The Type Theory of MDPs: Rare Events and Uncertainty Quantification

Imon Banerjee
Department of Statistics
Purdue University
West Lafayette, Indiana, USA 47906

Sayak Chakrabarty
Department of Computer Science
Northwestern University
Evanston, Illinois, USA 60201

Riddhiman Bhattacharya
Department of Statistics
University of California Santa Cruz
Santa Cruz, California, USA 95064

Ramkrishna Jyoti Samanta
Department of Statistical Science
University College London
London, UK WC1E 6BT

Abstract

The method of types, pioneered by Csiszar [1998], uses combinatorial arguments to approximate the probabilities of empirical measures and has become a standard tool for deriving error rates in hypothesis testing, source coding, and related information-theoretic problems. Although the technique is widely used across machine learning, computer science, and econometrics, rigorous type-theoretic guarantees for stationary Markov decision processes and reinforcement learning remain limited. This paper addresses that gap by developing a method-of-types framework for the empirical pair measure of the state–action process of a stationary Markov decision process. We establish a finite-sample upper bound and a matching subsequential lower exponent for the state–action process and characterize the associated entropy cost through the conditional Kullback–Leibler divergence relative to the true pair kernel and transfer it to transition kernels via a normalisation map. This allows us to analyze plug-in value- and Q-function estimators. Finally, we demonstrate the relevance of the framework by deriving explicit coding error exponents, highlighting its usefulness for computer science and machine-learning applications.

1 Introduction

Markov decision processes (MDPs) [Howard and Matheson, 1972, Howard, 1960] form a foundational model for reinforcement learning (RL) and sequential decision making [Sutton and Barto, 2018]. We write a stationary controlled process as $\{(X_i, a_i)\}_{i\geq 0}$, where X_i denotes the state and a_i denotes the action, taking values in the finite state–action space $\chi \times \mathbb{I}$. Since the work of Borkar [1988], estimation problems for MDPs have been studied extensively; recent developments include Li et al. [2022a,b], Banerjee et al. [2025a,b]. A core statistical object in this literature is the collection of empirical pair measures and empirical transition measures generated by the observed state–action trajectory. The objective of this paper is to develop a method-of-types framework for these empirical objects, with particular emphasis on their finite-sample type probabilities and the associated entropy-cost functionals.

To formally introduce our results, we fix some notation. Let $\{(X_i, a_i)\}_{i\geq 0}$ be a stationary MDP. In particular,

$$\mathbb{P}\left(X_i = t \mid X_{i-1} = s, a_{i-1} = l\right) = M^{(l)}_{s,t},$$

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

where, for every $(s, l) \in \chi \times \mathbb{I}$, $M_{s,t}^{(l)} \geq 0$ and $\sum_t M_{s,t}^{(l)} = 1$.

Remark 1.1. Beyond stationarity, we impose no parametric structure on the transition kernels or on the stationary action-selection rule. Thus, our framework is fully nonparametric in nature.

Now, with $N_s^{(l)}$ denote the number of visits to the state–action pair (s, l), and $N_{s,t}^{(l)}$ denote the number of observed transitions from s to t under action l, the empirical transition kernel is

$$\hat{M}_{s,t}^{(l)} = \frac{N_{s,t}^{(l)}}{N_s^{(l)}},$$

whenever $N_s^{(l)} > 0$, with the usual convention on unvisited state–action pairs. We write $\hat{\mathbb{M}} := (\hat{M}^{(1)}, \ldots, \hat{M}^{(|\mathbb{I}|)})$ for the resulting empirical MDP transition kernel, where $\hat{M}^{(l)} = [\hat{M}_{s,t}^{(l)}]_{s,t}$. The main objective of this paper is to derive method-of-types bounds for the empirical state–action pair measure and to transfer the resulting entropy-cost characterization to the empirical MDP kernel $\hat{\mathbb{M}}$.

While sample complexity upper bounds provide insight into the amount of data sufficient for learning, they fail to capture the manner in which errors diminish. In contrast, a method of types provides a sharper finite-alphabet perspective by characterizing the exponentially small probabilities of rare or atypical transitions. This complements worst-case sample-complexity theory by providing a sharper viewpoint that is useful in hypothesis testing, source coding [Natarajan, 1985], importance sampling [Deo and Murthy, 2025], and related machine-learning applications.

An immediate consequence in machine learning is the characterization of atypical behavior of offline RL estimators through their dependence on the empirical transition structure. In particular, common RL objects such as V- and Q-functions can be written as deterministic transformations of the transition kernel through the Bellman equations; see equations (5.1) and (5.2). Thus, once the empirical MDP kernel is related to the empirical pair type, the same entropy-cost framework yields variational formulas for plug-in value- and Q-function estimators.

Initially, our work was motivated by the goal of characterizing atypical behavior of value-, Q-, and advantage-functions in offline reinforcement learning. This led us to a more basic methodological question: can one obtain a method-of-types theory for empirical MDP kernels without imposing strict positivity of the transition kernel? Classical treatments of Markovian type probabilities and their asymptotic consequences often work under positivity assumptions [Vidyasagar, 2014, Dembo and Zeitouni, 2010, Natarajan, 1985]. These assumptions are restrictive in modern machine-learning settings, where sparsity and structural zeros are common [Agarwal et al., 2019, 2020]. Our work addresses this issue directly by developing type bounds that accommodate non-positive transition structure. We also note that subtleties arise in extending Markovian type arguments beyond the strictly positive setting. For example, Daures [2025] points out a technical oversight in the treatment of admissible empirical measures in Fayolle and de La Fortelle [2002], while also noting that this oversight does not affect the main proofs there. Our analysis avoids this issue by working directly with balanced empirical pair measures and an explicit bounded return-path augmentation argument. We now briefly summarize our main contributions

1. **Type bounds for empirical pair measures.** In Section 4, we establish a finite-sample upper bound and a matching subsequential lower exponent (Theorem 4.1) for the empirical pair measure $L_n^{(2)}$ of the state–action process. The result is supported on the balanced set and identifies the conditional Kullback–Leibler divergence relative to the true pair kernel as the relevant entropy cost. The result is exponentially sharp asymptotically, with an explicit finite-sample upper bound; finite-length lower estimates apply to individual realizable types. Large deviation principles for Markov chains [Donsker and Varadhan, 1975, 1976, 1983] are a downstream consequence (see Section 5 for more details).
2. **Applications to RL and source coding.** We apply the kernel-level entropy-cost characterization to plug-in offline RL. In particular, Proposition 5.1 give finite-sample variational entropy-cost formulas for discounted value- and Q-function estimates through the corresponding Bellman maps. We also apply the type bounds to universal source coding (Proposition 5.2).

The rest of the paper is organized as follows: Section 2 reviews relevant research works. In Section 3, we formally introduce the model and the relevant notations. In Section 4, we provide our key

theoretical results and the proof sketches for these, while the full proofs have been deferred till the appendix due to lack of space. Various theoretical applications for our results are presented in Section 5.

2 Background and Related Literature

Markov decision processes have found wide appeal in both theoretical [Jedra and Proutiere, 2020, Wang et al., 2020, Mania et al., 2020] and applied [Zhang et al., 2022, Kao et al., 2025, Zekri et al., 2024] aspects of reinforcement learning. We refer the reader to standard monographs [Hernández-Lerma et al., 1991, Puterman, 2014, Sutton and Barto, 2018] for comprehensive treatments. In parallel, the connection between MDP's and controlled Markov chains first appeared in a seminal PTRF paper by Borkar [1988] (see also Borkar and Varaiya [1979]) before appearing in the monograph Borkar [1991], and then being further expanded in Su et al. [2025], Banerjee et al. [2025a,b], Lee et al. [2025], among others.

We make note that Davisson, Longo, and Sgarro [1981] develop Markov-type arguments permitting structural zeros which permits a downstream LDP formulation. However, as Natarajan [1985, Remark 2] observes that Davisson's repair does not extend his Theorems 1 and 2 to general irreducible–aperiodic kernels, whereas our bounded return-path augmentation constructs nearby balanced realizable types with $O(1/n)$ perturbation.

Daures [2025] establishes weak level-2 and level-3 LDPs for possibly reducible Markov chains on discrete state spaces. Under a fixed Markov policy, $Z_i = (X_i, a_i)$ is itself a Markov chain, so these asymptotic results apply to the state–action process. Our contribution concerns finite-length type estimates and explicit finite-sample upper bounds, together with their transfer to empirical MDP kernels and Bellman functionals.

Our results also complement the minimax literature on finite Markov models. Agarwal, Kakade, and Yang [2020] study worst-case sample complexity for learning a near-optimal policy from generative-model samples, while Wolfer and Kontorovich [2019, 2021] study minimax transition estimation for ergodic Markov chains. We instead characterize the distribution-dependent decay of atypical empirical transitions and plug-in V- and Q-function estimates from a single stationary trajectory under a fixed policy. The resulting conditional-KL rate is sharp at the exponential scale under the stated lower-bound conditions, but we do not claim optimal polynomial prefactors or improved minimax rates. Our finite-state and finite-action assumptions are essential to the explicit type-counting bounds.

This paper's method-of-types framework for empirical MDP kernels is relevant to several areas of modern machine learning research [Kao et al., 2025, Zekri et al., 2024]. Related empirical transition and similarity structures also arise in clustering [Makarychev and Chakrabarty, 2023, Sanders et al., 2020], Markovian bootstraps [Banerjee and Chakrabarty, 2025], and time-series analysis [Douc et al., 2018, Mao et al., 2025]. In these settings, type-counting and entropy-cost arguments provide a way to quantify the probability of atypical empirical structures and to understand when the learned objects are reliable. We therefore move on to the main open question.

Open question The paper studies the following question in the theory of types for MDPs: *Can one establish a finite-sample upper bound and a matching subsequential lower exponent for the empirical state–action pair measure of an MDP under nothing but an ergodicity assumption on the pair process and, as downstream consequences, derive explicit exponential-rate characterizations for plug-in reinforcement-learning estimators and universal source-coding trajectories?*

Remark 2.1. Classical asymptotic LDPs for Markov chains also allow structural zeros [Dembo and Zeitouni, 2010, Daures, 2025]. Section E outlines the Gärtner–Ellis argument under strict positivity for simplicity. The asymptotic argument does not supply the finite-length type estimates or explicit finite-sample upper bound established here by path counting.

3 Problem Statement

Notations Let $\mathbb{N}$ and $\mathbb{R}$ denote the set of natural and real numbers, respectively, and let the symbol $\lfloor \cdot \rfloor$ denote the floor function. Throughout this paper, all random variables are defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where $\mathcal{F}$ is a σ-algebra and $\mathbb{F} := \{\mathcal{F}_i\}$ is a filtration with $\mathcal{F}_i \subset \mathcal{F}$.

