left(-\frac{(1-C)^2}{8(4\eta_p^2 + 5\log \eta_p)} \gamma_{\mathrm{ps}} (H-1)\mathcal{D}_\pi(t; k') \right) \\
&\leq \exp \left(-\frac{(1-C)^2}{8(4\eta_p^2 + 5\log \eta_p)} \gamma_{\mathrm{ps}} (H-1)\mathcal{D}_\pi \right).
\end{aligned}$$

We choose $C = \frac{1}{2}$ and invoke Lemma B.5, which yields

$$\begin{aligned}
&\mathbb{P} \left(\sum_{s,s' \in \mathcal{S}} \widehat{N}_t(s, s') \log \frac{p^{(k)}(s'|s)}{p^{(k')}(s'|s)} < \frac{(H-1)\mathcal{D}_\pi(t; k')}{2} \right) \\
&\leq \sqrt{\frac{1}{\pi_{\min}} \exp \left(-\frac{1}{32(4\eta_p^2 + 5\log \eta_p)} \gamma_{\mathrm{ps}} (H-1)\mathcal{D}_\pi \right)} \\
&= \sqrt{\exp \left(\log \frac{1}{\pi_{\min}} - \frac{1}{32(4\eta_p^2 + 5\log \eta_p)} \gamma_{\mathrm{ps}} (H-1)\mathcal{D}_\pi \right)}
\end{aligned}$$

$$\leq \exp\left(-\frac{\gamma_{\mathrm{ps}}(H-1)\mathcal{D}_\pi}{128(4\eta_p^2+5\log\eta_p)}\right),$$

where the last inequality is true when $H-1 > \frac{64(4\eta_p^2+5\log\eta_p)}{\gamma_{\mathrm{ps}}\mathcal{D}_\pi}\log\frac{1}{\pi_{\min}}$. □

Proof of Lemma C.1. For notational simplicity, we omit the dependency on t. Let $\boldsymbol{P}$ and $\widetilde{\boldsymbol{P}}$ be the (bounded) linear operators corresponding to the original Markov chain s_h and the augmented Markov chain $X_h := (s_h, s_{h+1})$, respectively. Then, it is clear that $\widetilde{\boldsymbol{P}} = \bigoplus_{s\in\mathcal{S}} \boldsymbol{P}$, where $\bigoplus$ denotes the (matrix) direct sum.

We now denote $\boldsymbol{P}^*$ and $\widetilde{\boldsymbol{P}}^*$ be their adjoints on $L^2(\pi)$ and $L^2(\tilde{\pi})$, respectively, i.e., their time reversals (Lezaud, 1998). By a well-known property of direct sum (Horn and Johnson, 2012), we have that $\widetilde{\boldsymbol{P}}^* = \bigoplus_{s\in\mathcal{S}} \boldsymbol{P}^*$. This can also be seen via explicit computation:

$$\tilde{p}^*(y,y'|x,x') := \frac{\tilde{\pi}(y,y')\tilde{p}(x,x'|y,y')}{\tilde{\pi}(x,x')} = \frac{\pi(y)p(y'|y)p(x'|x)\mathbb{1}[y'=x]}{\pi(x)\tilde{p}(x'|x)} = \underbrace{\frac{\pi(y)p(x|y)}{\pi(x)}}_{=p^*(y|x)}\mathbb{1}[y'=x].$$

The off-diagonal zero blocks correspond to where $\mathbb{1}[y'=x]=0$.

Thus, we have that for every $k\in\mathbb{N}$, $(\widetilde{\boldsymbol{P}}^*)^k\widetilde{\boldsymbol{P}}^k = \bigoplus_{s\in\mathcal{S}}(\boldsymbol{P}^*)^k\boldsymbol{P}^k$, i.e., the spectrum of $(\widetilde{\boldsymbol{P}}^*)^k\widetilde{\boldsymbol{P}}^k$ is precisely S copies of that of $(\boldsymbol{P}^*)^k\boldsymbol{P}^k$. This immediately implies that $\gamma_{\mathrm{ps}}(\widetilde{\boldsymbol{P}}) = \gamma_{\mathrm{ps}}(\boldsymbol{P})$. □

C.2 Proof of Proposition C.2

Recall that

$$\mathcal{L}(k;t) := \sum_{h=1}^{H-1}\log\widehat{p}_0^{(k)}(s_{t,h+1}|s_{t,h}) = \sum_{s,s'\in\mathcal{S}}\widehat{N}_t(s,s')\log\widehat{p}_0^{(k)}(s'|s) \tag{32}$$

where

$$\widehat{p}_0^{(k)}(s'|s) := \frac{\sum_{t\in(\widehat{f}_0)^{-1}(k)}\sum_{h\in[H-1]}\mathbb{1}[s_{t,h}=s, s_{t,h+1}=s']}{\sum_{t\in(\widehat{f}_0)^{-1}(k)}\sum_{h\in[H-1]}\mathbb{1}[s_{t,h}=s]}. \tag{33}$$

Then, Algorithm 2 ensures that t is misclassified if and only if

$$E_t \triangleq \mathcal{L}(\widehat{f}(t);t) - \mathcal{L}(f(t);t) > 0. \tag{34}$$

We decompose E_t as

$$E_t = \underbrace{\sum_{s,s'\in\mathcal{S}}\widehat{N}_t(s,s')\log\frac{p^{(\widehat{f}_0(t))}(s'|s)}{p^{(f(t))}(s'|s)}}_{\triangleq E_{1,t}} + \underbrace{\sum_{s,s'\in\mathcal{S}}\widehat{N}_t(s,s')\log\frac{\widehat{p}_0^{(\widehat{f}_0(t))}(s'|s)}{p^{(\widehat{f}_0(t))}(s'|s)}}_{\triangleq E_{2,t}} + \underbrace{\sum_{s,s'\in\mathcal{S}}\widehat{N}_t(s,s')\log\frac{p^{(f(t))}(s'|s)}{\widehat{p}_0^{(f(t))}(s'|s)}}_{\triangleq E_{3,t}}$$

We now upper bound $E_{2,t}$ and $E_{3,t}$.

To do so, we start with the following concentration for $\widehat{N}_t(s,s')$, which can be derived via Lemma B.4 (similar to Eqn. (23)): for each $k\in[K]$ and $t\in f^{-1}(k)$

$$\mathbb{P}\left(\left|\widehat{N}_t(s,s') - (H-1)\pi^{(k)}(s)p^{(k)}(s'|s)\right| \gtrsim \sqrt{\frac{H\pi^{(k)}(s)p^{(k)}(s'|s)}{\gamma_{\mathrm{ps}}}\log\frac{1}{\pi_{\min}\delta^2}}\right) \leq \delta, \tag{35}$$

i.e., $\widehat{N}_t(s,s') \lesssim H\pi^{(k)}(s)p^{(k)}(s'|s) + \sqrt{\frac{H\pi^{(k)}(s)p^{(k)}(s'|s)}{\gamma_{\rm ps}}\log T}$ with probability at least $1 - \frac{1}{\sqrt{\pi_{\min}}T^2}$

We then bound the intermediate estimation error of the transition probabilities:

Lemma C.2. *Let $\delta \in (0,1)$, and suppose that $H \gtrsim \frac{1}{\gamma_{\rm ps}\pi_{\min}}\log\frac{1}{\delta}$. Then, for each $k \in [K]$ and $s, s' \in \mathcal{S}$, the following holds with probability at least $1-\delta$:*

$$\left|\log\frac{\widehat{p}_0^{(k)}(s'|s)}{p^{(k)}(s'|s)}\right| \lesssim \sqrt{\frac{1}{\gamma_{\rm ps}\alpha_k T H\pi^{(k)}(s)p^{(k)}(s'|s)}\log\frac{\alpha_k T}{\sqrt{\pi_{\min}}\delta}\log\frac{1}{\sqrt{\pi_{\min}}\delta}} + \frac{\eta_p\eta_\pi e^{(0)}}{\alpha_k T}. \tag{36}$$

Then, for each $t \in f^{-1}(k)$, we have that with probability at least $1 - \frac{1}{\sqrt{\pi_{\min}}T^2}$,

$$\begin{aligned}
E_{2,t} &\lesssim \sum_{s,s'\in\mathcal{S}} H\pi^{(k)}(s)p^{(k)}(s'|s)\left(\sqrt{\frac{1}{\gamma_{\rm ps}\alpha_k T H\pi^{(k)}(s)p^{(k)}(s'|s)}}\log T + \frac{\eta_p\eta_\pi e^{(0)}}{\alpha_k T}\right) \\
&\quad + \sum_{s,s'\in\mathcal{S}}\sqrt{\frac{H\pi^{(k)}(s)p^{(k)}(s'|s)}{\gamma_{\rm ps}}\log T}\left(\sqrt{\frac{1}{\gamma_{\rm ps}\alpha_k T H\pi^{(k)}(s)p^{(k)}(s'|s)}}\log T + \frac{\eta_p\eta_\pi e^{(0)}}{\alpha_k T}\right) \\
&= \sum_{s,s'\in\mathcal{S}}\left(\sqrt{\frac{H\pi^{(k)}(s)p^{(k)}(s'|s)}{\gamma_{\rm ps}\alpha_k T}}\log T\right) + \frac{H\eta_p\eta_\pi e^{(0)}}{\alpha_k T} \\
&\quad + \sum_{s,s'\in\mathcal{S}}\left(\sqrt{\frac{1}{\alpha_k T}}\frac{(\log T)^{3/2}}{\gamma_{\rm ps}} + \sqrt{\frac{H\pi^{(k)}(s)p^{(k)}(s'|s)}{\gamma_{\rm ps}}\log T}\frac{\eta_p\eta_\pi e^{(0)}}{\alpha_k T}\right) \\
&\leq S\sqrt{\frac{H}{\gamma_{\rm ps}\alpha_k T}}\log T + \frac{H\eta_p\eta_\pi e^{(0)}}{\alpha_k T} + \frac{S^2(\log T)^{3/2}}{\gamma_{\rm ps}\sqrt{\alpha_k T}} + S\sqrt{\frac{H}{\gamma_{\rm ps}}\log T}\frac{\eta_p\eta_\pi e^{(0)}}{\alpha_k T} \\
&\qquad\qquad \left(\textstyle\sum_{s,s'\in\mathcal{S}}\sqrt{\pi^{(k)}(s)p^{(k)}(s'|s)} \leq S\right) \\
&\lesssim S\sqrt{\frac{H}{\gamma_{\rm ps}\alpha_k T}}\log T + \frac{H\eta_p\eta_\pi e^{(0)}}{\alpha_k T} \qquad \left(\text{when } H \gtrsim \frac{S^2\log T}{\gamma_{\rm ps}}\right)
\end{aligned}$$

and the same for $E_{3,t}$.

Now invoking Theorem 4.1 and the union bound, we have that with probability at least $1 - \frac{1}{\sqrt{\pi_{\min}}T} - \frac{1}{TH}$,

$$E_{2,t}, E_{3,t} \lesssim S\sqrt{\frac{H}{\gamma_{\rm ps}\alpha_{\min}T}}\log T + \frac{H\eta_p\eta_\pi}{\alpha_{\min}T}\frac{TRS}{\gamma_{\rm ps}H\Delta_{\boldsymbol{W}}^2}\log(TH). \tag{37}$$

With this, we deduce that $t \in \mathcal{H}$ is misclassified *only if*

$$H\mathcal{D}_\pi \lesssim -E_{1,t} < E_{2,t} + E_{3,t} \lesssim S\sqrt{\frac{H}{\gamma_{\rm ps}\alpha_{\min}T}}\log T + \frac{\eta_p\eta_\pi}{\alpha_{\min}}\frac{RS}{\gamma_{\rm ps}\Delta_{\boldsymbol{W}}^2}\log(TH),$$

where the first inequality follows from the definition of $\mathcal{H}$. By taking the contrapositive, we can deduce that if

$$H\mathcal{D}_\pi \gtrsim S\sqrt{\frac{H}{\gamma_{\rm ps}\alpha_{\min}T}}\log T \vee \frac{\eta_p\eta_\pi}{\alpha_{\min}}\frac{RS}{\gamma_{\rm ps}\Delta_{\boldsymbol{W}}^2}\log(TH)$$

hold, then $t \in \mathcal{H}$ must be classified correctly after one-shot likelihood improvement.