Let $\{X_i\}$ represent a discrete-time stochastic process adapted to $\mathbb{F}$, taking values in a finite state space χ with $|\chi| = d$. A matrix $Q \in \mathbb{R}^{d\times d}$ is a stochastic matrix if each row of Q sums to 1; we denote the (s,t)-th entry of Q by $Q_{s,t}$. Let $\mathcal{A}$ represent a finite action space with $|\mathcal{A}| = k$, and $\{a_i\}_{i\geq 0}$ represent the (not necessarily Markovian) sequence of actions with $a_i \in \mathcal{A}$. $\widetilde{\chi}$ denotes $\chi \times \mathbb{I}$. Throughout the paper, we follow the convention that $Z_i := (X_i, a_i)$ denotes the pair process on $\widetilde{\chi} := \chi \times \mathbb{I}$, with true stationary distribution for the pair process $\{(Z_i, Z_{i+1})\}$ denoted by ξ.

O, Θ, Ω denotes the usual Bachman-Landau notations [Cormen et al., 2022] for order terms. $\bar{\mathcal{O}}$ shall denote the closure of a set $\mathcal{O}$ while $\mathcal{O}^\circ$ shall denote its interior. $\mathbb{1}$ is the indicator function whereas Id is the identity matrix whose dimension may change from line to line but shall always match that of other matrices in the equation. Let $\mathcal{M}_c(\widetilde{\chi}^2)$ be the set of all balanced measures $\boldsymbol{\nu} =: (\nu_{o,p})_{o,p\in\widetilde{\chi}}$ on $\widetilde{\chi}^2$, which means that $\sum_{o\in\widetilde{\chi}} \nu_{o,p} = \sum_{o\in\widetilde{\chi}} \nu_{p,o}$. For any measure $\boldsymbol{\nu}$, $H(\boldsymbol{\nu})$ shall denote its Shannon entropy, and $D(\boldsymbol{\nu}) := H(\boldsymbol{\nu}) - H(\bar{\boldsymbol{\nu}})$ where $\bar{\boldsymbol{\nu}} := (\bar{\nu}_o)_{o\in\widetilde{\chi}}$ and $\bar{\nu}_o := \sum_{p\in\widetilde{\chi}} \nu_{o,p}$. Observe that $D(\cdot)$ is the conditional-entropy functional

$$D(\boldsymbol{\nu}) \;:=\; -\sum_{o\in\widetilde{\chi}} \bar{\nu}_o \sum_{p\in\widetilde{\chi}} \frac{\nu_{o,p}}{\bar{\nu}_o} \log\Big(\frac{\nu_{o,p}}{\bar{\nu}_o}\Big).$$

Definition 3.1. A Markov decision process (MDP) is an $\mathbb{F}$-adapted pair process $\{(X_i, a_i)\}_{i\geq 0}$ such that

$$\mathbb{P}\left(X_{i+1} = s_{i+1} | \mathcal{F}_i\right) = M^{(l)}_{s_i,s_{i+1}} := \mathbb{P}(X_{i+1} = s_{i+1} | X_i = s_i, a_i = l).$$

Remark 3.2. A Markov decision process (MDP) is a CMC with a reward process $\{r_i\}$.

For each state $s \in \chi$, let M_s represent the matrix

$$\begin{bmatrix} M^{(1)}_{s,1}, \ldots, M^{(k)}_{s,1} \\ M^{(1)}_{s,2}, \ldots, M^{(k)}_{s,2} \\ \vdots \\ M^{(1)}_{s,d-1}, \ldots, M^{(k)}_{s,d-1} \\ M^{(1)}_{s,d}, \ldots, M^{(k)}_{s,d} \end{bmatrix}. \tag{3.1}$$

As d is finite, the collection of such matrices is finite.

The action sequence $\{a_i\}_{i\geq 0}$ is adapted to $\{\mathcal{F}_i\}$ for all $i \geq 0$. We have assumed that $\{a_i\}$ depends only on $\{X_i\}$, which makes the pair process $\{(X_i, a_i)\}$ form a time-homogeneous Markov chain with transition matrices obtained similarly.

Let $N^{(l)}_s$ be the number of visits to the (state, action) pair (s, l), and $N^{(l)}_{s,t}$ as the number of transitions from state s to state t under l. Formally, with $\mathbb{1}[\cdot]$ as indicator function,

$$N^{(l)}_s := \sum_{i=0}^{n-1} \mathbb{1}[X_i = s, a_i = l], \quad N^{(l)}_{s,t} := \sum_{i=0}^{n-1} \mathbb{1}[X_i = s, X_{i+1} = t, a_i = l], \quad \hat{M}^{(l)}_{s,t} := \frac{N^{(l)}_{s,t}}{N^{(l)}_s}$$

where $\hat{M}^{(l)}_{s,t}$ is our count-based estimator of $M^{(l)}_{s,t}$, the transition probability from state s to t conditioned on action l.

3.1 Goals

Our aim is to supply a complete exponential theory for the rare events of a controlled Markov chain under the weakest credible assumptions. Two concrete targets guide the paper:

1. **Establishing a method of types for controlled chains.** We aim to generalize the classical method of types to the setting of controlled state–action chains with non-positive transition kernels. This involves a combinatorial analysis of the empirical pair measure, with rate function characterized by the conditional Kullback–Leibler divergence derived via multinomial entropy.
2. **Characterizing fluctuations of learned dynamics.** Applying our method-of-types results to the normalization map, we precisely characterize the fluctuations of learned transition dynamics from offline data. This kernel-level entropy-cost characterization is what enables the downstream applications to plug-in V- and Q-function estimators and to source coding developed in Section 5.

4 Theoretical Results

Recall from Section 3 that our first objective is to produce a method of type bound for the empirical pair process. To that end, we recall the definition of $\mathcal{M}_c(\widetilde{\chi}^2)$, and $\bar{\nu}_o$ from the introduction. Next, we define the *pair empirical transition* matrix. Recall that for stationary MDPs, a_i depends only on X_i, and the pair process $Z_i := (X_i, a_i)$ forms a time-homogeneous Markov chain on $\widetilde{\chi}$. We denote by M_{pair} its transition matrix, defined by

$$M_{pair}(o,p) := \mathbb{P}(Z_{i+1} = p \mid Z_i = o), \qquad o,p \in \widetilde{\chi}.$$

We say that the pair process is *ergodic* if there exists $\kappa \in \mathbb{N}$ such that M_{pair}^{κ} has strictly positive entries. With these definitions in place, we can state our main type-counting result for the empirical pair measure $L_n^{(2)} := (\frac{1}{n}\sum_{i=1}^{n-1} \mathbb{1}[Z_i = o, Z_{i+1} = p])_{o,p\in\widetilde{\chi}}$.

Theorem 4.1 (Finite-sample upper bound and subsequential lower exponent)**.** *Let $\{Z_i\}_{i\geq 1}$ be the pair process on the finite space $\widetilde{\chi} := \chi \times \mathbb{I}$, and let*

$$L_n^{(2)} := \left(\frac{1}{n}\sum_{i=1}^{n-1} \mathbb{1}[Z_i = o, Z_{i+1} = p]\right)_{o,p\in\widetilde{\chi}}.$$

Assume that $\{Z_i\}$ is a time-homogeneous Markov chain with transition matrix M_{pair}, and that it is ergodic in the sense that there exists $\kappa \in \mathbb{N}$ such that $M_{pair}^{\kappa} > 0$. Trivially defining $\frac{0}{0} = 0$ and $0\log 0 = 0$

$$I(\boldsymbol{\nu}) := \begin{cases} \displaystyle\sum_{o\in\widetilde{\chi}} \bar{\boldsymbol{\nu}}_o \sum_{p\in\widetilde{\chi}} \frac{\boldsymbol{\nu}_{o,p}}{\bar{\boldsymbol{\nu}}_o} \log\left(\frac{\boldsymbol{\nu}_{o,p}/\bar{\boldsymbol{\nu}}_o}{M_{pair}(o,p)}\right), \end{cases} \tag{4.1}$$

for $\boldsymbol{\nu} \in \mathcal{M}_c(\widetilde{\chi}^2)$ and ∞ otherwise. With $C_{\widetilde{\chi}} = |\widetilde{\chi}|^2 + \lceil \log_2 |\widetilde{\chi}| \rceil$, for every Borel set $\mathcal{O} \subseteq \mathcal{M}_c(\widetilde{\chi}^2)$ and every $n \in \mathbb{N}$,

$$\mathbb{P}\Big(L_n^{(2)} \in \mathcal{O}\Big) \leq (n+1)^{C_{\widetilde{\chi}}} \exp\left(-n \inf_{\boldsymbol{\nu}\in\mathcal{O}} I(\boldsymbol{\nu})\right).$$

Moreover, for every Borel set $\mathcal{O} \subseteq \mathcal{M}_c(\widetilde{\chi}^2)$, there exists a subsequence $n_\ell \to \infty$ such that

$$\liminf_{\ell\to\infty} \frac{1}{n_\ell} \log \mathbb{P}\Big(L_{n_\ell}^{(2)} \in \mathcal{O}\Big) \geq -\inf_{\boldsymbol{\nu}\in\mathcal{O}^\circ} I(\boldsymbol{\nu}).$$

Remark 4.2. For the upper bound alone, it suffices that $\{Z_i\}$ is a time-homogeneous Markov chain on the finite space $\widetilde{\chi}$, with any initial distribution; neither irreducibility nor aperiodicity is required. Looking at the proof, we note that the finite-sample upper bound in Theorem 4.1 remains valid under the weaker assumptions above, while retaining the unified theorem statement for ease of presentation. In particular, κ is not required for the upper bound.

Proof Sketch: We provide a sketch of the proof while the full details are relegated to Section D.

Step 1: In general, $L_n^{(2)} \notin \mathcal{M}_c(\widetilde{\chi}^2)$, but, by ergodicity of the pair process, $\exists\, \kappa \in \mathbb{N}$ such that any state can be returned to in at most κ steps with positive probability. We append a shortest return path of length $k \leq \kappa$ to get an augmented trajectory of length $n+k$ (Lemma C.2 proves that this can be done in the same probability space) whose empirical measure $L_{n+k}^{(2)}$ lies in $\mathcal{M}_c(\widetilde{\chi}^2)$ and satisfies

$$\left\|L_{n+k}^{(2)} - L_n^{(2)}\right\|_1 \leq \frac{C}{n}, \tag{4.2}$$

for a constant C depending only on κ and $|\widetilde{\chi}|$. This allows us to work exclusively with consistent empirical measures.

Step 2: Fix $\boldsymbol{\nu} \in \mathcal{M}_c(\widetilde{\chi}^2)$ and let $\mathcal{E}(n,k,\boldsymbol{\nu})$ denote the set of augmented trajectories whose empirical measure until time n equals $\boldsymbol{\nu}$. This is our **type class**. By representing $\boldsymbol{\nu}$ as edge frequencies of a directed multigraph on $\widetilde{\chi}$, we use standard counting arguments on Eulerian multigraphs to get for $\tilde{n} = n(1 + \Theta(\frac{\kappa|\widetilde{\chi}|}{n}))$

$$\tfrac{1}{(2\tilde{n})^{|\widetilde{\chi}|^2}} \exp\big(\tilde{n} D(\boldsymbol{\nu})\big) \;\leq\; |\mathcal{E}(n,k,\boldsymbol{\nu})| \;\leq\; \tilde{n}\, \exp\big(\tilde{n} D(\boldsymbol{\nu})\big), \tag{4.3}$$

where $D(\boldsymbol{\nu}) = H(\boldsymbol{\nu}) - H(\bar{\boldsymbol{\nu}})$.

Step 3: For any trajectory $(z_1, \ldots, z_n) \in \mathcal{T}_n(\boldsymbol{\nu})$, the Markov property yields

$$\log \mathbb{P}\big(\text{trajectory}\big) = -n\Big(\mathsf{J}(\boldsymbol{\nu}, M_{pair}) - \mathsf{J}(\bar{\boldsymbol{\nu}}, \bar{M}_{pair})\Big) + O(1), \tag{4.4}$$

where $\mathsf{J}(\cdot,\cdot)$ denotes the cross-entropy functional. Combining (4.3) and (4.4), for each realizable balanced type $\boldsymbol{\nu}$

$$\frac{1}{n}\log \mathbb{P}\big(L_n^{(2)} = \boldsymbol{\nu}\big) = -\mathsf{D}(\boldsymbol{\nu}\|M_{pair}) + O(\tfrac{\log n}{n}). \tag{4.5}$$

Step 4: For the upper bound, write $m = |\widetilde{\chi}|$. Direct path counting gives a type-class upper bound $m\exp(nD(\boldsymbol{\nu}))$ and at most $(n+1)^{m^2}$ transition-count arrays, so

$$\mathbb{P}(L_n^{(2)} \in \mathcal{O}) \le m(n+1)^{m^2}\exp\left(-n\inf_{\boldsymbol{\nu}\in\mathcal{O}} I(\boldsymbol{\nu})\right).$$

This argument does not use the return-path augmentation. For the lower statement, the finite-length estimates for realizable augmented types and the approximation argument in Section D give a subsequence $n_\ell \to \infty$ for which

$$\liminf_{\ell\to\infty}\frac{1}{n_\ell}\log\mathbb{P}(L_{n_\ell}^{(2)} \in \mathcal{O}) \ge -\inf_{\boldsymbol{\nu}\in\mathcal{O}^\circ} I(\boldsymbol{\nu}).$$

Here $\mathcal{O}^\circ$ denotes the interior of $\mathcal{O}$; this is an asymptotic lower statement, rather than an all-n finite-sample bound. □

The following proposition concerns the obstruction to an all-n lower bound on balanced types when the pair process is periodic.

Proposition 4.3. *There exists an MDP $\{(X_i, a_i)\}_{i\ge1}$ in the sense of Definition 3.1, with finite χ and $\mathbb{I}$, and with a (deterministic) Markov policy $a_i = \pi(X_i)$, such that the induced time-homogeneous Markov chain $\{Z_i = (X_i, a_i)\}$ on $\widetilde{\chi} = \chi \times \mathbb{I}$ is irreducible but periodic. For this process, the empirical pair measures*

$$L_n^{(2)}(o,p) := \left(\frac{1}{n}\sum_{i=1}^{n-1}\mathbb{1}[Z_i = o, Z_{i+1} = p]\right)_{o,p\in\widetilde{\chi}}$$

do not admit an all-n lower bound on balanced types in $\mathcal{M}_c(\widetilde{\chi}^2)$.

Proposition 4.3 uses a recurrent (in the sense that each state-action pair appears infinitely often) Markov policy, so the failure is not due to non-Markovianity. The obstruction is purely *periodicity* of the induced pair chain $\{Z_i\}$, which forces subsequence-dependent behavior of admissible types and is precisely what the primitivity assumption $M_{\text{pair}}^{\kappa} > 0$ rules out. In Appendix G we note certain weaker forms of Theorem 4.1 for in-ergodic Markov chains containing communicating classes, although research in this field is in its infancy (see Appendix G for details).

4.1 Extensions to Uncertainty Quantification and Hypothesis Testing

We now extend the results of the previous section to derive finite-sample guarantees on uncertainty quantification for Markov chains. Let $\mathcal{U}_\beta$ be the set valued uncertainty map

$$\mathcal{U}_\beta(\boldsymbol{\nu}) := \{\theta \in \mathcal{M}_c(\widetilde{\chi}^2) : \mathsf{D}(\boldsymbol{\nu}\|\theta) \le \beta\}$$

where $\mathsf{D}(\boldsymbol{\nu}\|\theta)$ is the same rate function as Equation (4.1) with $M_{pair}(o,p)$ replaced by the candidate transition probability

$$\frac{\theta_{o,p}}{\bar{\theta}_o}, \qquad \bar{\theta}_o := \sum_{p\in\widetilde{\chi}}\theta_{o,p}.$$

Here $\beta > 0$ is the radius of uncertainty.

For any lower semi-continuous target functional $\psi : \mathcal{M}_c(\widetilde{\chi}^2) \to \mathbb{R}$, define the pessimistic estimator $\psi_\beta(\boldsymbol{\nu}) := \inf_{\theta\in\mathcal{U}_\beta(\boldsymbol{\nu})}\psi(\theta)$. Recall from Section 3 that ξ is true stationary distribution for the pair process $\{(Z_i, Z_{i+1})\}$. We now have the following proposition who proof is in Section B.1.

Proposition 4.4. *For every $\beta > 0$ and every $n \in \mathbb{N}$,*

$$\frac{1}{n} \log \mathbb{P}\left(\psi(\xi) < \psi_\beta\left(L_n^{(2)}\right)\right) \leq \frac{C_{\tilde{\chi}} \log(n+1)}{n} - \beta.$$

A consequence of the previous proposition is the following one-sided hypothesis test (we refer to Lehmann and Romano [2006] for details of hypothesis testing). Fix a threshold $c \in \mathbb{R}$ and consider

$$H_0 : \psi(\xi) \leq c \quad \text{vs} \quad H_1 : \psi(\xi) > c.$$

We reject H_0 whenever $\psi_\beta(L_n^{(2)}) > c$. Indeed, under H_0, the rejection event implies $\psi(\xi) < \psi_\beta(L_n^{(2)})$ and therefore

$$\mathbb{P}_\xi(\text{reject } H_0) \leq (n+1)^{C_{\tilde{\chi}}} e^{-n\beta}.$$

Thus the finite-sample significance level of the test is at most $\rho_n(\beta) = (n+1)^{C_{\tilde{\chi}}} e^{-n\beta}$. Equivalently, to obtain a prescribed significance level $\alpha_{\text{sig}} \in (0,1)$, it suffices to choose the radius of uncertainty

$$\beta \geq \frac{C_{\tilde{\chi}} \log(n+1) + \log(1/\alpha_{\text{sig}})}{n}.$$

For rare or unobserved transitions, conditional-KL uncertainty sets can be tighter than Hoeffding-type sets, while their local behavior is variance-adaptive as in Bernstein bounds.

5 Applications

5.1 Uncertainty Quantification for Reinforcement Learning Under Model Misspecification

Characterising the behavior of the V- and Q- function in offline MDP's is an important and a long-standing problem in reinforcement learning [Sutton and Barto, 2018]. We continue to work within the framework of Section 3 and briefly introduce concepts related to V- and Q- functions.

V-function: Let $\mathbb{M} := (M^{(1)}, M^{(2)}, \ldots, M^{(k)})$, with $\hat{\mathbb{M}}_n$ being the empirical counterpart, and for each state s let $\pi_s = (\pi(s,1), \ldots, \pi(s,k))$ denote the action distribution conditional on s. Define $\Pi = \text{diag}(\pi_1, \ldots, \pi_d)$, a $d \times dk$ block-diagonal matrix. Assume that g is a positive reward vector.

Let $\tilde{g} := (\tilde{g}(x,a) : (x,a) \in \chi \times \mathbb{I}) \in \mathbb{R}^{dk}$, where the component indexed by $(s-1)k + l$ corresponds to the cost incurred at state s under action l. The per-stage expected reward vector is $g := (g(x) : x \in \chi) \in \mathbb{R}^d$, with components given by $g(x) = \sum_{a \in \mathbb{I}} \pi(x,a)\tilde{g}(x,a)$.

For a fixed discount factor $0 < \alpha_{\text{dis}} < 1$, the value function $V_\Pi := (V_\Pi(x) : x \in \chi) \in \mathbb{R}^d$ is determined by the Bellman equation [Bertsekas, 2011], namely

$$V_\Pi = \Big(\text{Id} - \alpha_{\text{dis}}\, \Pi\mathbb{M}\Big)^{-1} g. \tag{5.1}$$

Replacing $\mathbb{M}$ with its empirical counterpart $\hat{\mathbb{M}}$ yields the plug-in estimate $\hat{V}_\Pi = \Big(\text{Id} - \alpha_{\text{dis}}\, \Pi\hat{\mathbb{M}}\Big)^{-1} g$.

Q-Function: Let $\tilde{r} := \big(\tilde{r}(x,a,y) : (x,a,y) \in \chi \times \mathbb{I} \times \chi\big)$ denote a $d^2k \times 1$ vector, where the component indexed by $(s-1)dk + (l-1)d + t$ corresponds to the cost associated with the tuple (s,l,t). The expected reward associated with each state–action pair is defined as $r := \big(r(x,a) : (x,a) \in \chi \times \mathbb{I}\big) \in \mathbb{R}^{dk}, r(x,a) = \sum_{y \in \chi} M_{x,y}^{(a)} \tilde{r}(x,a,y)$. Analogous to the value function, for a fixed discount factor $0 < \alpha_{\text{dis}} < 1$, the Q-function $Q_\Pi := \big(Q_\Pi(x,a) : (x,a) \in \chi \times \mathbb{I}\big) \in \mathbb{R}^{dk}$ is determined by the Bellman equation [Bertsekas, 2011],

$$Q_\Pi = \Big(\text{Id} - \alpha_{\text{dis}}\, \mathbb{M}\Pi\Big)^{-1} r. \tag{5.2}$$

Replacing $\mathbb{M}$ with its empirical counterpart $\hat{\mathbb{M}}$ yields the plug-in estimator $\hat{Q}_\Pi = (\text{Id} - \alpha_{\text{dis}}\, \hat{\mathbb{M}}\Pi)^{-1} r$.