We now collect all the requirements in blue throughout the proof, which gives:

$$H \gtrsim \frac{S^2\log T}{\gamma_{\rm ps}} \vee \frac{\log T}{\gamma_{\rm ps}\pi_{\min}} \vee \frac{\eta_p\eta_\pi RS\log(TH)}{\mathcal{D}_\pi\alpha_{\min}\gamma_{\rm ps}\Delta_{\boldsymbol{W}}^2}, \quad TH \gtrsim \frac{(S\log T)^2}{\mathcal{D}_\pi^2\gamma_{\rm ps}\alpha_{\min}}. \tag{38}$$

The proof then concludes with union bound over $t \in \mathcal{H}$. ∎

C.3 Proof of Lemma C.2

We denote $N_t(s) := (H-1)\pi^{(f(t))}(s)$ and $N_t(s,s') := (H-1)\pi^{(f(t))}(s)p^{(f(t))}(s'|s)$ as the expected number of visitations/transitions under the respective chain's stationary distribution.

As $\frac{x}{1+x} \le \log(1+x) \le x$ for $x > -1$, we have that

$$\begin{aligned}
\left|\log \frac{\widehat{p}_0^{(k)}(s'|s)}{p^{(k)}(s'|s)}\right| &\le \left|\frac{\widehat{p}_0^{(k)}(s'|s)}{p^{(k)}(s'|s)} - 1\right| \\
&= \left|\frac{\sum_{t\in(\widehat{f}_0)^{-1}(k)} \widehat{N}_t(s,s')}{\sum_{t\in(\widehat{f}_0)^{-1}(k)} \widehat{N}_t(s)} \frac{1}{p^{(k)}(s'|s)} - 1\right| \\
&= \left|\underbrace{\frac{\sum_{t\in(\widehat{f}_0)^{-1}(k)} \widehat{N}_t(s,s')}{\sum_{t\in(\widehat{f}_0)^{-1}(k)} N_t(s,s')} \frac{\sum_{t\in(\widehat{f}_0)^{-1}(k)} N_t(s)}{\sum_{t\in(\widehat{f}_0)^{-1}(k)} \widehat{N}_t(s)}}_{\triangleq E_{est}} \underbrace{\frac{\sum_{t\in(\widehat{f}_0)^{-1}(k)} N_t(s,s')}{\sum_{t\in(\widehat{f}_0)^{-1}(k)} N_t(s)} \frac{1}{p^{(k)}(s'|s)}}_{\triangleq E_{cluster}} - 1\right|.
\end{aligned}$$

Throughout the proof, let $\mathcal{E}_0^{(k)} := \widehat{f}_0^{-1}(k) \setminus f^{-1}(k)$ and $\mathcal{C}_0^{(k)} := f^{-1}(k) \cap \widehat{f}_0^{-1}(k)$, which satisfies $\widehat{f}_0^{-1}(k) = \mathcal{E}_0^{(k)} \dot\cup \mathcal{C}_0^{(k)}$. We first bound E_{est}:

$$\begin{aligned}
\left|\frac{\sum_{t\in\widehat{f}_0^{-1}(k)} \widehat{N}_t(s,s')}{\sum_{t\in\widehat{f}_0^{-1}(k)} N_t(s,s')} - 1\right| &= \frac{\left|\sum_{t\in\widehat{f}_0^{-1}(k)} \left(\widehat{N}_t(s,s') - N_t(s,s')\right)\right|}{\sum_{t\in\mathcal{E}_0^{(k)}} N_t(s,s') + \sum_{t\in\mathcal{C}_0^{(k)}} N_t(s,s')} \\
&\le \frac{\left|\sum_{t\in\widehat{f}_0^{-1}(k)} \left(\widehat{N}_t(s,s') - N_t(s,s')\right)\right|}{\sum_{t\in\mathcal{C}_0^{(k)}} N_t(s,s')} \\
&\le \frac{\left|\sum_{t\in\widehat{f}_0^{-1}(k)} \left(\widehat{N}_t(s,s') - N_t(s,s')\right)\right|}{(\alpha_k T - |\mathcal{E}_0^{(k)}|)(H-1)\pi^{(k)}(s)p^{(k)}(s'|s)} && (|\mathcal{C}_0^{(k)}| \ge \alpha_k T - |\mathcal{E}_0^{(k)}|) \\
&\le \frac{2\left|\sum_{t\in\widehat{f}_0^{-1}(k)} \left(\widehat{N}_t(s,s') - N_t(s,s')\right)\right|}{\alpha_k T(H-1)\pi^{(k)}(s)p^{(k)}(s'|s)}. && (|\mathcal{E}_0^{(k)}| \le e^{(0)} \le \tfrac{\alpha_{\min}T}{2} \le \tfrac{\alpha_k T}{2})
\end{aligned}$$

Defining $X_t := \widehat{N}_t(s,s') - N_t(s,s')$, note that X_t's are independent across t and satisfy $\mathbb{E}_\pi[X_t] = 0$. Let us define an event $\mathcal{M} := \left\{|X_t| \le M,\ \forall t \in \widehat{f}_0^{-1}(k)\right\}$, where $M > 0$ is chosen later. By Hoeffding's inequality (Hoeffding, 1963) conditioned on $\mathcal{M}$, we have that for any $z \ge 0$,

$$\mathbb{P}_\pi\left(\left|\sum_{t\in\widehat{f}_0^{-1}(k)} \left(\widehat{N}_t(s,s') - N_t(s,s')\right)\right| \ge z \,\middle|\, \mathcal{M}\right) \le \exp\left(-\frac{z^2}{2M^2|\widehat{f}_0^{-1}(k)|}\right).$$

Using Lemma B.5 to change $\mathbb{P}_\pi$ to $\mathbb{P}$ and the fact that $|\widehat{f}_0^{-1}(k)| \le \frac{3\alpha_k T}{2}$, we have

$$\mathbb{P}\left(\left|\sum_{t\in\widehat{f}_0^{-1}(k)} \left(\widehat{N}_t(s,s') - N_t(s,s')\right)\right| \ge z \,\middle|\, \mathcal{M}\right) \le \frac{1}{\sqrt{\pi_{\min}}} \exp\left(-\frac{z^2}{6M^2\alpha_k T}\right).$$

We choose $M \asymp \sqrt{\frac{H\pi(s)p(s'|s)}{\gamma_{\text{ps}}} \log \frac{\alpha_k T}{\sqrt{\pi_{\min}}\delta}}$. Then, by Eqn. (35), we have that $\mathbb{P}(\mathcal{M}^\complement) \le \frac{\delta}{2}$. By reparametrizing

$z \asymp M\sqrt{\alpha_k T \log \frac{1}{\sqrt{\pi_{\min}}\delta}}$ and union bound, we have that

$$\mathbb{P}\left(\left|\sum_{t\in \widehat{f}_0^{-1}(k)} \left(\widehat{N}_t(s,s') - N_t(s,s')\right)\right| \geq \sqrt{\frac{\alpha_k T H \pi(s) p(s'|s)}{\gamma_{\mathrm{ps}}} \log \frac{\alpha_k T}{\sqrt{\pi_{\min}}\delta} \log \frac{1}{\sqrt{\pi_{\min}}\delta}}\right) \leq \delta.$$

Then, combining everything, the following holds with probability at least $1-\delta$:

$$\left|\frac{\sum_{t\in \widehat{f}_0^{-1}(k)} \widehat{N}_t(s,s')}{\sum_{t\in \widehat{f}_0^{-1}(k)} N_t(s,s')} - 1\right| \lesssim \sqrt{\frac{1}{\gamma_{\mathrm{ps}}\alpha_k T H \pi(s) p(s'|s)} \log \frac{\alpha_k T}{\sqrt{\pi_{\min}}\delta} \log \frac{1}{\sqrt{\pi_{\min}}\delta}}. \tag{39}$$

For the other term, we have that

$$\begin{aligned}
\left|\frac{\sum_{t\in \widehat{f}_0^{-1}(k)} N_t(s)}{\sum_{t\in \widehat{f}_0^{-1}(k)} \widehat{N}_t(s)} - 1\right| &\leq \frac{1}{\sum_{t\in \widehat{f}_0^{-1}(k)} \widehat{N}_t(s)} \left|\sum_{t\in \widehat{f}_0^{-1}(k)} \left(\widehat{N}_t(s) - N_t(s)\right)\right| \\
&\lesssim \frac{1}{\sum_{t\in \widehat{f}_0^{-1}(k)} \widehat{N}_t(s)} \sqrt{\frac{\alpha_k T H \pi(s)}{\gamma_{\mathrm{ps}}} \log \frac{\alpha_k T}{\sqrt{\pi_{\min}}\delta} \log \frac{1}{\sqrt{\pi_{\min}}\delta}} \\
&\qquad\qquad \text{(with probability at least } 1-\delta/2) \\
&\overset{(*)}{\lesssim} \frac{1}{\sum_{t\in \widehat{f}_0^{-1}(k)} N_t(s)} \sqrt{\frac{\alpha_k T H \pi(s)}{\gamma_{\mathrm{ps}}} \log \frac{\alpha_k T}{\sqrt{\pi_{\min}}\delta} \log \frac{1}{\sqrt{\pi_{\min}}\delta}} \\
&\qquad\qquad \text{(with probability at least } 1-\delta/2\text{, assuming that } H \gtrsim \tfrac{1}{\gamma_{\mathrm{ps}}\pi_{\min}} \log \tfrac{1}{\delta}) \\
&\lesssim \sqrt{\frac{1}{\gamma_{\mathrm{ps}}\alpha_k T H \pi(s)} \log \frac{\alpha_k T}{\sqrt{\pi_{\min}}\delta} \log \frac{1}{\sqrt{\pi_{\min}}\delta}},
\end{aligned}$$

where $(*)$ follows from Lemma B.4.

We now bound $E_{cluster}$:

$$\begin{aligned}
\frac{\sum_{t\in (\widehat{f}_0)^{-1}(k)} N_t(s,s')}{\sum_{t\in (\widehat{f}_0)^{-1}(k)} N_t(s)} \frac{1}{p^{(k)}(s'|s)} &= \frac{\sum_{t\in \mathcal{C}_0^{(k)}} p^{(k)}(s'|s) N_t(s) + \sum_{t\in \mathcal{E}_0^{(k)}} p^{(\widehat{f}_0(t))}(s'|s) N_t(s)}{\left(\sum_{t\in \mathcal{C}_0^{(k)}} N_t(s) + \sum_{t\in \mathcal{E}_0^{(k)}} N_t(s)\right) p^{(k)}(s'|s)} \\
&\leq \frac{\sum_{t\in \mathcal{C}_0^{(k)}} N_t(s) + \eta_p \sum_{t\in \mathcal{E}_0^{(k)}} N_t(s)}{\sum_{t\in \mathcal{C}_0^{(k)}} N_t(s) + \sum_{t\in \mathcal{E}_0^{(k)}} N_t(s)} \qquad \text{(Assumption 2)} \\
&\leq \frac{\sum_{t\in \mathcal{C}_0^{(k)}} N_t(s) + \eta_p \sum_{t\in \mathcal{E}_0^{(k)}} N_t(s)}{\sum_{t\in \mathcal{C}_0^{(k)}} N_t(s)} \\
&= 1 + \frac{\eta_p \sum_{t\in \mathcal{E}_0^{(k)}} N_t(s)}{\sum_{t\in \mathcal{C}_0^{(k)}} N_t(s)}.
\end{aligned}$$