We now derive a finite-sample uncertainty-quantification guarantee for the plug-in value functional introduced in the previous subsection. Define the mapping

$$\Phi(\boldsymbol{\nu})((x,a),x') = \begin{cases} \sum_{a' \in \mathbb{I}} \frac{\nu_{(x,a),(x',a')}}{\bar{\nu}_{(x,a)}}, & \text{if} \quad \bar{\nu}_{(x,a)} > 0 \\ 0 & \text{otherwise} \end{cases}$$

to the space of substochastic kernels. From the definition of the empirical measure, $\hat{\mathbb{M}}_n = \Phi(L_n^{(2)})$. Thus

$$V_\Pi = \big(\mathrm{Id} - \alpha_{\mathrm{dis}} \Pi \mathbb{M}\big)^{-1} g = \big(\mathrm{Id} - \alpha_{\mathrm{dis}} \Pi \Phi(\xi)\big)^{-1} g.$$

Fix a probability vector $\ell \in (\mathbb{R}^+)^d$. Define the scalar target functional

$$\psi_{\Pi,\ell}(\theta) := \ell^\top \big(\mathrm{Id} - \alpha_{\mathrm{dis}} \Pi \Phi(\theta)\big)^{-1} g, \qquad \theta \in \mathcal{M}_c(\widetilde{\chi}^2).$$

Equivalently, if we define

$$V_\Pi(\nu) := \big(\mathrm{Id} - \alpha_{\mathrm{dis}} \Pi \nu\big)^{-1} g,$$

then $\psi_{\Pi,\ell}(\theta) = \ell^\top V_\Pi(\Phi(\theta))$ represents the expected value under initial distribution ℓ.

For $\beta > 0$, the following estimator quantifies the uncertainty under misspecification. Recall the definition of the uncertainty set $\mathcal{U}_\beta(\boldsymbol{\nu})$ from Section 4.1 and define

$$V_{\Pi,\ell,\beta}(\boldsymbol{\nu}) := \inf_{\theta \in \mathcal{U}_\beta(\boldsymbol{\nu})} \psi_{\Pi,\ell}(\theta),$$

$Q_{\Pi,\ell,\beta}(\boldsymbol{\nu})$ is defined analogously. Then, we have the following proposition proved in Section B.3.

Proposition 5.1. *For every $\beta > 0$ and every $n \in \mathbb{N}$,*

$$\frac{1}{n} \log \mathbb{P}\left(\ell^\top V_\Pi < V_{\Pi,\ell,\beta}(L_n^{(2)})\right) \le \frac{C_{\widetilde{\chi}} \log(n+1)}{n} - \beta.$$

Furthermore, if $r \ge 0$ and $\eta \in (\mathbb{R}^+)^{dk}$, then

$$\frac{1}{n} \log \mathbb{P}\left(\eta^\top Q_\Pi < Q_{\Pi,\eta,\beta}(L_n^{(2)})\right) \le \frac{C_{\widetilde{\chi}} \log(n+1)}{n} - \beta.$$

5.2 Finite Sample Rates in Source Coding

Source coding is a fundamental problem in information theory and computer science dating back to the source coding theorem of Shannon [Moser and Chen, 2012, Chapter 5]. In this section we use the results of the previous section to derive the rate of decay of error probability in multi-source coding. To formalise the results of this section, we introduce some notation pertaining to source coding. The **source** is the Markov chain generating sequences (X_i, a_i), a **codebook** is a subset $A_n \subset \widetilde{\chi}^n$ of admissible source sequences (called **blocks**), and a **code** is the encoding scheme that assigns distinct **binary** codewords to elements of A_n and *declares an error when $Z^n \notin A_n$*.

Let $\bar{\xi}_o$ denote the stationary probability of $Z_i = o \in \widetilde{\chi}$, and define the stationary pair measure

$$\xi_{o,p} := \bar{\xi}_o \, M_{\mathrm{pair}}(o,p), \qquad o,p \in \widetilde{\chi}.$$

As before, we let

$$D(\nu) := H(\nu) - H(\bar{\nu})$$

denote the conditional-entropy functional on $\mathcal{M}_c(\widetilde{\chi}^2)$. For $R > 0$, define the universal codebook

$$A_n(R) := \left\{ z^n \in \widetilde{\chi}^n : D\Big(L_n^{(2)}(z^n)\Big) < R \right\}.$$

Then there exists a constant $C_{\mathrm{sc}} < \infty$, depending only on $|\widetilde{\chi}|$, such that

$$\frac{1}{n} \log |A_n(R)| \le R + \frac{C_{\mathrm{sc}} \log(n+1)}{n}.$$

Observe that the encoding rate is $\frac{1}{n} \log |A_n|$ and the decoding error probability for a given rate is

$$\mathbb{P}\{Z^n \notin A_n(R)\}.$$

Let I be the entropy-cost function appearing in Theorem 4.1. We now derive the finite sample version of the source coding theorem [Natarajan, 1985] as a downstream result.

Proposition 5.2 (Finite-sample universal source-coding bound)**.** *Assume the setting of Theorem 4.1. Then, the decoding error probability satisfies*

$$\mathbb{P}\{Z^n \notin A_n(R)\} \leq (n+1)^{C_{\tilde{\mathbb{X}}}} \exp\{-n\, E_+(R)\},$$

where

$$E_+(R) := \inf\left\{I(\nu) : \nu \in \mathcal{M}_c(\widetilde{\chi}^2), D(\nu) \geq R\right\}.$$

In particular, if $R > D(\xi)$, then $E_+(R) > 0$, and the decoding error probability decays exponentially up to a polynomial prefactor.

Similarly, if $R < D(\xi)$, then the success probability satisfies

$$\mathbb{P}\{Z^n \in A_n(R)\} \leq (n+1)^{C_{\tilde{\mathbb{X}}}} \exp\{-n\, E_-(R)\},$$

where

$$E_-(R) := \inf\left\{I(\nu) : \nu \in \mathcal{M}_c(\widetilde{\chi}^2), D(\nu) < R\right\}.$$

Informally, Proposition 5.2 depicts the error tradeoff for decoding in terms of the size of the codebook. When $R > D(\xi)$, the type codebook is large enough to contain the typical source trajectories, and the decoding error probability decays exponentially up to a polynomial prefactor. Conversely, when $R < D(\xi)$, the codebook is too small, and the probability of successful decoding is exponentially small up to a polynomial prefactor.

6 Conclusion

This paper develops a finite-sample method-of-types framework for empirical state–action pair measures in stationary finite MDPs. Under an ergodicity assumption, and without requiring strictly positive transition kernels, we obtain an explicit finite-sample upper bound and a matching subsequential lower exponent whose entropy cost is the conditional Kullback–Leibler divergence relative to the true pair kernel.

We then transfer this pair-measure characterization to empirical transition kernels through the normalization map, yielding large deviation principles for learned MDP dynamics. As applications, we derive variational rate functions for plug-in value- and Q-function estimators, finite-sample uncertainty-quantification and hypothesis-testing guarantees, and universal source-coding error exponents.

Limitations and Future Outlook. Although we addressed stationary MDPs in this paper, the question of bayesian [Ghavamzadeh et al., 2015, Vlassis et al., 2012] and contextual [Hallak et al., 2015] remain open. Of particular interest is the non-stationary setting [Bhattacharyya et al., 2026], although it remains challenging even for moderate deviations [Sarig, 2021, Dolgopyat and Hafouta, 2023]. An important direction for future work is to use the finite-sample KL sets developed here as ambiguity sets for offline reinforcement learning, optimizing policies against plausible transition kernels and propagating this uncertainty through robust V- and Q-functions for average reward MDP's.

Acknowledgments and Disclosure of Funding

Imon Banerjee acknowledges the Alumni Fellowship at Northwestern University, and the Operations Management Department at Booth School of Business for partially funding during the period when this research was conducted. Ramkrishna Jyoti Samanta acknowledges support from the Additional Funding Programme for Mathematical Sciences, delivered by EPSRC (EP/V521917/1) and the Heilbronn Institute for Mathematical Research.

References

A. Agarwal, N. Jiang, S. M. Kakade, and W. Sun. Reinforcement learning: Theory and algorithms. *CS Dept., UW Seattle, Seattle, WA, USA, Tech. Rep*, 32:96, 2019.

A. Agarwal, S. Kakade, and L. F. Yang. Model-Based Reinforcement Learning with a Generative Model is Minimax Optimal. In *Proceedings of Thirty Third Conference on Learning Theory*, pages 67–83. PMLR, July 2020. URL https://proceedings.mlr.press/v125/agarwal20b.html.

I. Banerjee and S. Chakrabarty. CLT and Edgeworth Expansion for m-out-of-n Bootstrap Estimators of The Studentized Median, May 2025. URL http://arxiv.org/abs/2505.11725. arXiv:2505.11725 [cs].

I. Banerjee, H. Honnappa, and V. Rao. Off-line Estimation of Controlled Markov Chains: Minimaxity and Sample Complexity. *Operations Research*, Feb. 2025a. ISSN 0030-364X. doi: 10.1287/opre.2023.0046. URL https://pubsonline.informs.org/doi/abs/10.1287/opre.2023.0046.

I. Banerjee, V. Rao, and H. Honnappa. Adaptive Estimation of the Transition Density of Controlled Markov Chains, May 2025b. URL http://arxiv.org/abs/2505.14458. arXiv:2505.14458 [math].

D. P. Bertsekas. Dynamic programming and optimal control 3rd edition, volume II. *Belmont, MA: Athena Scientific*, 2011.

R. Bhattacharyya, S. Chakrabarty, and I. Banerjee. Adaptive Estimation and Optimal Control in Offline Contextual MDPs without Stationarity, May 2026. URL http://arxiv.org/abs/2605.03393. arXiv:2605.03393 [stat].

V. Borkar and P. Varaiya. Adaptive control of markov chains, i: Finite parameter set. *IEEE Transactions on Automatic Control*, 24(6):953–957, 1979.

V. S. Borkar. A convex analytic approach to markov decision processes. *Probability Theory and Related Fields*, 78(4):583–602, 1988.

V. S. Borkar. *Topics in controlled Markov chains*. Longman Scientific & Technical, Harlow, UK, 1991.

T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms, fourth edition*. MIT Press, Apr. 2022. ISBN 978-0-262-36750-9.