Recalling that $\pi^{(k)}(s)$'s satisfy η-regularity (Assumption 2), we then have that

$$\begin{aligned}
\frac{\sum_{t\in (\widehat{f}_0)^{-1}(k)} N_t(s,s')}{\sum_{t\in (\widehat{f}_0)^{-1}(k)} N_t(s)} \frac{1}{p^{(k)}(s'|s)} - 1 &\leq \frac{\eta_p \sum_{t\in \mathcal{E}_0^{(k)}} (H-1)\pi^{(\widehat{f}_0(t))}(s)}{(\alpha_k T - |\mathcal{E}_0^{(k)}|)(H-1)\pi^{(k)}(s)} \qquad (|\mathcal{C}_0^{(k)}| \geq \alpha_k T - |\mathcal{E}_0^{(k)}|) \\
&\leq \frac{2\eta_p \eta_\pi |\mathcal{E}_0^{(k)}|}{\alpha_k T} \qquad \text{(Assumption 2, } |\mathcal{E}_0^{(k)}| \leq e^{(0)} \leq \tfrac{\alpha_{\min} T}{2} \leq \tfrac{\alpha_k T}{2})
\end{aligned}$$

We also have that

$$\frac{\sum_{t\in (\widehat{f}_0)^{-1}(k)} N_t(s,s')}{\sum_{t\in (\widehat{f}_0)^{-1}(k)} N_t(s)} \frac{1}{p^{(k)}(s'|s)} = \frac{\sum_{t\in \mathcal{C}_0^{(k)}} p^{(k)}(s'|s) N_t(s) + \sum_{t\in \mathcal{E}_0^{(k)}} p^{(\widehat{f}_0(t))}(s'|s) N_t(s)}{\left(\sum_{t\in \mathcal{C}_0^{(k)}} N_t(s) + \sum_{t\in \mathcal{E}_0^{(k)}} N_t(s)\right) p^{(k)}(s'|s)}$$

$$\geq \frac{\sum_{t\in\mathcal{C}_0^{(k)}} N_t(s) + \eta_p^{-1}\sum_{t\in\mathcal{E}_0^{(k)}} N_t(s)}{\sum_{t\in\mathcal{C}_0^{(k)}} N_t(s) + \sum_{t\in\mathcal{E}_0^{(k)}} N_t(s)} \qquad \text{(Assumption 2)}$$

$$= 1 + \frac{(\eta_p^{-1}-1)\sum_{t\in\mathcal{E}_0^{(k)}} N_t(s)}{\sum_{t\in\mathcal{C}_0^{(k)}} N_t(s) + \sum_{t\in\mathcal{E}_0^{(k)}} N_t(s)} > 0,$$

i.e., we can bound $E_{cluster}$ as

$$|E_{cluster} - 1| \leq \frac{2\eta_p\eta_\pi e^{(0)}}{\alpha_k T}. \tag{40}$$

Combining everything, we are done. □

D PROOF OF PROPOSITION 5.1 – RELATIONSHIP BETWEEN DIFFERENT GAPS

D.1 (a) Relationship between $\mathcal{D}_\pi$ and Kausik et al. (2023)'s α, Δ

We first recall the following lemma connecting the KL divergence and the L_2-distance between two probability measures on a finite state space $\mathcal{S}$:

Lemma D.1. *For two probability mass functions p, q over a finite state space $\mathcal{S}$, we have*

$$\frac{\log \frac{e}{2}}{q_{\max} \vee p_{\max}} L_2(p,q) \le \mathsf{KL}(p,q) \le \frac{1}{\min_{s\in\mathcal{S}} q(s)} L_2(p,q), \tag{41}$$

where $p_{\max} := \max_{s\in\mathcal{S}} p(s)$, $q_{\max} := \max_{s\in\mathcal{S}} q(s)$, and $L_2(p,q) := \sum_{s\in\mathcal{S}} (p(s)-q(s))^2$

Then, note that for each $k \neq k'$,

$$\mathcal{D}^{(k,k')} \ge \pi^{(k)}(s(k,k'))\, \mathsf{KL}\left(p^{(k)}(\cdot|s(k,k')), p^{(k')}(\cdot|s(k,k'))\right) \ge \left(\log \frac{e}{2}\right) \frac{\alpha\Delta^2}{p_{\max}},$$

where we recall that $p_{\max} := \max_{k\in[K]} \max_{s,s'\in\mathcal{S}} p^{(k)}(s'|s)$. □

Remark 4. *Note that the first inequality is tight if such $s(k,k')$ is unique because $s(k,k')$ is the only state from which k and k' can be meaningfully distinguished. Also, if there are $W(k,k')$ number of such $s(k,k')$'s, then the lower bound can be improved to $\left(\log \frac{e}{2}\right) W(k,k') \frac{\alpha\Delta^2}{p_{\max}}$.*

Proof of Lemma D.1. The second inequality is due to Csiszar and Talata (2006, Lemma 6.1). The first inequality is due to a Stackoverflow post and user Ze-Nan Li's answer, whose proof we reproduce here for completeness.

Let us define $\eta(s) := \frac{q(s)-p(s)}{p(s)}$ and partition $\mathcal{S}$ as follows:

$$\mathcal{S} = \underbrace{\left\{s \in \mathcal{S} : \eta(s) > 1\right\}}_{\triangleq \mathcal{S}_+} \dot\cup \underbrace{\left\{s \in \mathcal{S} : \eta(s) \le 1\right\}}_{\triangleq \mathcal{S}_-}. \tag{42}$$

Then it is easy to see that

(i) For $s \in \mathcal{S}_+$, $1+\eta(s) \le 2^{\eta(s)} = e^{\eta(s)\log 2}$, and

(ii) For $s \in \mathcal{S}_-$, $1+\eta(s) \le e^{\eta(s) - c\eta(s)^2}$, where $c := \log \frac{e}{2}$.

Then,

$$\begin{aligned}
\mathsf{KL}(p,q) &:= \sum_{s\in\mathcal{S}} p(s) \log \frac{p(s)}{q(s)} \\
&= -\sum_{s\in\mathcal{S}_+} p(s)\log(1+\eta(s)) - \sum_{s\in\mathcal{S}_-} p(s)\log(1+\eta(s)) \\
&\ge -\log 2 \sum_{s\in\mathcal{S}_+} p(s)\eta(s) - \sum_{s\in\mathcal{S}_-} p(s)(\eta(s) - c\eta(s)^2) \\
&= c\sum_{s\in\mathcal{S}_+} p(s)\eta(s) - \cancelto{0}{\sum_{s\in\mathcal{S}} p(s)\eta(s)} + c\sum_{s\in\mathcal{S}_-} p(s)\eta(s)^2 \\
&= c\sum_{s\in\mathcal{S}_+} (q(s)-p(s)) + c\sum_{s\in\mathcal{S}_-} \frac{(q(s)-p(s))^2}{p(s)} \\
&\ge c\sum_{s\in\mathcal{S}_+} q(s)\frac{q(s)-p(s)}{q(s)} + \frac{c}{p_{\max}} \sum_{s\in\mathcal{S}_-} (q(s)-p(s))^2.
\end{aligned}$$

Now, for $s \in \mathcal{S}_+$, we have that $0 < \frac{q(s)-p(s)}{q(s)} \leq 1$, and thus,

$$\begin{aligned}
\mathsf{KL}(p,q) &\geq c \sum_{s\in\mathcal{S}_+} q(s)\frac{q(s)-p(s)}{q(s)} + \frac{c}{p_{\max}} \sum_{s\in\mathcal{S}_-} (q(s)-p(s))^2 \\
&\geq c \sum_{s\in\mathcal{S}_+} q(s)\left(\frac{q(s)-p(s)}{q(s)}\right)^2 + \frac{c}{p_{\max}} \sum_{s\in\mathcal{S}_-} (q(s)-p(s))^2 \\
&\geq \frac{c}{q_{\max} \vee p_{\max}} \underbrace{\sum_{s\in\mathcal{S}} (q(s)-p(s))^2}_{=L_2(p,q)} .
\end{aligned}$$

□

D.2 (b) Relationship between $\mathcal{D}_\pi$ and $\Delta_{\mathbf{W}}$

We first provide the full statement of (b), then provide its proof:

Proposition D.1. *We have*

$$\Delta_{\boldsymbol{W}}^2 \leq \min_{k\neq k'} \left\{ \frac{2p_{\max}}{\log\frac{e}{2}} \mathcal{D}_\pi^{(k,k')} + 4H^2\left(\pi^{(k)}, \pi^{(k')}\right) \right\}, \tag{43}$$

where $H^2(\cdot,\cdot)$ is the Hellinger distance: $H^2(p,q) := \frac{1}{2}\left(\sum_{s\in\mathcal{S}}(\sqrt{p(s)}-\sqrt{q(s)})^2\right)^{\frac{1}{2}}$. Furthermore, we also have the following slightly more relaxed bound:

$$\Delta_{\boldsymbol{W}}^2 \leq 7\left(p_{\max}\mathcal{D}_\pi + \left(\lceil \log_\rho M^{-1}\rceil + \frac{1}{1-\rho}\right)\sqrt{\frac{1}{2\pi_{\min}}\mathcal{D}_\pi}\right), \tag{44}$$

where ρ and M are the contraction constants from the definition of uniformly ergodic Markov chains[7].

Proof. For each $s \in \mathcal{S}$,

$$\begin{aligned}
&\left\|\sqrt{\pi^{(k)}(s)}p^{(k)}(\cdot|s) - \sqrt{\pi^{(k')}(s)}p^{(k')}(\cdot|s)\right\|_2^2 \\
&= \left\|\sqrt{\pi^{(k)}(s)}\left(p^{(k)}(\cdot|s) - p^{(k')}(\cdot|s)\right) + \left(\sqrt{\pi^{(k)}(s)} - \sqrt{\pi^{(k')}(s)}\right)p^{(k')}(\cdot|s)\right\|_2^2 \\
&\leq 2\pi^{(k)}(s)\left\|p^{(k)}(\cdot|s) - p^{(k')}(\cdot|s)\right\|_2^2 + 2\left(\sqrt{\pi^{(k)}(s)} - \sqrt{\pi^{(k')}(s)}\right)^2\left\|p^{(k')}(\cdot|s)\right\|_2^2 \\
&\leq 2\pi^{(k)}(s)\left\|p^{(k)}(\cdot|s) - p^{(k')}(\cdot|s)\right\|_2^2 + 2\left(\sqrt{\pi^{(k)}(s)} - \sqrt{\pi^{(k')}(s)}\right)^2 \\
&\leq \frac{2p_{\max}}{\log\frac{e}{2}}\pi^{(k)}(s)\,\mathsf{KL}\left(p^{(k)}(\cdot|s), p^{(k')}(\cdot|s)\right) + 2\left(\sqrt{\pi^{(k)}(s)} - \sqrt{\pi^{(k')}(s)}\right)^2 && \text{(Lemma D.1)}
\end{aligned}$$

We conclude the first part by summing over $s \in \mathcal{S}$ and taking a minimum over $k \neq k'$.

For the second part, we start by recalling Le Cam's inequality (Tsybakov, 2009, Lemma 2.3):

$$H^2\left(\pi^{(k)}, \pi^{(k')}\right) \leq \mathsf{TV}\left(\pi^{(k)}, \pi^{(k')}\right).$$

[7] Recall that the Markov chain with stationary distribution π and transition matrix P is uniformly ergodic if there exists $\rho \in (0,1)$ and $M > 0$ such that $\max_{s\in\mathcal{S}}\left\|P^H(s,\cdot),\pi\right\|_1 \leq M\rho^H$ for all $H \in \mathbb{N}$.