I. Csiszar. The method of types [information theory]. *IEEE Transactions on Information Theory*, 44(6):2505–2523, Oct. 1998. ISSN 1557-9654. doi: 10.1109/18.720546. URL https://ieeexplore.ieee.org/document/720546.

L. Daures. Large deviations for general markov chains on discrete state spaces. *arXiv preprint arXiv:2507.11166*, 2025.

L. Davisson, G. Longo, and A. Sgarro. The error exponent for the noiseless encoding of finite ergodic Markov sources. *IEEE Transactions on Information Theory*, 27(4):431–438, 1981. doi: 10.1109/TIT.1981.1056377.

A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer, Berlin, Heidelberg, 2010. ISBN 978-3-642-03310-0 978-3-642-03311-7. doi: 10.1007/978-3-642-03311-7. URL http://link.springer.com/10.1007/978-3-642-03311-7.

A. Deo and K. Murthy. Achieving efficiency in black-box simulation of distribution tails with self-structuring importance samplers. *Operations Research*, 73(1):325–343, 2025.

D. Dolgopyat and Y. Hafouta. A Berry–Esseen theorem and Edgeworth expansions for uniformly elliptic inhomogeneous Markov chains. *Probability Theory and Related Fields*, 186(1):439–476, June 2023. ISSN 1432-2064. doi: 10.1007/s00440-022-01177-2. URL https://doi.org/10.1007/s00440-022-01177-2.

M. D. Donsker and S. R. S. Varadhan. Asymptotic evaluation of certain markov process expectations for large time, i. *Communications on Pure and Applied Mathematics*, 28(1):1–47, 1975.

M. D. Donsker and S. R. S. Varadhan. Asymptotic evaluation of certain markov process expectations, ii. *Communications on Pure and Applied Mathematics*, 28(2):279–301, 1976.

M. D. Donsker and S. R. S. Varadhan. Asymptotic evaluation of certain markov process expectations, iv. *Communications on Pure and Applied Mathematics*, 36(2):183–212, 1983.

R. Douc, E. Moulines, P. Priouret, and P. Soulier. *Markov chains*, volume 4. Springer, 2018.

G. Fayolle and A. de La Fortelle. Entropy and large deviations for discrete-time markov chains. *Problems of Information Transmission*, 38(4):354–367, 2002.

M. Ghavamzadeh, S. Mannor, J. Pineau, and A. Tamar. Bayesian Reinforcement Learning: A Survey. *Foundations and Trends in Machine Learning*, 8(5-6):359–483, Nov. 2015. ISSN 1935-8237. doi: 10.1561/2200000049. URL `https://doi.org/10.1561/2200000049`.

A. Hallak, D. D. Castro, and S. Mannor. Contextual Markov Decision Processes, Feb. 2015. URL `http://arxiv.org/abs/1502.02259`. arXiv:1502.02259 [stat].

O. Hernández-Lerma, R. Montes-de Oca, and R. Cavazos-Cadena. Recurrence conditions for Markov decision processes with Borel state space: a survey. *Annals of Operations Research*, 28(1):29–46, 1991.

R. A. Howard. *Dynamic programming and Markov processes*. Technology Press of Massachusetts Institute of Technology, 1960. Google-Books-ID: fXJEAAAAIAAJ.

R. A. Howard and J. E. Matheson. Risk-Sensitive Markov Decision Processes. *Management Science*, 18(7):356–369, 1972. ISSN 0025-1909. URL `https://www.jstor.org/stable/2629352`. Publisher: INFORMS.

Y. Jedra and A. Proutiere. Optimal best-arm identification in linear bandits. *Advances in Neural Information Processing Systems*, 33:10007–10017, 2020.

C.-C. Kao, C.-M. Yu, C.-S. Lu, and C.-S. Chen. Safety alignment depth in large language models: A markov chain perspective. *arXiv preprint arXiv:2502.00669*, 2025.

P. Lancaster. *Theory of Matrices*. Academic Press, New York, 1969.

J. Lee, Y. Jedra, A. Proutière, and S.-Y. Yun. Near-optimal clustering in mixture of markov chains. *arXiv preprint arXiv:2506.01324*, 2025.

E. L. Lehmann and J. P. Romano. *Testing statistical hypotheses*. Springer Science & Business Media, 2006.

G. Li, L. Shi, Y. Chen, Y. Chi, and Y. Wei. Settling the sample complexity of model-based offline reinforcement learning. *arXiv preprint arXiv:2204.05275*, 2022a.

Y. Li, R. Wang, and L. F. Yang. Settling the horizon-dependence of sample complexity in reinforcement learning. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 965–976. IEEE, 2022b.

K. Makarychev and S. Chakrabarty. Single-pass pivot algorithm for correlation clustering. keep it simple! *Advances in Neural Information Processing Systems*, 36:6412–6421, 2023.

H. Mania, M. I. Jordan, and B. Recht. Active learning for nonlinear system identification with guarantees. *arXiv preprint arXiv:2006.10277*, 2020.

Y. Mao, M. Hasan, M. M. Billah, Y. Li, S. Chakrabarty, C. S. Lee, K. Wang, M. N. T. Kilic, V. Gupta, W.-k. Liao, et al. An ai framework for time series microstructure prediction from processing parameters. *Scientific Reports*, 15(1):24074, 2025.

S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Springer Science & Business Media, 2012.

S. M. Moser and P.-N. Chen. *A student's guide to coding and information theory*. Cambridge University Press, 2012.

S. Natarajan. Large deviations, hypotheses testing, and source coding for finite markov chains. *IEEE Transactions on Information Theory*, 31(3):360–365, 1985.

M. L. Puterman. *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.

J. Sanders, A. Proutière, and S.-Y. Yun. Clustering in block markov chains. *The Annals of Statistics*, 48(6):3488–3512, 2020.

D. D. a. O. Sarig. Local limit theorems for inhomogeneous Markov chains, Sept. 2021. URL `http://arxiv.org/abs/2109.05560`. arXiv:2109.05560 [math].

Z. Su, I. Banerjee, and D. Klabjan. Central limit theorems for transition probabilities of controlled markov chains. *arXiv preprint arXiv:2508.01517*, 2025.

R. S. Sutton and A. G. Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

M. Vidyasagar. An elementary derivation of the large deviation rate function for finite state m arkov chains. *Asian Journal of Control*, 16(1):1–19, 2014.

N. Vlassis, M. Ghavamzadeh, S. Mannor, and P. Poupart. Bayesian Reinforcement Learning. In M. Wiering and M. van Otterlo, editors, *Reinforcement Learning: State-of-the-Art*, pages 359–386. Springer, Berlin, Heidelberg, 2012. ISBN 978-3-642-27645-3. doi: 10.1007/978-3-642-27645-3_11. URL `https://doi.org/10.1007/978-3-642-27645-3_11`.

P.-A. Wang, A. Proutiere, K. Ariu, Y. Jedra, and A. Russo. Optimal algorithms for multiplayer multi-armed bandits. In *International Conference on Artificial Intelligence and Statistics*, pages 4120–4129. PMLR, 2020.

G. Wolfer and A. Kontorovich. Minimax learning of ergodic Markov chains. In *Proceedings of the 30th International Conference on Algorithmic Learning Theory*, volume 98 of *Proceedings of Machine Learning Research*, pages 904–930. PMLR, 2019. URL `https://proceedings.mlr.press/v98/wolfer19a.html`.

G. Wolfer and A. Kontorovich. Statistical estimation of ergodic Markov chain kernel over discrete state space. *Bernoulli*, 27(1):532–553, 2021.

O. Zekri, A. Odonnat, A. Benechehab, L. Bleistein, N. Boullé, and I. Redko. Large language models as markov chains. *arXiv preprint arXiv:2410.02724*, 2024.

Y. Zhang, S. Chakrabarty, R. Liu, A. Pugliese, and V. Subrahmanian. A new dynamically changing attack on review fraud systems and a dynamically changing ensemble defense. in 2022 ieee intl conf on dependable, autonomic and secure computing, 2022.

A Further Applications

A.1 Downstream Type-bounds for Transition Measures

We now record the some consequences of the type bound from Theorem 4.1. Let $\tilde{\Gamma}$ be the set of stochastic kernels, and define the continuous mapping

$$\Phi : \mathcal{M}^+(\widetilde{\chi}^2) \to \Gamma, \qquad \Phi(\boldsymbol{\nu})((x,a),x') = \sum_{a' \in \mathbb{I}} \frac{\nu_{(x,a),(x',a')}}{\bar{\nu}_{(x,a)}},$$

From the definition of the empirical measure, $\hat{\mathbb{M}}_n = \Phi(L_n^{(2)})$. Then, we have the following proposition that is proved in Section B.2.

Proposition A.1 (Finite-sample kernel-level type bound)**.** *Assume the setting of Theorem 4.1. Let $\Phi : \mathcal{M}_c(\widetilde{\chi}^2) \to \widetilde{\Gamma}$ be the extended normalization map. Then, for every Borel set $\mathcal{B} \subseteq \widetilde{\Gamma}$, and defining*

$$J_{\ker}(K) := \inf \left\{ I(\nu) : \nu \in \mathcal{M}_c(\widetilde{\chi}^2),\ \Phi(\nu) = K \right\},$$

we have

$$\mathbb{P}\{\hat{\mathbb{M}}_n \in \mathcal{B}\} \le (n+1)^{C_{\tilde{\mathbb{X}}}} \exp\left\{ -n \inf_{K \in \mathcal{B}} J_{\ker}(K) \right\}.$$

B Proofs of Application

B.1 Proof of Proposition 4.4

Proof. Fix $\beta > 0$. Since the state space $\widetilde{\chi}$ is finite, $\mathcal{M}_c(\widetilde{\chi}^2)$ is compact. Moreover, it is easy to see that the map

$$(\boldsymbol{\nu}, \theta) \mapsto \mathsf{D}(\boldsymbol{\nu} \| \theta)$$

is lower semi-continuous on $\mathcal{M}_c(\widetilde{\chi}^2) \times \mathcal{M}_c(\widetilde{\chi}^2)$. Hence the graph of the set-valued map

$$\mathcal{U}_\beta(\boldsymbol{\nu}) = \left\{ \theta \in \mathcal{M}_c(\widetilde{\chi}^2) : \mathsf{D}(\boldsymbol{\nu} \| \theta) \le \beta \right\}$$

is closed. Since $\mathcal{M}_c(\widetilde{\chi}^2)$ is compact, $\mathcal{U}_\beta(\boldsymbol{\nu})$ is compact for every $\boldsymbol{\nu} \in \mathcal{M}_c(\widetilde{\chi}^2)$.

We first show that ψ_β is lower semi-continuous. To see this, let $\boldsymbol{\nu}_m \to \boldsymbol{\nu}$ and let $\theta_m \in \mathcal{U}_\beta(\boldsymbol{\nu}_m)$ be such that

$$\psi_\beta(\boldsymbol{\nu}_m) = \psi(\theta_m).$$

Such a minimizer exists because $\mathcal{U}_\beta(\boldsymbol{\nu}_m)$ is compact and ψ is lower semi-continuous. By compactness, along a subsequence, still denoted by θ_m, we may assume that $\theta_m \to \theta$. By lower semi-continuity of D,

$$\mathsf{D}(\boldsymbol{\nu} \| \theta) \le \liminf_{m \to \infty} \mathsf{D}(\boldsymbol{\nu}_m \| \theta_m) \le \beta.$$

Thus $\theta \in \mathcal{U}_\beta(\boldsymbol{\nu})$. Using the lower semi-continuity of ψ, we obtain

$$\psi_\beta(\boldsymbol{\nu}) \le \psi(\theta) \le \liminf_{m \to \infty} \psi(\theta_m) = \liminf_{m \to \infty} \psi_\beta(\boldsymbol{\nu}_m).$$

This proves the lower semi-continuity of ψ_β.

Now define the bad set

$$\mathcal{D}_\beta := \left\{ \boldsymbol{\nu} \in \mathcal{M}_c(\widetilde{\chi}^2) : \psi(\xi) < \psi_\beta(\boldsymbol{\nu}) \right\}.$$

Since ψ_β is lower semi-continuous, $\mathcal{D}_\beta$ is open, and in particular Borel. Moreover,

$$\left\{ \psi(\xi) < \psi_\beta\left(L_n^{(2)}\right) \right\} = \left\{ L_n^{(2)} \in \mathcal{D}_\beta \right\}.$$

We next lower bound the entropy cost of $\mathcal{D}_\beta$. Let $\boldsymbol{\nu} \in \mathcal{D}_\beta$. We claim that

$$\mathsf{D}(\boldsymbol{\nu} \| \xi) > \beta.$$

Indeed, suppose otherwise that $\mathsf{D}(\boldsymbol{\nu} \| \xi) \le \beta$. Then $\xi \in \mathcal{U}_\beta(\boldsymbol{\nu})$, and therefore

$$\psi_\beta(\boldsymbol{\nu}) = \inf_{\theta \in \mathcal{U}_\beta(\boldsymbol{\nu})} \psi(\theta) \le \psi(\xi),$$

which contradicts the fact that $\boldsymbol{\nu} \in \mathcal{D}_\beta$. Hence every $\boldsymbol{\nu} \in \mathcal{D}_\beta$ satisfies $\mathsf{D}(\boldsymbol{\nu}\|\xi) > \beta$, and so

$$\inf_{\boldsymbol{\nu}\in\mathcal{D}_\beta} \mathsf{D}(\boldsymbol{\nu}\|\xi) \geq \beta.$$

Applying the finite-sample method-of-types upper bound of Theorem 4.1 to the Borel set $\mathcal{D}_\beta$ gives

$$\begin{aligned}\mathbb{P}\left(\psi(\xi) < \psi_\beta\Big(L_n^{(2)}\Big)\right) &= \mathbb{P}\left(L_n^{(2)} \in \mathcal{D}_\beta\right) \\ &\leq (n+1)^{C_{\tilde{\chi}}} \exp\left(-n \inf_{\boldsymbol{\nu}\in\mathcal{D}_\beta} \mathsf{D}(\boldsymbol{\nu}\|\xi)\right) \\ &\leq (n+1)^{C_{\tilde{\chi}}} e^{-n\beta}.\end{aligned}$$

Taking logarithms and dividing by n yields

$$\frac{1}{n}\log\mathbb{P}\left(\psi(\xi) < \psi_\beta\Big(L_n^{(2)}\Big)\right) \leq \frac{C_{\widetilde{\chi}}\log(n+1)}{n} - \beta.$$

This completes the proof. □

B.2 Proof of Proposition A.1

Proof. By definition,

$$\{\hat{\mathbb{M}}_n \in \mathcal{B}\} = \{\Phi(L_n^{(2)}) \in \mathcal{B}\} = \{L_n^{(2)} \in \Phi^{-1}(\mathcal{B})\}.$$

Since Φ is Borel measurable, and Φ^{-1} is lower-semicontinuous, $\Phi^{-1}(\mathcal{B})$ is Borel in $\mathcal{M}_c(\widetilde{\chi}^2)$. Applying Theorem 4.1 to the Borel set $\Phi^{-1}(\mathcal{B})$ gives

$$\mathbb{P}\{\hat{\mathbb{M}}_n \in \mathcal{B}\} \leq (n+1)^{C_{\tilde{\chi}}} \exp\left\{-n \inf_{\nu\in\Phi^{-1}(\mathcal{B})} I(\nu)\right\}.$$

Since

$$\inf_{\nu\in\Phi^{-1}(\mathcal{B})} I(\nu) = \inf_{\nu\in\mathcal{M}_c(\widetilde{\chi}^2):\Phi(\nu)\in\mathcal{B}} I(\nu) = \inf_{K\in\mathcal{B}} J_{\ker}(K),$$

the result follows. □

B.3 Proof of Proposition 5.1

Proof. We only prove the first part concerning the expected value. The second part concerning the Q-function follows analogously.

Lower semicontinuity of $\psi_{\Pi,\ell}$: The key point is to verify that the target functional

$$\psi_{\Pi,\ell}(\theta) := \ell^\top\big(\mathrm{Id} - \alpha_{\mathrm{dis}}\Pi\Phi(\theta)\big)^{-1} g$$

is lower semicontinuous on $\mathcal{M}_c(\widetilde{\chi}^2)$.

By construction, $\Phi(\theta)$ is a substochastic kernel for every $\theta \in \mathcal{M}_c(\widetilde{\chi}^2)$. Hence $\Pi\Phi(\theta)$ is also row-substochastic. Therefore

$$\rho\big(\Pi\Phi(\theta)\big) \leq 1,$$

and since $0 < \alpha_{\mathrm{dis}} < 1$, the inverse

$$\big(\mathrm{Id} - \alpha_{\mathrm{dis}}\Pi\Phi(\theta)\big)^{-1}$$

is well-defined and admits the Neumann expansion

$$\big(\mathrm{Id} - \alpha_{\mathrm{dis}}\Pi\Phi(\theta)\big)^{-1} = \sum_{j=0}^{\infty} \alpha_{\mathrm{dis}}^j \big(\Pi\Phi(\theta)\big)^j.$$

Now, by the $0/0 = 0$ convention, each coordinate of Φ is lower semicontinuous and nonnegative on $\mathcal{M}_c(\widetilde{\chi}^2)$. Since Π has nonnegative entries, each coordinate of $\Pi\Phi(\theta)$ is also lower semicontinuous

and nonnegative. Consequently, for each fixed $j \geq 0$, each coordinate of $(\Pi\Phi(\theta))^j$ is a finite sum of products of nonnegative lower-semicontinuous functions, and is therefore lower semicontinuous.

Because $\ell \geq 0$ and $g \geq 0$, the finite partial sums

$$S_m(\theta) := \sum_{j=0}^{m} \alpha_{\mathrm{dis}}^j \ell^\top \big(\Pi\Phi(\theta)\big)^j g$$

are lower semicontinuous. Moreover,

$$\psi_{\Pi,\ell}(\theta) = \sup_{m \geq 0} S_m(\theta).$$

A pointwise supremum of lower-semicontinuous functions is lower semicontinuous. Hence $\psi_{\Pi,\ell}$ is lower semicontinuous on $\mathcal{M}_c(\widetilde{\chi}^2)$.

Application of Proposition 4.4: We now apply Proposition 4.4 with

$$\psi = \psi_{\Pi,\ell}, \qquad \psi_{\Pi,\ell}(\theta) = \ell^\top \big(\mathrm{Id} - \alpha_{\mathrm{dis}}\Pi\Phi(\theta)\big)^{-1} g.$$

Since $\Phi(\xi) = \mathbb{M}$, we have

$$\psi_{\Pi,\ell}(\xi) = \ell^\top \big(\mathrm{Id} - \alpha_{\mathrm{dis}}\Pi\mathbb{M}\big)^{-1} g = \ell^\top V_\Pi.$$

Moreover,

$$\psi_{\Pi,\ell,\beta}(L_n^{(2)}) = \inf_{\theta : \mathsf{D}(L_n^{(2)} \| \theta) \leq \beta} \psi_{\Pi,\ell}(\theta) = V_{\Pi,\ell,\beta}(L_n^{(2)}).$$

Therefore Proposition 4.4 gives

$$\frac{1}{n} \log \mathbb{P}_\xi \left(\ell^\top V_\Pi < V_{\Pi,\ell,\beta}(L_n^{(2)}) \right) \leq \frac{C_{\widetilde{\chi}} \log(n+1)}{n} - \beta.$$

This proves the claimed finite-sample bound for the value functional. □

B.4 Proof of Proposition 5.2

Proof. We first prove the codebook-size bound. Let

$$\mathcal{T}_n := \left\{ \nu \in \mathcal{M}_c(\widetilde{\chi}^2) : \nu = L_n^{(2)}(z^n) \text{ for some } z^n \in \widetilde{\chi}^n \right\}$$

denote the collection of admissible empirical pair types. For each $\nu \in \mathcal{T}_n$, let

$$\mathcal{C}_n(\nu) := \left\{ z^n \in \widetilde{\chi}^n : L_n^{(2)}(z^n) = \nu \right\}$$

be the corresponding type class. By the type-class counting bound from the method-of-types argument, there exists a constant $C_0 < \infty$, depending only on $|\widetilde{\chi}|$, such that for every admissible type ν,

$$|\mathcal{C}_n(\nu)| \leq (n+1)^{C_0} \exp\{nD(\nu)\}.$$

Moreover, the number of admissible empirical pair types is polynomial in n; that is, for some constant $C_1 < \infty$, depending only on $|\widetilde{\chi}|$,

$$|\mathcal{T}_n| \leq (n+1)^{C_1}.$$

Therefore,

$$\begin{aligned}
|A_n(R)| &= \left| \bigcup_{\nu \in \mathcal{T}_n :\, D(\nu) < R} \mathcal{C}_n(\nu) \right| \\
&\leq \sum_{\nu \in \mathcal{T}_n :\, D(\nu) < R} |\mathcal{C}_n(\nu)| \\
&\leq \sum_{\nu \in \mathcal{T}_n :\, D(\nu) < R} (n+1)^{C_0} \exp\{nD(\nu)\} \\
&\leq (n+1)^{C_0 + C_1} \exp(nR).
\end{aligned}$$