We utilize the perturbation bound for Markov chains (Mitrophanov, 2005, Corollary 3.1) to obtain

$$\begin{aligned}
\mathsf{TV}\left(\pi^{(k)}, \pi^{(k')}\right) &\leq \left(\lceil \log_\rho M^{-1} \rceil + \frac{1}{1-\rho}\right) \max_{s\in\mathcal{S}} \mathsf{TV}\left(p^{(k)}(\cdot|s), p^{(k')}(\cdot|s)\right) \\
&\leq \left(\lceil \log_\rho M^{-1} \rceil + \frac{1}{1-\rho}\right) \sqrt{\frac{1}{2} \max_{s\in\mathcal{S}} \mathsf{KL}\left(p^{(k)}(\cdot|s), p^{(k')}(\cdot|s)\right)} \qquad \text{(Pinsker's inequality)} \\
&\leq \left(\lceil \log_\rho M^{-1} \rceil + \frac{1}{1-\rho}\right) \sqrt{\frac{1}{2\pi_{\min}} \sum_{s\in\mathcal{S}} \pi^{(k)}(s)\, \mathsf{KL}\left(p^{(k)}(\cdot|s), p^{(k')}(\cdot|s)\right)} \\
&= \left(\lceil \log_\rho M^{-1} \rceil + \frac{1}{1-\rho}\right) \sqrt{\frac{1}{2\pi_{\min}} \mathcal{D}_\pi^{(k,k')}}.
\end{aligned}$$

As $\frac{2}{\log \frac{\varepsilon}{2}} < 7$, we finally have that

$$\begin{aligned}
\Delta_{\boldsymbol{W}}^2 &\leq 7 \min_{k\neq k'} \left\{ p_{\max} \mathcal{D}_\pi^{(k,k')} + \left(\lceil \log_\rho M^{-1} \rceil + \frac{1}{1-\rho}\right) \sqrt{\frac{1}{2\pi_{\min}} \mathcal{D}_\pi^{(k,k')}} \right\} \\
&= 7 \left(p_{\max} \mathcal{D}_\pi + \left(\lceil \log_\rho M^{-1} \rceil + \frac{1}{1-\rho}\right) \sqrt{\frac{1}{2\pi_{\min}} \mathcal{D}_\pi} \right),
\end{aligned}$$

where the last equality follows from the fact that $z \mapsto c_1 z + c_2 \sqrt{z}$ is strictly increasing for $z > 0$ and $c_1, c_2 > 0$. □

D.3 (c) Relationship between $\Delta_{\mathbf{W}}$ and Kausik et al. (2023)'s α, Δ

Similarly, for each $s \in \mathcal{S}$,

$$\begin{aligned}
&\left\| \sqrt{\pi^{(k)}(s)} p^{(k)}(\cdot|s) - \sqrt{\pi^{(k')}(s)} p^{(k')}(\cdot|s) \right\|_2^2 \\
&= \left\| \sqrt{\pi^{(k)}(s)} \left(p^{(k)}(\cdot|s) - p^{(k')}(\cdot|s)\right) + \left(\sqrt{\pi^{(k)}(s)} - \sqrt{\pi^{(k')}(s)}\right) p^{(k')}(\cdot|s) \right\|_2^2 \\
&\geq \frac{1}{2} \pi^{(k)}(s) \left\| p^{(k)}(\cdot|s) - p^{(k')}(\cdot|s) \right\|_2^2 - \left(\sqrt{\pi^{(k)}(s)} - \sqrt{\pi^{(k')}(s)}\right)^2 \left\| p^{(k')}(\cdot|s) \right\|_2^2 \\
&\geq \frac{1}{2} \pi^{(k)}(s) \left\| p^{(k)}(\cdot|s) - p^{(k')}(\cdot|s) \right\|_2^2 - \pi^{(k)}(s) \left(1 - \sqrt{\frac{\pi^{(k')}(s)}{\pi^{(k)}(s)}}\right)^2 \\
&\geq \frac{1}{2} \pi^{(k)}(s) \left\| p^{(k)}(\cdot|s) - p^{(k')}(\cdot|s) \right\|_2^2 - \pi^{(k)}(s) \left((\sqrt{\eta_\pi} - 1)^2 \vee \left(1 - \frac{1}{\sqrt{\eta_\pi}}\right)^2 \right). \qquad \text{(Assumption 2)}
\end{aligned}$$

Summing over $s \in \mathcal{S}$ and taking the min over $k \neq k'$, we have:

$$\begin{aligned}
\Delta_{\boldsymbol{W}}^2 &\geq \frac{1}{2} \sum_{s\in\mathcal{S}} \pi^{(k)}(s) \left\| p^{(k)}(\cdot|s) - p^{(k')}(\cdot|s) \right\|_2^2 - \left((\sqrt{\eta_\pi} - 1)^2 \vee \left(1 - \frac{1}{\sqrt{\eta_\pi}}\right)^2 \right) \\
&\geq \frac{1}{2} \alpha \Delta^2 - \left((\sqrt{\eta_\pi} - 1)^2 \vee \left(1 - \frac{1}{\sqrt{\eta_\pi}}\right)^2 \right). \qquad (45)
\end{aligned}$$

□

E PROOF OF PROPOSITION 5.2 – HARD INSTANCE WITH S-DEPENDENT SEPARATION ACROSS GAPS

Fix $S \geq 2$, $K \geq 1$, horizon H, and $0 < \zeta < 1/S$. For simplicity, we assume that $S \geq K$, and we also fix the initial distribution to be uniform, i.e., $\mu = \text{Unif}([S])$ with $\mathcal{S} = [S]$.

E.1 Construction of Cyclic-Bump MMC

We now introduce our construction, which we refer to **cyclic-bump MMC**. Intuitively, this is done by starting from an everywhere-uniform chain, then repeatedly adding doubly-stochastic perturbations via a small circulant flow of magnitude ζ.

Rigorously speaking, we first define the uniform chain's transition kernel $U = \frac{1}{S}\mathbf{1}\mathbf{1}^\top$, where $\mathbf{1}$ is an S-dimensional all-1 column vector. We set the first chain to be $P^{(1)} = U$. Then, define the remaining $K-1$ kernels as

$$P^{(k+1)} := U + \zeta\left(C^k - C^{k+1}\right), \tag{46}$$

where the k-step cyclic right-shift permutation matrix C^k is defined as

$$\left(C^k\right)_{i,j} := \mathbb{1}[j \equiv i + k \mod S]. \tag{47}$$

Claim E.1. *For all $k \geq 1$, $P^{(k)}$ defines an ergodic Markov chain with uniform stationary distribution: $\pi^{(k)}(\cdot|s) = \text{Unif}([S])$ for all $s \in [S]$.*

Proof. The statement is trivial for $k = 1$, and so we focus on $k \geq 2$.

$P^{(k)}$ is well-defined, doubly-stochastic transition matrix. Entrywise-positivity follows from our choice $\zeta < \frac{1}{S}$. As for doubly-stochasticity, first note that as each of C^k and C^{k+1} is a permutation matrix, every row and column sums to 1. Thus, every row and column of $C^k - C^{k+1}$ sums to 0. As U is doubly-stochastic, so is $P^{(k)}$.

Uniform Stationarity. Let $u^\top = \frac{1}{S}\mathbf{1}^\top$ be the row-vector describing the uniform distribution over $[S]$. Because $u^\top$ is invariant under any permutation, we have that $u^\top C^k = u^\top$. Hence

$$u^\top P^{(k)} \;=\; u^\top U + \zeta\big(u^\top C^k - u^\top C^{k+1}\big) \;=\; u^\top + \zeta(u^\top - u^\top) \;=\; u^\top,$$

so $\pi^{(k)} = u$ is stationary. Finally, as $P^{(k)}$ is entrywise strictly positive, it defines an ergodic Markov chain, hence $\pi^{(k)}$ (as defined previously) is indeed the unique stationary distribution of $P^{(k)}$. □

E.2 Explicit Computations of the Gaps

With the above construction, we now compute each gap explicitly.

Computing $\alpha\Delta^2$. First, we know that $\alpha = \frac{1}{S}$, and so it suffices to compute Δ^2, which is the (maximum) ℓ_2-gap of the outgoing transitions. From direct computations, we have that for $k, k' \geq 2$,

$$\left\|P^{(k)}(\cdot|s) - P^{(1)}(\cdot|s)\right\|_2^2 = 2\zeta^2$$

and

$$\left\|P^{(k)}(\cdot|s) - P^{(k')}(\cdot|s)\right\|_2^2 = \begin{cases} 4\zeta^2, & k' - k \neq \pm 1, \\ 6\zeta^2, & k' - k = \pm 1. \end{cases}$$

All in all, we have that $\alpha\Delta^2 = \frac{6\zeta^2}{S}$.

Computing $\Delta^2_{\boldsymbol{W}}$. By definition and the above computations, we have that

$$\begin{aligned}\Delta^2_{\boldsymbol{W}} &:= \min_{k\neq k'} \sum_{s\in\mathcal{S}} \left\| \sqrt{\pi^{(k)}(s)} P^{(k)}(\cdot|s) - \sqrt{\pi^{(k')}(s)} P^{(k')}(\cdot|s) \right\|_2^2 \\ &= \frac{1}{S} \min_{k\neq k'} \sum_{s\in\mathcal{S}} \left\| P^{(k)}(\cdot|s) - P^{(k')}(\cdot|s) \right\|_2^2 \\ &= \frac{1}{S} \sum_{s\in\mathcal{S}} 2\zeta^2 = 2\zeta^2.\end{aligned}$$

Computing D_π. From direct computations, we have that for $k, k' \geq 2$,

$$\begin{aligned}\mathsf{KL}\left(P^{(k)}(\cdot|s), P^{(1)}(\cdot|s)\right) &= \frac{1}{S}\log\big((1+\zeta S)(1-\zeta S)\big) + \zeta \log\frac{1+\zeta S}{1-\zeta S}, \\ \mathsf{KL}\left(P^{(1)}(\cdot|s), P^{(k)}(\cdot|s)\right) &= \frac{1}{S}\log\frac{1}{(1+\zeta S)(1-\zeta S)},\end{aligned}$$

and

$$\mathsf{KL}\left(P^{(k)}(\cdot|s), P^{(k')}(\cdot|s)\right) = \begin{cases} \zeta\log\frac{(1+\zeta S)^2}{1-\zeta S}, & k'-k=1, \\ \zeta\log\frac{1+\zeta S}{(1-\zeta S)^2}, & k'-k=-1, \\ \zeta\log\frac{1+\zeta S}{1-\zeta S}, & k'-k\neq\pm 1. \end{cases}$$

As the above computation holds across all s's and as $\pi^{(k)}$'s are all uniform, it suffices to take the minimum over the above five outcomes. As $\zeta S \in (0,1)$, it is clear that

$$\zeta\log\frac{(1+\zeta S)^2}{1-\zeta S} \vee \zeta\log\frac{1+\zeta S}{(1-\zeta S)^2} > \zeta\log\frac{1+\zeta S}{1-\zeta S} > \frac{1}{S}\log\big((1+\zeta S)(1-\zeta S)\big) + \zeta\log\frac{1+\zeta S}{1-\zeta S},$$

and thus, it suffices to compare between $\mathsf{KL}\left(P^{(k)}(\cdot|s), P^{(1)}(\cdot|s)\right)$ and $\mathsf{KL}\left(P^{(1)}(\cdot|s), P^{(k)}(\cdot|s)\right)$. The following claim, whose proof is deferred to the end of this subsection, shows that the former quantity is smaller than the latter:

Claim E.2. $\mathsf{KL}\left(P^{(k)}(\cdot|s), P^{(1)}(\cdot|s)\right) \leq \mathsf{KL}\left(P^{(1)}(\cdot|s), P^{(k)}(\cdot|s)\right)$.