Setting $C_{\mathrm{sc}} := C_0 + C_1$, taking logarithms, and dividing by n gives

$$\frac{1}{n}\log|A_n(R)| \le R + \frac{C_{\mathrm{sc}}\log(n+1)}{n}.$$

We now prove the finite-sample decoding-error bound. By definition of the universal codebook,

$$A_n(R) = \left\{ z^n \in \widetilde{\chi}^n : D\Big(L_n^{(2)}(z^n)\Big) < R \right\}.$$

Hence the decoding error event is

$$\{Z^n \notin A_n(R)\} = \left\{ D\Big(L_n^{(2)}\Big) \ge R \right\}.$$

Define the Borel set

$$\mathcal{O}_R^+ := \left\{ \nu \in \mathcal{M}_c(\widetilde{\chi}^2) : D(\nu) \ge R \right\}.$$

Then

$$\{Z^n \notin A_n(R)\} = \left\{ L_n^{(2)} \in \mathcal{O}_R^+ \right\}.$$

Applying Theorem 4.1 to $\mathcal{O}_R^+$, we obtain

$$\begin{aligned}
\mathbb{P}\{Z^n \notin A_n(R)\} &= \mathbb{P}\left\{ L_n^{(2)} \in \mathcal{O}_R^+ \right\} \\
&\le (n+1)^{C_{\tilde{\mathbf{x}}}} \exp\left(-n \inf_{\nu \in \mathcal{O}_R^+} I(\nu) \right) \\
&= (n+1)^{C_{\tilde{\mathbf{x}}}} \exp\left\{ -nE_+(R) \right\},
\end{aligned}$$

where

$$E_+(R) := \inf\left\{ I(\nu) : \nu \in \mathcal{M}_c(\widetilde{\chi}^2),\ D(\nu) \ge R \right\}.$$

Next suppose $R > D(\xi)$. Since D is continuous on the finite-dimensional simplex and $\mathcal{O}_R^+$ is closed, $\xi \notin \mathcal{O}_R^+$. Moreover, under the ergodicity assumption in Theorem 4.1, the entropy-cost function I has the unique zero ξ. Therefore the infimum of I over the closed set $\mathcal{O}_R^+$ is strictly positive. Hence

$$E_+(R) > 0,$$

and the decoding error probability decays exponentially up to the polynomial prefactor $(n+1)^{C_{\tilde{\mathbf{x}}}}$.

We finally prove the finite-sample success-probability bound in the low-rate regime. If $R < D(\xi)$, then

$$\{Z^n \in A_n(R)\} = \left\{ D\Big(L_n^{(2)}\Big) < R \right\}.$$

Define

$$\mathcal{O}_R^- := \left\{ \nu \in \mathcal{M}_c(\widetilde{\chi}^2) : D(\nu) < R \right\}.$$

Then

$$\{Z^n \in A_n(R)\} = \left\{ L_n^{(2)} \in \mathcal{O}_R^- \right\}.$$

Applying Theorem 4.1 again gives

$$\begin{aligned}
\mathbb{P}\{Z^n \in A_n(R)\} &= \mathbb{P}\left\{ L_n^{(2)} \in \mathcal{O}_R^- \right\} \\
&\le (n+1)^{C_{\tilde{\mathbf{x}}}} \exp\left(-n \inf_{\nu \in \mathcal{O}_R^-} I(\nu) \right) \\
&= (n+1)^{C_{\tilde{\mathbf{x}}}} \exp\left\{ -nE_-(R) \right\},
\end{aligned}$$

where

$$E_-(R) := \inf\left\{ I(\nu) : \nu \in \mathcal{M}_c(\widetilde{\chi}^2),\ D(\nu) < R \right\}.$$

Again, because $R < D(\xi)$, the set

$$\{\nu \in \mathcal{M}_c(\widetilde{\chi}^2) : D(\nu) \le R\}$$

does not contain ξ. Since I has the unique zero ξ, the exponent is strictly positive. Therefore the probability of successful decoding is exponentially small up to the same polynomial prefactor. $\square$

C Technical Desiderata

Definition C.1 (Irreducible matrix)**.** A nonnegative square matrix $A \in \mathbb{R}^{n\times n}$ is called *irreducible* if for every pair of indices $i, j \in \{1, \dots, n\}$ there exists an integer $\kappa \geq 1$ such that

$$(A^{\kappa})_{ij} > 0.$$

Equivalently, if one associates to A the directed graph with vertices $\{1, \dots, n\}$ and an edge $i \to j$ whenever $A_{ij} > 0$, then A is irreducible precisely when this digraph is strongly connected.

Let $(X_t, a_t)_{t\geq 0}$ be a finite-state irreducible Markov chain with state space $\widetilde{\chi} = \chi \times \mathbb{I}$ and define the pair process

$$Z_i^{(2)} := \big((X_i, a_i), (X_{i+1}, a_{i+1})\big), \qquad i \geq 1,$$

with finite state space $\widetilde{\chi}^2$. Observe that $Z_i^{(2)}$ is Markovian with transition kernel

$$M_{pair}^{(2)}\big((y,b)\times(y',b'),(x,a)\times(x',a')\big) = \mathbb{I}_{(y',b')}((x,a))\,\mathbb{P}\left(Z_{i+1} = (x',a') \mid Z_i = (x,a)\right).$$

where $\mathbb{P}\left(Z_{i+1} = (x',a') \mid Z_i = (x,a)\right) := \mathbb{P}(X_{i+1} = x' \mid X_i = x, a_i = a)\,\mathbb{P}(a_{i+1} = a' \mid X_{i+1} = x)$. For simplicity, we have assumed that $M_{pair}(i,j) > 0,\ \forall\, i,j \in \widetilde{\chi}$, which easily implies that the paired process $Z_t^{(2)}$ is irreducible. The empirical transition measure of $Z_i^{(2)}$ is

$$L_n^{(2)}(o,p) \;=\; \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}[Z_i = o, Z_{i+1} = p], \qquad s_1, s_2 \in \widetilde{\chi}.$$

Let $M_{pair}^{(2)}$ denote the transition matrix of the chain $(Z_i^{(2)})$ and, for $\lambda = \Big(\lambda(y,b,y',b')\Big)_{(y,b,y',b')\in\widetilde{\chi}^2} \in \mathbb{R}^{d\times k\times d\times k} (= \mathbb{R}^{|\widetilde{\chi}|^2})$, define the tilted matrix

$$M_{pair,\lambda}^{(2)}(s,t) := M_{pair}^{(2)}(s,t)\, e^{\lambda_t}, \qquad s,t \in \widetilde{\chi}^2.$$

Lemma C.2 (Return to initial state via augmentation)**.** *Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a probability space and let $(W_t)_{t\geq 1}$ be an irreducible Markov chain on a finite state space $\mathcal{S}$ with transition matrix M.*

Fix $n \in \mathbb{N}$ and states $w_1, \dots, w_n \in \mathcal{S}$, and define the path event

$$E_n \;:=\; \{W_1 = w_1,\ W_2 = w_2,\ \dots,\ W_n = w_n\}.$$

If $\mathbb{P}(E_n) > 0$, then there exist an integer $\ell \geq 1$ and states $v_1, \dots, v_{\ell-1} \in \mathcal{S}$ such that the augmented event

$$E_{n+\ell} \;:=\; \{W_1 = w_1, \dots, W_n = w_n,\ W_{n+1} = v_1, \dots, W_{n+\ell-1} = v_{\ell-1},\ W_{n+\ell} = w_1\}$$

satisfies $E_{n+\ell} \in \mathcal{F}$ and $\mathbb{P}(E_{n+\ell}) > 0$. In particular, one can extend the finite path segment $(w_1, \dots, w_n)$ by finitely many additional states so that the final state equals the initial state.

Proof. For each $t \geq 1$ and $s \in \mathcal{S}$, the set $\{W_t = s\} = W_t^{-1}(\{s\})$ is in $\mathcal{F}$ because W_t is measurable. Hence E_n and $E_{n+\ell}$ are finite intersections of measurable sets, so $E_n, E_{n+\ell} \in \mathcal{F}$.

Since the chain is irreducible, for any $i, j \in \mathcal{S}$ there exists $k \geq 1$ such that $(M^{\ell})_{i,j} > 0$. Applying this with $i = w_n$ and $j = w_1$, we obtain some $\ell \geq 1$ such that

$$(M^{\ell})_{w_n,w_1} > 0.$$

By expanding the matrix power P^{ℓ}, there exist states $v_1, \dots, v_{\ell-1} \in \mathcal{S}$ such that

$$M_{w_n,v_1}\, M_{v_1,v_2} \cdots M_{v_{\ell-1},w_1} \;>\; 0.$$

Fix such ℓ and $(v_1, \dots, v_{\ell-1})$ and define $E_{n+\ell}$ accordingly. Assuming $\mathbb{P}(E_n) > 0$, we compute

$$\mathbb{P}(E_{n+\ell}) = \mathbb{P}(E_n)\,\mathbb{P}\left(W_{n+1} = v_1, \dots, W_{n+\ell} = w_1 \,|\, E_n\right).$$

On E_n we have $W_n = w_n$, and by Markov property,

$$\begin{aligned}\mathbb{P}\left(W_{n+1} = v_1, \dots, W_{n+\ell} = w_1 \,|\, E_n\right) &= \mathbb{P}\left(W_{n+1} = v_1, \dots, W_{n+\ell} = w_1 \,|\, W_n = w_n\right)\\ &= M_{w_n,v_1}\, M_{v_1,v_2} \cdots M_{v_{\ell-1},w_1}.\end{aligned}$$

Therefore,

$$\mathbb{P}(E_{n+\ell}) = \mathbb{P}(E_n)\, M_{w_n,v_1}\, M_{v_1,v_2} \cdots M_{v_{\ell-1},w_1} \;>\; 0,$$

since $\mathbb{P}(E_n) > 0$ and the product of transition probabilities is strictly positive by construction. □

D Proof of Theorem 4.1

Step I: We first introduce the notion of balanced measures, which form an important class of doublet measures in this paper. Let $\boldsymbol{\nu} : \widetilde{\chi}^{n-1} \to [0,1]^{|\widetilde{\chi}|^2}$ defined as $(\nu(Z_1, \ldots, Z_n))_{o,p} = \frac{1}{n}\sum_{i=1}^{n-1} \mathbb{1}[Z_i = o, Z_{i+1} = p]$ for all $o, p \in \widetilde{\chi}$. Let $\mathcal{A}_n = \{\boldsymbol{\nu} : (z_1, \ldots, z_{n-1}) \in \widetilde{\chi}^{n-1}\}$ and let $\mathcal{M}_c(\widetilde{\chi}) := \{\boldsymbol{\mu} : \sum_{o,p} \mu_{o,p} = 1 \text{ for } o, p \in \widetilde{\chi}\}$ be the simplex of probability measures on $\widetilde{\chi}$ which satisfies the stationary balance conditions

$$\sum_{o \in \widetilde{\chi}} \mu_{o,p} = \sum_{o \in \widetilde{\chi}} \mu_{p,o}. \qquad \text{(Balance condition)}$$

We call such measures *balanced*, and observe that any empirical doublet distribution on $Z_1, \ldots, Z_n$ is balanced *if and only if* $Z_1 = Z_n$. Further observe that, the stationary doublet distribution is balanced for any Markov chain, and therefore $\lim_{n\to\infty} \mathbb{P}(L_n^{(2)} = \boldsymbol{\nu}) = 0$ whenever $\boldsymbol{\nu}$ is unbalanced.