All in all, we then have that

$$\mathcal{D}_\pi = \frac{1}{S}\log\big((1+\zeta S)(1-\zeta S)\big) + \zeta\log\frac{1+\zeta S}{1-\zeta S}.$$

Comparing everything. Let $\zeta = \frac{1}{2S}$ for simplicity.[8] Then, we have that

$$\alpha\Delta^2 = \frac{3}{2S^3}, \quad \Delta^2_{\boldsymbol{W}} = \frac{1}{2S^2}, \quad \mathcal{D}_\pi = \frac{1}{S}\log\frac{3\sqrt{3}}{4},$$

i.e., for this instance of clustering in MMC, we have that

$$S^2\alpha\Delta^2 \asymp S\Delta^2_{\boldsymbol{W}} \asymp \mathcal{D}_\pi. \tag{48}$$

□

Proof of Claim E.2. Let $z = \zeta S \in (0,1)$. Then, one can rewrite the inequality as

$$g(z) \triangleq 2\log(1-z^2) + z\log\frac{1+z}{1-z} \leq 0, \quad \forall z \in (0,1).$$

As $g(0) = 0$, it suffices to show that $g(\cdot)$ is monotone decreasing for $z \in (0,1)$. To see this, we compute its first and second derivatives:

$$g'(z) = \log\frac{1+z}{1-z} - \frac{2z}{1-z^2}, \quad g''(z) = -\frac{4z^2}{(1-z^2)^2} \leq 0.$$

As $g'(0) = 0$ as well, we are done. □

[8] One could do an asymptotic expansion of the KL-divergence for a more generic asymptotic comparison.

E.3 Discussions

Various Perspectives on Cyclic–Bump MMCs. Our construction is most naturally viewed as *circulant perturbation* of an ergodic Markov kernel with a divergence-free *circulation* on the transition graph: we add and cancel small cyclic shifts so that row and column sums remain unchanged and the stationary distribution stays the same, directly connecting to classical analyses of circulant/group-based chains and their spectra (Diaconis, 1988; Levin and Peres, 2017). Geometrically, the perturbation is a small, *local* move inside the Birkhoff–von Neumann polytope[9]: starting from the point U, we stay within the affine subspace of doubly-stochastic matrices by moving along a divergence-free direction given by a scaled *signed difference of permutation vertices*, which is interesting in its own right from a polyhedral viewpoint. The construction presented here shares principles with irreversible and lifted Markov Chain Monte Carlo (MCMC) methods. Specifically, these methods introduce a cycle flow to modify the chain's spectral properties and accelerate mixing, all while preserving the target distribution (Chen et al., 1999; Diaconis, 1988). Our work diverges from this literature in a crucial way: whereas lifting requires expanding to a larger state space (Chen et al., 1999), our construction modifies the chain while preserving the original state space $[S]$.

What if $K > S$? Our construction assumed $K \le S$ so that single-bump shifts can be chosen distinct; when $K > S$,[10] indices coincide modulo S (e.g., 1 and $K+1 \equiv 1 \mod K$), and the corresponding single-bump components become identical (hence non-identifiable) if they use the same amplitude. To retain distinctness while preserving the uniform stationary distribution and ergodicity, one can: (i) vary amplitudes across components, taking $P^{(k)} = U + \zeta_k\big(C^k - C^{k+1}\big)$ with $\zeta_k \neq \zeta_{k'}$ and $0 < \zeta_k < 1/S$; (ii) use multi-bump combinations $P^{(k)} = U + \sum_{r\in R_k} \zeta_{k,r}\big(C^{k+r} - C^{k+r+1}\big)$ with $\sum_r \zeta_{k,r} < 1/S$ and $R_k \subset [K]$; or (iii) adopt a general small circulant combination $P^{(k)} = U + \sum_{m=1}^{S-1} a_{k,m} C^m$ with $\sum_m a_{k,m} = 0$ and $\sum_{m=1}^{S-1} |a_{k,m}| < 1/S$ (a simple sufficient positivity condition). Option (iii) has a natural Fourier/representation-theoretic interpretation on the cyclic group (circulant kernels diagonalized by the discrete Fourier transform); see Diaconis (1988) for background.

Why D_π is Large while $\Delta^2_{\boldsymbol{W}}$ is Small? From an information–geometric perspective, the gap stems from the *Fisher information metric* underlying KL divergence (Amari, 2016; Nielsen, 2020). This metric reweights directions by the inverse probabilities, amplifying changes along low-mass coordinates. Rigorously, let g_u be the Fisher information metric on the probability simplex at the uniform distribution $u = (1/S, \dots, 1/S)$, which reduces to $g_u = \mathrm{diag}(1/u) = S\boldsymbol{I}$ on the tangent space – hence locally $\mathrm{KL}(p\|u) \approx \frac{1}{2}(p-u)^\top g_u (p-u) = \frac{S}{2}\|p-u\|_2^2$. The cyclic bump changes exactly two coordinates per row by $\pm\varepsilon$, which is a tiny *absolute* move – hence $\Delta^2_{\boldsymbol{W}} = \Theta(\varepsilon^2)$ – but a *relative* change of order S larger, i.e., $\mathcal{D}_\pi = \Theta(\varepsilon S)$. This is why the KL-based separation satisfies $D_\pi \asymp S\Delta^2_{\boldsymbol{W}}$ compared to the ℓ_2-based separation $\Delta^2_{\boldsymbol{W}}$.

[9] This is the convex polytope generated by permutation matrices (Birkhoff, 1946).

[10] Here, we denote $C^k := C^{k \mod S}$ for $k > S$.

F DEFERRED EXPERIMENTAL DETAILS & ADDITIONAL ABLATIONS

F.1 Deferred Experimental Details from Section 6

For the algorithms of Kausik et al. (2023), the number next to 'th' is the threshold level that is used to output an initial estimate of the clustering from pairwise distances; see line 16 of Algorithm 2 of Kausik et al. (2023). The "right" choice of the threshold level explicitly depends on the separation gap (see Kausik et al. (2023, Theorem 1)), and thus, it is inherently heuristic in the absence of such knowledge. For this reason, we arbitrarily sweep over $\{10^{-5}, 10^{-4}, 10^{-3}\}$. We vary $H \in \{10, 20, \cdots, 90, 100, 200, \cdots, 500\}$ and $T \in \{100, 200, \cdots, 600\}$. For statistical significance, we report results with $\delta = 0.05$, shading the 95% bootstrapped confidence intervals (Efron, 1979; DiCiccio and Efron, 1996) over 30 independent repeats for each configuration. Each repeat corresponds to a fresh random draw of T trajectories. The codes are available in our GitHub repository.[11]

F.2 Ablation #1. Impact of Number of Iterations for Stage II

We utilize the same instance of MMC as in Section 6. Figure 2 examines how the clustering error evolves with additional EM iterations in Stage II. While our theory only requires a single EM step, we observe consistent empirical gains from running more iterations. The benefit is most pronounced when H is small or T is large, where each additional iteration pushes the error curve closer to the oracle benchmark. The improvement largely saturates after about 5–6 iterations, suggesting diminishing returns beyond this point. Importantly, all Stage II variants still outperform Stage I alone, underscoring the added value of EM refinement.

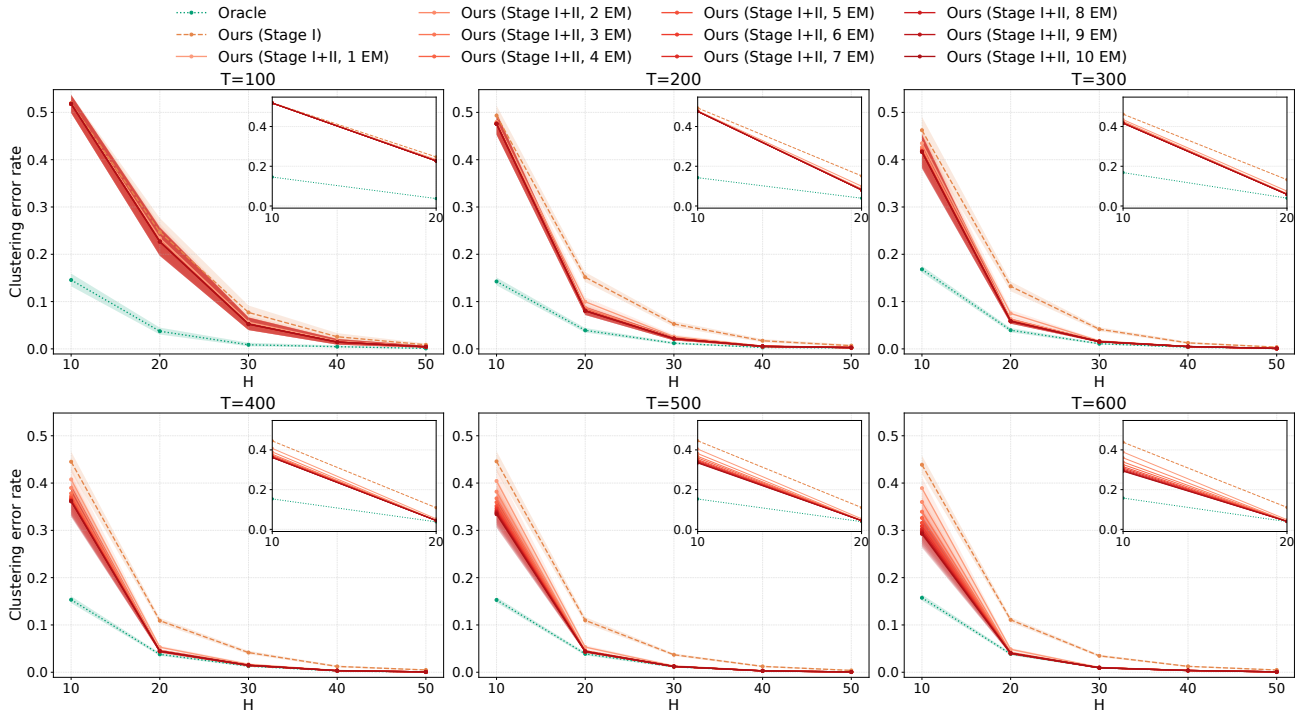

Figure 2: **(Ablation #1)** Clustering error on the synthetic MMC instance across EM iterations (up to 10). Panels vary across $T \in \{100, 200, \cdots, 600\}$; the x-axis varies across $H \in \{10, 20, \cdots, 50\}$. Darker red curves indicate more EM iterations; "Oracle" and "Stage I" are shown for reference.

F.3 Ablation #2. Impact of S and T on Performance and Runtime

To analyze the computational complexity of our two-stage algorithm, we measure wall-clock runtime as a function of the state-space size S and the number of trajectories T. We decompose the total runtime into four components:

[11] https://github.com/nick-jhlee/optimal-mixture-markov

(i) embedding construction (line 4 of Algorithm 1), (ii) singular value decomposition (SVD) (line 5), (iii) K-means clustering (lines 6–7), and (iv) likelihood reassignment (Algorithm 2). For the experiment varying S, we fix $T = 400$ and $H = 500$, and sweep $S \in \{10, 20, 30, 40, 50, 60\}$. For the experiment varying T, we fix $S = 40$ and $H = 500$, and sweep $T \in \{100, 200, 300, 400, 500, 600\}$. In both settings, we run 10 EM iterations in Stage II and average over 30 independent trials.

Figure 3 below summarizes the results. The top row shows that clustering performance degrades as S increases (left), due to the decreasing separation $\zeta = 0.9/S$ in our MMC construction, whereas performance improves as T increases (right), as expected. The middle row reveals that SVD dominates the runtime for large S (left), while likelihood reassignment dominates for large T (right). The bottom row quantifies the overall scaling: total runtime grows quadratically in S with $R^2 \approx 0.9583$, confirming the $\mathcal{O}(TS^2)$ SVD cost, and linearly in T with $R^2 \approx 0.9957$, consistent with the $\mathcal{O}(T)$ scaling of all algorithmic steps.

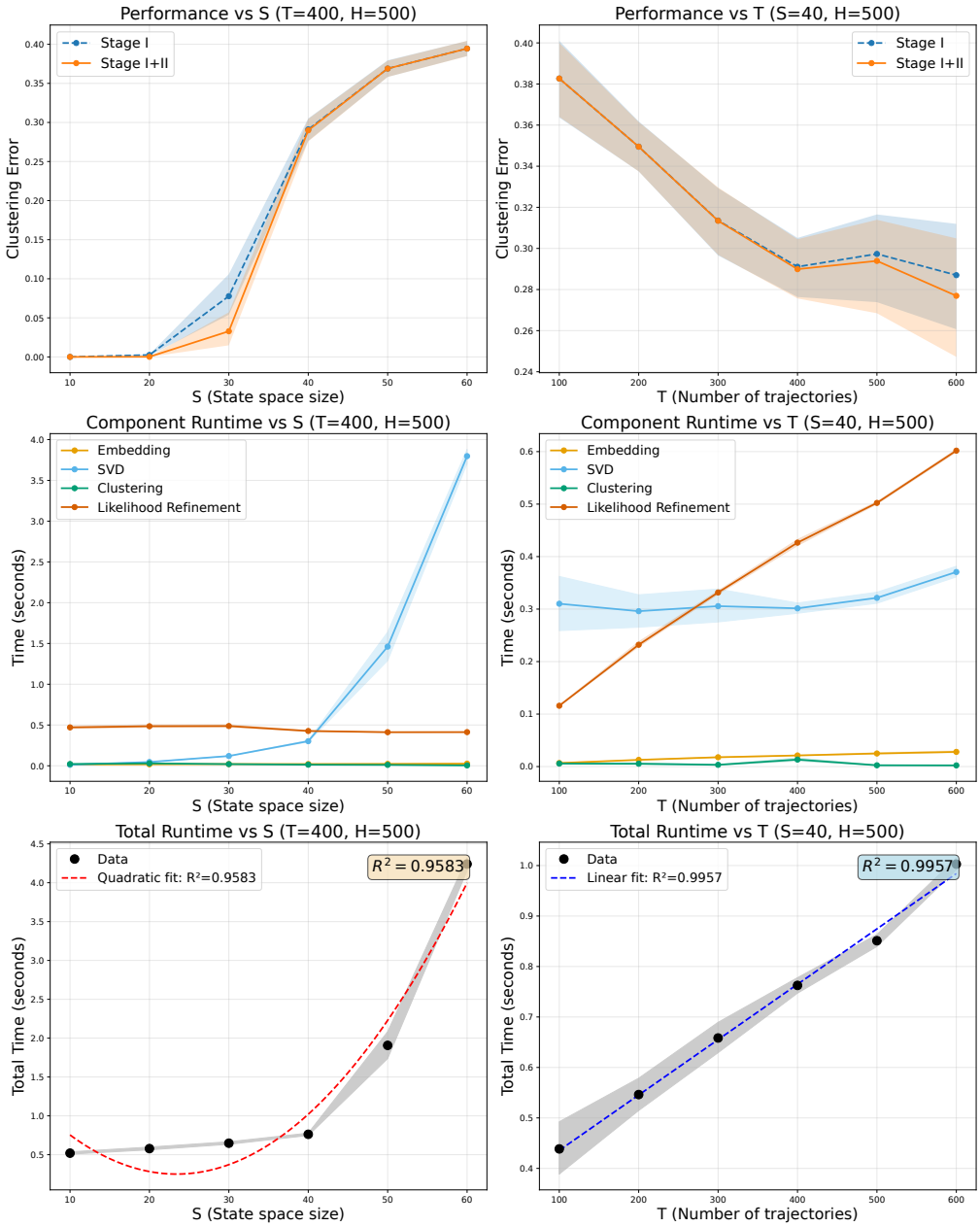

Figure 3: **(Ablation #2)** Impact of S and T on the resulting clustering error and runtime.

F.4 Ablation #3. Our Algorithm with Unknown K

Again, in the same instance of MMC as in Section 6, we evaluate the performance of our method when the number of clusters K is unknown a priori. Recall that Algorithm 1 uses two fixed constants (64 for adaptive singular value thresholding at line 6, and 32 for estimating K at lines 10–14). For this ablation, we replace these fixed values with tunable hyperparameters c_1 and c_2, respectively. In practice, careful tuning of (c_1, c_2) is essential, since the performance of the unknown-K variant is sensitive to the empirical tightness of the concentration bounds that these parameters effectively calibrate. Within a reasonable range, we believe that adjusting (c_1, c_2) does not alter the theoretical validity of our guarantee; the only difference is that the failure probability scales as $\Theta(\delta)$ rather than exactly δ.

We sweep over $c_1 \in \{10^{-3}, 10^{-2}, 10^{-1}\}$ and $c_2 \in \{10^{-3}, 5\cdot10^{-3}, 10^{-2}, 5\cdot10^{-2}, 10^{-1}\}$, and extend the trajectory length to $H \in \{100, 200, \ldots, 1000\} \cup \{1000, 2000, \ldots, 20000\}$ in order to probe the large-H regime where theory predicts consistent recovery of K. We fix $T = 100$, and Stage II is fixed to run 10 EM iterations. The known-K variant achieves perfect clustering and is therefore omitted from comparison.

Figure 4 summarizes the results. For small H (top row), the unknown-K variant often fails to recover clusters reliably, especially for moderate choices $c_1 \in \{10^{-2}, 10^{-1}\}$. In contrast, a more aggressive setting $c_1 = 10^{-3}$ succeeds even in the 100's regime across all considered c_2 values, indicating that performance is substantially more sensitive to c_1 than to c_2. In the large-H regime (bottom row), $c_1 = 10^{-2}$ also succeeds consistently across c_2, whereas $c_1 = 10^{-1}$ continues to fail, again underscoring the high sensitivity of empirical performance to the choice of c_1. While there remains some dependence on c_2—as reflected in differing convergence speeds with increasing H—the overall recovery trends are qualitatively similar.

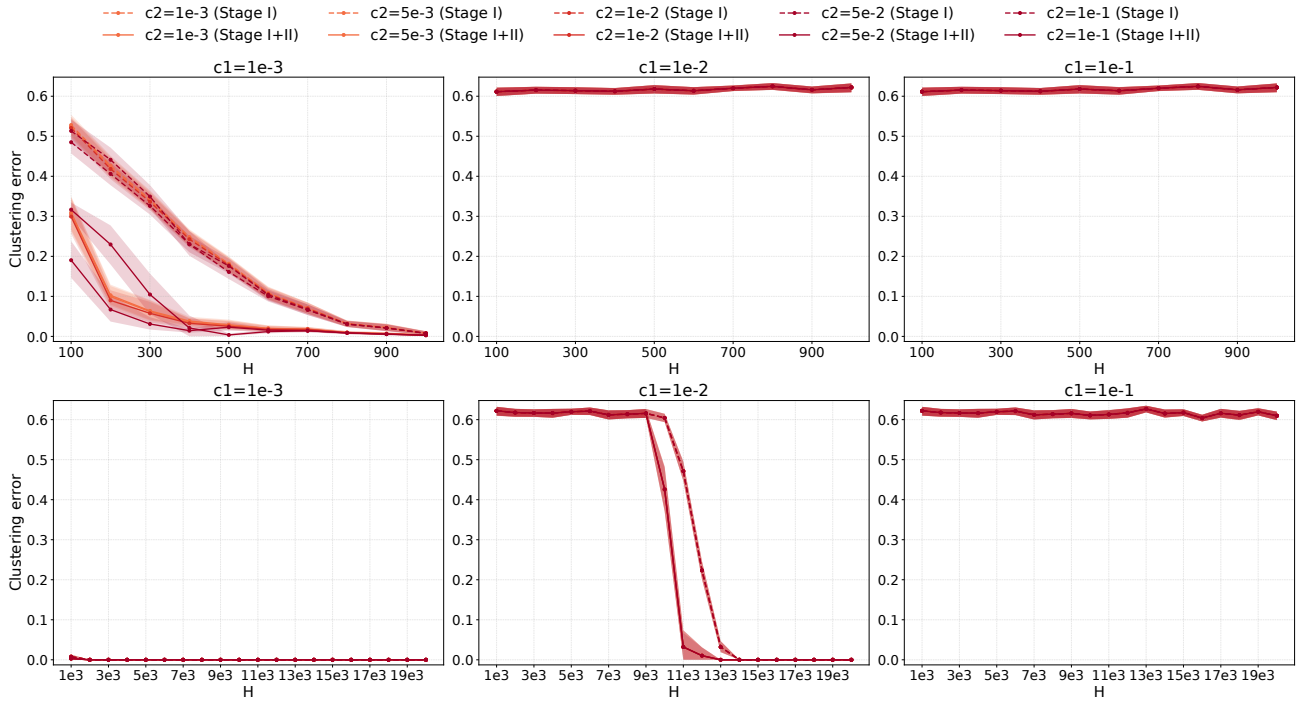

Figure 4: **(Ablation #3)** Clustering error of the unknown-K variant on the synthetic MMC instance across different (c_1, c_2) settings. Panels vary across $c_1 \in \{10^{-3}, 10^{-2}, 10^{-1}\}$ (columns) and trajectory length H (rows: top for $H \leq 1000$, bottom for $H \geq 1000$). Curves indicate different c_2 values; Stage I and Stage I+II are shown separately.

F.5 Ablation #4. Impact of γ_{ps}

To demonstrate that the dependence on the pseudo-spectral gap γ_{ps} in our final performance guarantee (Theorem 4.2) is fundamental (at least with our current algorithm design), we construct a sequence of MMC instances

where the information-theoretic gap $\mathcal{D}_\pi$ is strictly constant, but γ_{ps} varies.

We propose the "Signal in a Maze" construction. Let $S = 10$ and $K = 2$. We denote the state space as $\mathcal{S} = \{s_1, s_2, \ldots, s_{10}\}$. We partition this space into a signal region $\mathcal{S}_{\text{diff}} = \{s_1, s_2\}$ and a mixing maze $\mathcal{S}_{\text{mix}} = \{s_3, \ldots, s_{10}\}$. Conceptually, the *signal* region contains the only states where the two underlying Markov chains exhibit distinct transition behaviors. The *maze* acts as a sequence of identical, slow-mixing states that artificially bottlenecks the traversal of the chain without altering the stationary distribution. Finally, a small *leakage* parameter ϵ dictates the probability of transitioning between the signal region and the maze, ensuring the chain remains ergodic while compartmentalizing the signal.

We fix the leakage parameter $\epsilon = 0.05$ and distinct transition probabilities $p_1 = 0.4$ and $p_2 = 0.3$. We parameterize the sequence of MMCs by a bottleneck variable $q \in (0, 0.5)$, which dictates the traversal speed through the maze. For each $k \in \{1, 2\}$ and given $q \in (0, 0.5)$, we define the transition matrix $\boldsymbol{P}_q^{(k)}$ using a block-matrix structure:

$$\boldsymbol{P}_q^{(k)} = \begin{bmatrix} \boldsymbol{B}^{(k)} & \boldsymbol{E}^\top \\ \boldsymbol{E} & \boldsymbol{T}_q \end{bmatrix},$$

where the constituent blocks are defined as follows:

1. **Signal Block ($\boldsymbol{B}^{(k)} \in \mathbb{R}^{2\times 2}$):** Governs transitions within $\mathcal{S}_{\text{diff}}$ and dictates the distinguishing behavior between the chains:
$$\boldsymbol{B}^{(k)} = \begin{bmatrix} 1 - p_k - \epsilon & p_k \\ p_k & 1 - p_k - \epsilon \end{bmatrix}.$$

2. **Leakage Block ($\boldsymbol{E} \in \mathbb{R}^{8\times 2}$):** Connects the signal region to the ends of the maze. All entries are zero except for the entries mapping $s_3 \leftarrow s_1$ and $s_{10} \leftarrow s_2$, namely $\boldsymbol{E}_{1,1} = \epsilon$ and $\boldsymbol{E}_{8,2} = \epsilon$.

3. **Maze Block ($\boldsymbol{T}_q \in \mathbb{R}^{8\times 8}$):** A symmetric tridiagonal matrix representing the path graph of $\mathcal{S}_{\text{mix}}$, independent of k:
$$\boldsymbol{T}_q = \text{tridiag}(q, 1 - 2q, q) + \text{diag}(-\epsilon, 0, \ldots, 0, -\epsilon).$$
Explicitly, the main diagonal is $(1 - q - \epsilon, 1 - 2q, \ldots, 1 - 2q, 1 - q - \epsilon)$ and the first off-diagonals are all q.

By construction, $\boldsymbol{P}_q^{(k)}$ is symmetric and doubly stochastic. Thus, the stationary distribution is strictly uniform: $\pi^{(1)} = \pi^{(2)} = \pi = [1/10, \ldots, 1/10]^\top$, which is entirely independent of q. Consequently, the stationary-weighted KL divergence evaluates to:

$$\mathcal{D}_\pi = \frac{1}{10}\,\mathsf{KL}\left(\boldsymbol{P}_q^{(1)}(s_1, \cdot), \boldsymbol{P}_q^{(2)}(s_1, \cdot)\right) + \frac{1}{10}\,\mathsf{KL}\left(\boldsymbol{P}_q^{(1)}(s_2, \cdot), \boldsymbol{P}_q^{(2)}(s_2, \cdot)\right).$$

Because the bottleneck parameter q only appears in the $\boldsymbol{T}_q$ block (which is identical across $k = 1, 2$), $\mathcal{D}_\pi$ remains exactly constant for all q.

However, as $q \to 0$, the pseudo-spectral gap γ_{ps} strictly shrinks, requiring trajectories of increasingly longer length H to traverse the maze and observe the signal states. For $q \in \{0.01, 0.03, 0.05, 0.07, 0.09, 0.1, 0.2\}$ (which is the range of q's at which we evaluate our algorithm), the pseudo-spectral gaps are numerically computed as:

$$\gamma_{\text{ps}} \in \{0.008399, 0.023994, 0.037832, 0.049807, 0.059931, 0.064334, 0.090893\}.$$

We evaluate our algorithm on the above instance over $H \in \{100, 200, \ldots, 2000\}$, with 30 repeats per configuration. As shown in Figure 5, the rate of error decay against H is slower with lower γ_{ps} (i.e., higher mixing time).

F.6 Additional Experiment with Last.fm 1K

Preprocessing. For this dataset, we follow the same pre-processing steps of Kausik et al. (2023), as provided in their GitHub repository.[12] We explain the preprocessing pipeline here for completeness. We use the Last.fm

[12] https://github.com/hetankevin/mdpmix

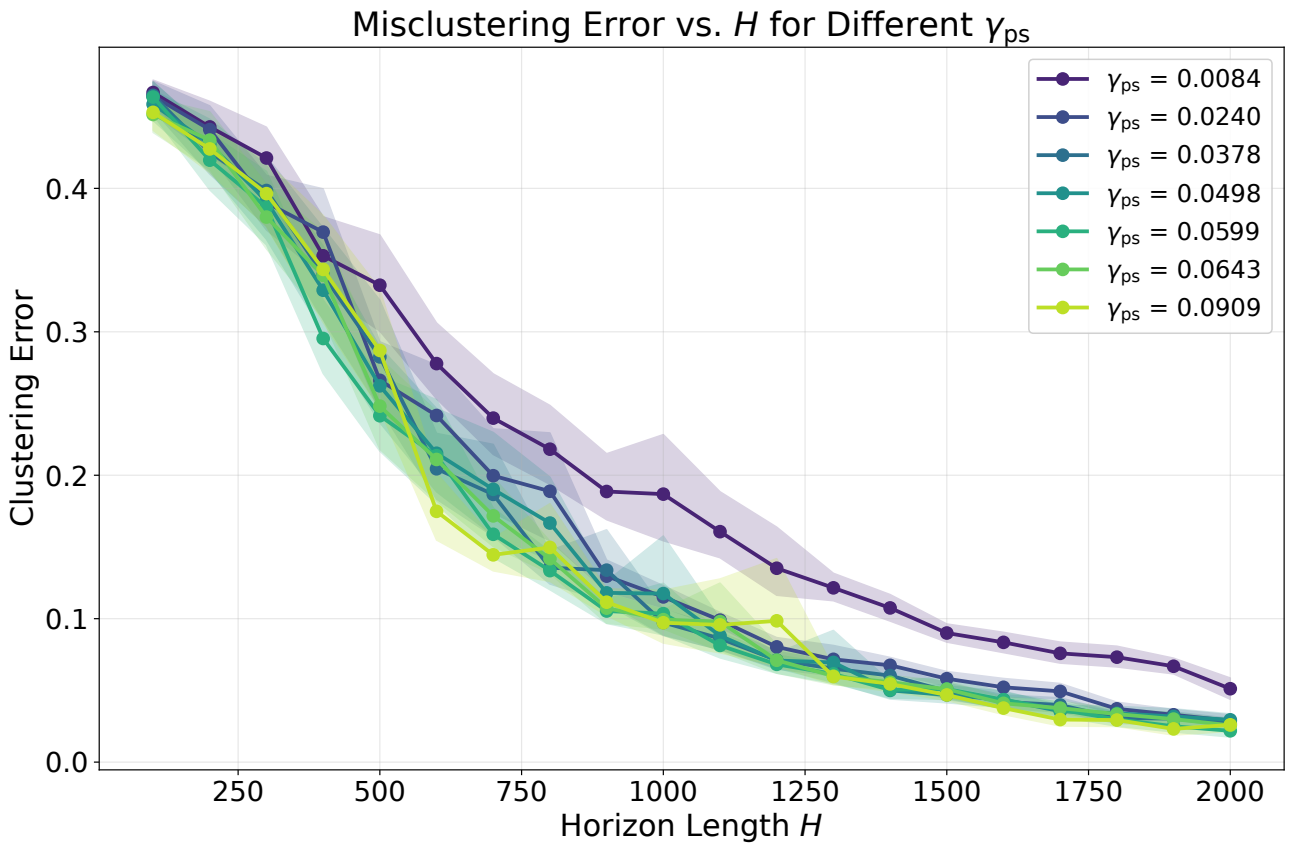

Figure 5: **(Ablation #4)** Clustering error on another synthetic MMC instances with varying γ_{ps}'s, over $H \in \{100, 200, \cdots, 2000\}$.

1K listening logs[13] (`userid-timestamp-artid-artname-traid-traname.tsv`) and the Lastfm-ArtistTags2007 files[14] `tags.txt` and `ArtistTags.dat`. The state space is the set of the top 100 tags (genres) from `tags.txt`. Each artist is mapped to exactly one genre by retaining, for each artist, the tag with the highest `rawtagcount` among the top 100 tags. We retain the top 10 users by the number of listening events, merge their events with artist tags, and sort the results by user and timestamp. We keep only transitions where the genre *changes* (and the user is unchanged), so consecutive plays of the same genre are collapsed into a single state. From each user, we then form non-overlapping trajectories of length H along this sequence, yielding 75 trajectories per user; the true cluster label of a trajectory is the user index. Trajectories are optionally shuffled with a fixed seed for evaluation. This yields $K = 10$ clusters, $S = 100$ states, horizon H, and total number of trajectories $T = 750$ (with next-step trajectories) for the mixture-of-Markov-chains experiment.

We vary the horizon $H \in \{10, 20, \ldots, 100\}$, repeat each setting 30 times with different seeds, and report permutation-invariant cluster error rate with bootstrap 95% confidence intervals. We compare our method (Stage I: spectral clustering only; Stage I+II: spectral clustering followed by 10 EM iterations) to the algorithms of Kausik et al. (2023) (clustering-only and clustering + 10 EM iterations), with thresholds sweeped over $\{10^{-6}, 5 \cdot 10^{-5}, 10^{-5}\}$. For all algorithms, we use Laplace smoothing with $\lambda = 0.1$ for numerical stability.

Results. Figure 6 below shows the overall results. We first note that we reproduce the general trend of the algorithms reported in Kausik et al. (2023), as shown in their Figures 4 and 5. Secondly, note that our algorithms, both Stage I and I+II, significantly outperform those of Kausik et al. (2023). This demonstrates that our algorithm is also practically efficient on real-world datasets. Importantly, our algorithm (with Laplace smoothing) is shown to perform beyond the theoretically predicted regime, i.e., when the regularity (Assumption 2) fails and when H is small.

[13] http://mtg.upf.edu/static/datasets/last.fm/lastfm-dataset-1K.tar.gz
[14] https://web.archive.org/web/20110827230919/http://static.echonest.com/Lastfm-ArtistTags2007.tar.gz

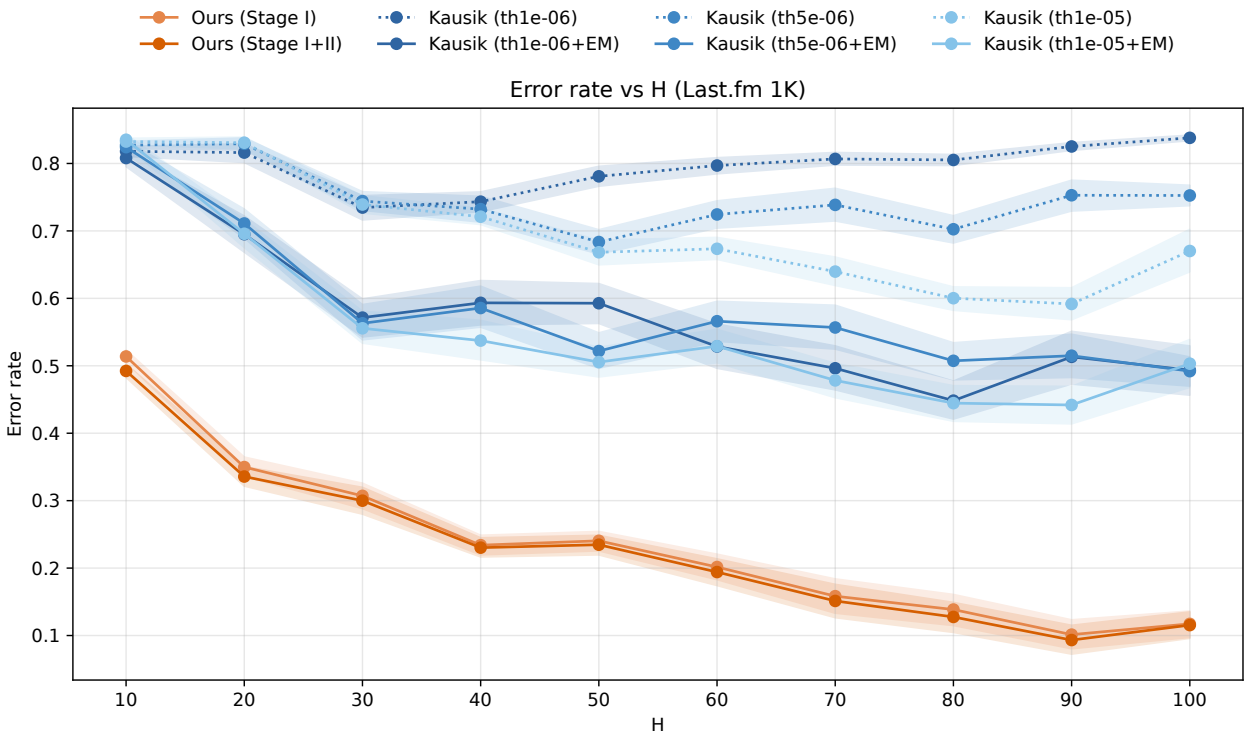

Figure 6: **(Last.fm 1K)** Clustering errors for Last.fm 1K, over $H \in \{10, 20, \cdots, 100\}$.

G RELATIONSHIPS BETWEEN OUR L-EMBEDDING AND OTHER SPECTRAL EMBEDDINGS

Diffusion Map (Coifman and Lafon, 2006). The objective of a diffusion map fundamentally differs from our approach. Let $\mathcal{X}$ be a measure space endowed with a kernel $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_{\geq 0}$. Diffusion maps provide a meaningful Euclidean embedding of *data points* $\mathcal{D} \subseteq \mathcal{X}$ via a reversible Markov process (random walk) over $\mathcal{X}$, induced by the kernel k which encodes prior knowledge of the local geometry. Thus, its purpose is not to embed a specific Markov chain, but rather to map a measure space onto a Euclidean space using a random walk.

(Weighted) Laplacian (Chung, 1997, Chapter 1.5). Similarly, the graph Laplacian is tailored for the spectral analysis of weighted undirected graphs. While it offers an elegant probabilistic interpretation of eigenvalues and eigenvectors, it is not directly applicable as an embedding for general, potentially non-reversible Markov chains.

Symmetrized Form (Levin and Peres, 2017, Chapter 12). Recall that the symmetrized form is $\boldsymbol{D}_\pi^{1/2}\boldsymbol{P}\boldsymbol{D}_\pi^{1/2}$, where $\boldsymbol{D}_\pi = \mathrm{diag}(\pi)$. The left factor, $\boldsymbol{D}_\pi^{1/2}\boldsymbol{P}$, when vectorized, corresponds precisely to our L-embedding. To the best of our knowledge, however, no prior work has analyzed the properties of the mapping $S : \mathcal{M} \mapsto \boldsymbol{D}_\pi^{1/2}\boldsymbol{P}\boldsymbol{D}_\pi^{1/2}$ as a Euclidean embedding for Markov chains. Specifically, its injectivity and concentration properties remain systematically unexplored. In contrast, for our L-embedding, we explicitly establish injectivity (Proposition 4.1) and derive sharp concentration bounds (Lemma B.1), both of which are critical for our Stage I analysis (Theorem 4.1).

Doublet Frequency (Vidyasagar, 2014; Wolfer and Kontorovich, 2021). Defined as $\boldsymbol{D}_\pi\boldsymbol{P}$, the doublet frequency closely resembles our L-embedding. It satisfies injectivity up to the initial distribution and exhibits similarly well-behaved concentration properties. However, its induced ℓ_2-gap scales differently and does not align as naturally with the information-theoretic gap $\mathcal{D}_\pi$ or Kausik et al. (2023)'s notion og gap $\alpha\Delta^2$, whereas our L-embedding explicitly facilitates these direct comparisons via $\Delta_{\boldsymbol{W}}$ (Section 5**(3)**).

H RANK OF $\boldsymbol{W}$ MAY BE MUCH SMALLER THAN THE NUMBER OF CLUSTERS K

Here, we show this via explicit constructions of $K+1$ transition matrices whose induced $\boldsymbol{W}$ is of rank 2.

Let P, P' be two distinct ergodic transition matrices with a same stationary distribution π. Define the other transition matrices as

$$P^{(k)} = \left(1 - \frac{k-1}{K}\right) P + \frac{k-1}{K} P', \quad k \in \{1, 2, \cdots, K+1\}. \tag{49}$$

As π is also a stationary distribution of $P^{(k)}$ and ergodicity is preserved under convex combination, it can be easily seen that every row of $\boldsymbol{W}$ is a convex combination of the 1-st and $(K+1)$-th row. In other words, $\text{rank}(\boldsymbol{W}) = 2$, yet there are $K+1$ distinct transition matrices.

Due to its simple form, we can also compute $\Delta^2_{\boldsymbol{W}}$ in a closed form:

$$\begin{aligned}
\Delta^2_{\boldsymbol{W}} &= \min_{k \neq k'} \sum_{s \in \mathcal{S}} \pi(s) \left\| p^{(k)}(\cdot|s) - p^{(k')}(\cdot|s) \right\|_2^2 \\
&= \min_{k \neq k'} \sum_{s \in \mathcal{S}} \pi(s) \left\| \frac{k'-k}{K} p(\cdot|s) + \frac{k-k'}{K} p'(\cdot|s) \right\|_2^2 \\
&= \sum_{s \in \mathcal{S}} \pi(s) \left\| p(\cdot|s) - p'(\cdot|s) \right\|_2^2 \min_{k \neq k'} \left(\frac{k'-k}{K} \right)^2 \\
&= \frac{1}{K^2} \sum_{s \in \mathcal{S}} \pi(s) \left\| p(\cdot|s) - p'(\cdot|s) \right\|_2^2 > 0.
\end{aligned}$$

I RELAXING THE KNOWLEDGE OF γ_{ps}

We first recall the naïve idea briefly sketched in Section 5**(4)**:

1. For each $t \in [T]$, compute the estimator $\hat{\gamma}_{\mathrm{ps},t}$ of Wolfer and Kontorovich (2024) that estimates $\hat{\gamma}_{\mathrm{ps}}^{(f(t))}$ while satisfying the following:
$$\mathbb{P}\left(\left|\hat{\gamma}_{\mathrm{ps},t} - \gamma_{\mathrm{ps}}^{(f(t))}\right| \leq \frac{1}{2}\gamma_{\mathrm{ps}}^{(f(t))}\right) \geq 1 - \frac{\delta}{T}. \tag{50}$$
2. Output $\hat{\gamma}_{\mathrm{ps}} := \frac{2}{3}\min_{t\in[T]} \hat{\gamma}_{\mathrm{ps},t}$.

The union bound then immediately implies that $\hat{\gamma}_{\mathrm{ps}}$ satisfies $\mathbb{P}(\hat{\gamma}_{\mathrm{ps}} \leq \gamma_{\mathrm{ps}}) \geq 1-\delta$. However, for Eqn. (50) to hold, we require $H \gtrsim \gamma_{\mathrm{ps}}^{-3}\,\pi_{\min}^{-1}\,\log\frac{1}{\pi_{\min}}\,\log\frac{T}{\pi_{\min}}\,\log\frac{T}{\pi_{\min}\gamma_{\mathrm{ps}}}$ (Wolfer and Kontorovich, 2024, Theorem 1). Compared to our requirement $H \gtrsim \gamma_{\mathrm{ps}}^{-1}\,\pi_{\min}^{-1}\,\log T$ (Theorem 4.2), this introduces an extra $\gamma_{\mathrm{ps}}^{-2}$ factor and an *extra* $\log T$ factor. As a lower bound of $H \gtrsim \gamma_{\mathrm{ps}}^{-3} S \log\frac{T}{S}$ is known (Wolfer and Kontorovich, 2019, Theorem 4), we conjecture that the extra $\gamma_{\mathrm{ps}}^{-2}$ factor is unavoidable.

As for the latter, we now describe a simple approach that could mitigate the extra $\log T$ factor:

1. Sample an index set $\mathcal{I} \subset [T]$ of cardinality $|\mathcal{I}| = \lceil \alpha_{\min}^{-1}\log\frac{K}{\delta}\rceil$, uniformly at random.
2. Repeat the "naïve" approach, with $\frac{\delta}{T}$ replaced with $\frac{\delta}{|\mathcal{I}|}$ in step 1 and $[T]$ replaced with $\mathcal{I}$ in step 2.

The following lemma shows that with high probability, $\mathcal{I}$ "covers" all K chains:

Lemma I.1. $\mathbb{P}\left(\mathcal{I} \cap f^{-1}(k) \neq \emptyset,\ \forall k \in [K]\right) \geq 1-\delta$.

Under this event, then note that we only need to take the union bound over $\mathcal{I}$, not $[T]$. This then results in the *additional*[15] requirement of $H \gtrsim \gamma_{\mathrm{ps}}^{-3}\,\pi_{\min}^{-1}\,\log\frac{1}{\pi_{\min}}\,\log\frac{\log\frac{K}{\delta}}{\alpha_{\min}\pi_{\min}}\,\log\frac{\log\frac{K}{\delta}}{\alpha_{\min}\pi_{\min}\gamma_{\mathrm{ps}}}$, which becomes *free* of any $\log T$ factors.

We leave to future work the precise details of the proof of our entire algorithm (γ_{ps} estimation, Stage I, and Stage II), but we believe that it should be straightforward.

Proof of Lemma I.1. This follows from a simple combinatorial argument:

$$\begin{aligned}
\mathbb{P}\left(\exists k \in [K]\ :\ \mathcal{I} \cap f^{-1}(k) = \emptyset\right) &\leq \sum_{k\in[K]} \mathbb{P}\left(\mathcal{I} \cap f^{-1}(k) = \emptyset\right) && \text{(union bound)}\\
&= \sum_{k\in[K]} \frac{\binom{T-\alpha_k T}{|\mathcal{I}|}}{\binom{T}{|\mathcal{I}|}}\\
&= \sum_{k\in[K]} \prod_{i=0}^{|\mathcal{I}|-1} \frac{T-\alpha_k T - i}{T-i}\\
&\leq \sum_{k\in[K]} (1-\alpha_k)^{|\mathcal{I}|} && \text{(the multiplicands monotonically decrease with } i\text{)}\\
&\leq K e^{-\alpha_{\min}|\mathcal{I}|}.
\end{aligned}$$

Setting $|\mathcal{I}| = \lceil \alpha_{\min}^{-1}\log\frac{K}{\delta}\rceil$ and simplifying gives the desired statement. □

[15] Here, we emphasize that this requirement is *in addition* to that of our Theorem 4.2.

J ADDITIONAL FUTURE DIRECTIONS

In this section, we present several directions for future work that extend beyond those mentioned in the main text.

(1) Towards Computationally Efficient Algorithms. As noted in Section 5**(4)**, our algorithm prioritizes parameter-free design and statistical efficiency over computational scalability. The construction of the empirical data matrix $\widehat{\boldsymbol{W}}$ and its subsequent exact SVD in Stage I entail $\mathcal{O}(TS^2)$ space and $\mathcal{O}\big(TS^2 \min(T, S^2)\big)$ time complexities, respectively. For environments with massive state spaces or trajectory counts, this can become prohibitively heavy.

To alleviate this computational burden, future large-scale deployments could integrate randomized linear algebra techniques, such as matrix sketching or fast low-rank approximations (Woodruff, 2014). Alternatively, one could adapt the clustering pipeline to streaming environments (Yun et al., 2014), which would significantly reduce the memory footprint. Exploring whether incorporating such approximations inherently induces a fundamental trade-off between computational scalability and the tight statistical efficiency bounds established in this work remains an important open question.

(2) Further Study of the L-Embedding. It is natural to explore how our L-based ℓ_2 distance (Definition 4.1) relates to broader information divergences beyond KL (Wang and Choi, 2023), as well as to bisimulation metrics (Calo et al., 2024, 2025). Even modest equivalence or comparison results could clarify when L-geometry best captures dynamics and when alternative metrics offer sharper statistical or algorithmic behavior.

Relatedly, it is important to determine whether the requirement $H \gtrsim 1/\Delta_{\boldsymbol{W}}^2$—or, more generally, an inverse-squared dependence on an ℓ_2 separation induced by any Euclidean embedding of the chains—is fundamental or merely an artifact of our two-stage design. Establishing either (i) a matching lower bound phrased in terms of $\Delta_{\boldsymbol{W}}$ (or any embedding that dominates it), or (ii) an alternative initialization whose sample complexity depends directly on the information gap $\mathcal{D}_\pi$ rather than an ℓ_2 proxy, would sharpen statistical limits and guide the choice of representations. We believe this is related to refining the lower bound by considering the uncertainty of the probability kernels; see the last paragraph of Section 5**(1)**.

(3) Extension to MDPs. Extending our approach to estimation *and* learning in mixtures of MDPs (Banerjee et al., 2025; Kausik et al., 2023) is another important direction. Progress here would connect with spectral methods in structured RL (e.g., POMDPs (Azizzadenesheli et al., 2016, 2017; Jedra et al., 2023), HMMs (Anandkumar et al., 2012; Hsu et al., 2012), and latent MDPs (Kwon et al., 2021, 2024b,a)) and could reveal performance separations analogous to those observed when state abstraction or clustering is known *a priori*. We believe that this extension would yield unique challenges due to controllable actions and policy-induced distributions (Jedra et al., 2023); even harder when one considers history-dependent policies and active exploration-style approaches (Tarbouriech and Lazaric, 2019; Tarbouriech et al., 2020).