Next, for any subset $\mathcal{O} \subseteq \mathcal{M}_c(\widetilde{\chi})$

$$\mathbb{P}\left(L_n^{(2)} \in \mathcal{O}\right) = \sum_{\boldsymbol{\nu} \in \mathcal{A}_n \cap \mathcal{O}} \mathbb{P}\left(L_n^{(2)} = \boldsymbol{\nu}\right)$$

Let $z_n, \ldots, z_{n+k-1}$ be the shortest path such that $z_{n+k-1} = z_1$, and that

$$\mathbb{P}\left(Z_{n-1} = z_{n-1}, Z_n = z_n, \ldots, Z_{n+k-1} = z_{n+k-1}\right) > 0.$$

Lemma C.2 guarantees that this can be done in the same probability space. This path has some useful properties. Let $n_{o,p} := \sum_{i=1}^{n+k-1} \mathbb{1}[z_i = o, z_{i+1} = p]$, $o, p \in \widetilde{\chi}$, and $\bar{n}_o = \sum_p n_{o,p}$, then:

- Either $n\bar{\nu}_o = \bar{n}_o$ or $\bar{n}_o = n\bar{\nu}_o + 1$. Otherwise we get a cycle which contracts the shortest requirement.
- If $n\bar{\nu}_o = \bar{n}_o$, then $n\nu_{o,p} = n_{o,p}$ and otherwise, for exactly one p, $n_{o,p} = n\nu_{o,p} + 1$.
- The length of the path is at most κ, where κ is such that $M_{pair}^{\kappa} > 0$. Such a κ exists by irreducibility of the pair process.

all three of which are upper bounded by 1 (which follows from the fact that $\nu_{o,p} \le \nu_o$).

$$\frac{n_{o,p}}{\bar{n}_o} = \begin{cases} \dfrac{\nu_{o,p}}{\bar{\nu}_o}, & \text{if } \bar{n}_o = n\bar{\nu}_o \text{ (hence } n_{o,p} = n\nu_{o,p} \text{ for all } p), \\ \dfrac{\nu_{o,p} + \frac{1}{n}}{\bar{\nu}_o + \frac{1}{n}}, & \text{if } \bar{n}_o = n\nu_o + 1 \text{ and } p \text{ is such that } n_{o,p} = n\nu_{o,p} + 1, \\ \dfrac{\nu_{o,p}}{\bar{\nu}_o + \frac{1}{n}}, & \text{otherwise.} \end{cases} \qquad \text{(D.1)}$$

Next, let $\mathcal{E}(n, k, \boldsymbol{\nu})$ be defined as follows

$$\mathcal{E}(n,k,\boldsymbol{\nu}) := \left\{ L_{n+k}^{(2)} : Z_i = z_i, \forall\, i \in \{1, \ldots, n-1\}, L_n^{(2)} = \boldsymbol{\nu}, \boldsymbol{\nu} \in \mathcal{A}_n, \text{ and } L_{n+k}^{(2)} \text{ is balanced} \right\}.$$

In words, $\mathcal{E}(n, k, \boldsymbol{\nu})$ is the set of all measures created by returning to z_1 in k step starting from z_n without repeating states. Observe by this can be done within at most κ steps and that $\mathcal{A}_n = \bigcup_{k=1}^{\kappa} \mathcal{E}(n, k, \boldsymbol{\nu})$ and $\mathcal{E}(n, i, \nu) \cap \mathcal{E}(n, j, \nu) = \varnothing$ for all i, j pairs. We therefore get,

$$\begin{aligned} \mathbb{P}\left(L_n^{(2)} \in \mathcal{O}\right) &= \sum_{\boldsymbol{\nu} \in \mathcal{A}_n \cap \mathcal{O}} \mathbb{P}\left(L_n^{(2)} = \boldsymbol{\nu}\right) \\ &= \sum_{\substack{k \in \{1,\ldots,\kappa\} \\ \boldsymbol{\nu}^{(k)} \in \mathcal{E}(n,k,\boldsymbol{\nu}) \cap \mathcal{O}}} \mathbb{P}\left(L_{n+k}^{(2)} = \boldsymbol{\nu}^{(k)}\right) \\ &\le \left(\sum_{k=1}^{\kappa} |\mathcal{E}(n,k,\boldsymbol{\nu})|\right) \left(\sup_{\boldsymbol{\nu}^{(k)} \in \mathcal{E}(n,k,\boldsymbol{\nu})} \mathbb{P}\left(L_{n+k}^{(2)} = \boldsymbol{\nu}^{(k)}\right)\right) \\ \implies \log \mathbb{P}\left(L_{n+k}^{(2)} \in \mathcal{O}\right) &\le \log\left(\sum_{k=1}^{\kappa} |\mathcal{E}(n,k,\boldsymbol{\nu})|\right) + \log \sup_{\boldsymbol{\nu}^{(k)} \in \mathcal{E}(n,k,\boldsymbol{\nu})} \mathbb{P}\left(L_{n+k}^{(2)} = \boldsymbol{\nu}^{(k)}\right). \end{aligned}$$

Now, since $\log$ is a continuous and monotone function,

$$\log \sup_{\boldsymbol{\nu}^{(k)} \in \mathcal{E}(n,k,\boldsymbol{\nu})} \mathbb{P}\left(L_{n+k}^{(2)} = \boldsymbol{\nu}^{(k)}\right) = \sup_{\boldsymbol{\nu}^{(k)} \in \mathcal{E}(n,k,\boldsymbol{\nu})} \log \mathbb{P}\left(L_{n+k}^{(2)} = \boldsymbol{\nu}^{(k)}\right).$$

Dividing by n

$$\tfrac{1}{n} \log \mathbb{P}\left(L_n^{(2)} \in \mathcal{O}\right) \leq \tfrac{1}{n} \log \left(\sum_{k=1}^{\kappa} |\mathcal{E}(n,k,\boldsymbol{\nu})|\right) + \sup_{\boldsymbol{\nu}^{(k)} \in \mathcal{E}(n,k,\boldsymbol{\nu})} \tfrac{1}{n} \log \mathbb{P}\left(L_{n+k}^{(2)} = \boldsymbol{\nu}^{(k)}\right).$$

We now require bounds on $(\sum_{k=1}^{\kappa} |\mathcal{E}(n,k,\boldsymbol{\nu})|)$ which is what we derive next. Observe that,

$$\left(\sum_{k=1}^{\kappa} |\mathcal{E}(n,k,\boldsymbol{\nu})|\right) \leq (2^{\kappa} - 1)|\mathcal{A}_n|$$

so it is sufficient to bound $|\mathcal{A}_n|$. To this end, we need to count an upper bound to the number of possible distributions from a sample of length n, which is simply $\binom{n+|\widetilde{\chi}|-1}{\widetilde{\chi}-1}$. At this point, we make note that the previous bound does not make use of the stationary balance conditions. A tighter bound $\binom{n+|\widetilde{\chi}|^2-|\widetilde{\chi}|}{\widetilde{\chi}^2-|\widetilde{\chi}|}$ can be derived when the balance conditions are imposed[1], but the main point is that this is polynomial in n. Thus,

$$\tfrac{1}{n} \log \mathbb{P}\left(L_n^{(2)} \in \mathcal{O}\right) \leq O\left(\frac{\log n}{n}\right) + \sup_{\boldsymbol{\nu}^{(k)} \in \mathcal{E}(n,k,\boldsymbol{\nu})} \tfrac{1}{n} \log \mathbb{P}\left(L_{n+k}^{(2)} = \boldsymbol{\nu}^{(k)}\right). \tag{D.2}$$

Step II: We move on to the next step. Observe that for each $\boldsymbol{\nu} \in \mathcal{O}$ and for each sample path $z_1, \ldots, z_n$ such that for that sample path $L_n^{(2)} = \boldsymbol{\nu}$, we get

$$\mathbb{P}\left(L_n^{(2)} = \boldsymbol{\nu}\right) = \sum_{\{z_1,\ldots,z_{n-1}: L_n^{(2)} = \boldsymbol{\nu}\}} \mathbb{P}\left((Z_1, \ldots, Z_{n-1}) = (z_1, \ldots, z_{n-1})\right). \tag{D.3}$$

Now,

$$\begin{aligned}
&\mathbb{P}\left((Z_1, \ldots, Z_{n-1}) = (z_1, \ldots, z_{n-1})\right) \\
&\quad = \mathbb{P}(Z_1 = z_1) \prod_{i=2}^{n-1} \mathbb{P}(Z_i = z_i | Z_{i-1} = z_{i-1}) \\
&\quad = \frac{\mathbb{P}(Z_1 = z_1)}{\prod_{i=n}^{n+k-1} \mathbb{P}\left(Z_i = z_i | Z_{i-1} = z_{i-1}\right)} \prod_{i=2}^{n+k-1} \mathbb{P}(Z_i = z_i | Z_{i-1} = z_{i-1}) \\
&\quad = \frac{\mathbb{P}(Z_1 = z_1)}{\prod_{i=1}^{k} \mathbb{P}\left(Z_i = z_i | Z_{i-1} = z_{i-1}\right)} \prod_{i=2}^{n+k-1} \mathbb{P}(Z_i = z_i | Z_{i-1} = z_{i-1})
\end{aligned} \tag{D.4}$$

Let $\bar{\xi}_o$ be the stationary probability of being in state o and let $\xi_{o,p} := \mathbb{P}(Z_2 = p \mid Z_1 = o)\bar{\xi}_o$. Then, for some $\boldsymbol{\nu}^{(k)} \in \mathcal{E}(n,k,\boldsymbol{\nu})$

$$\begin{aligned}
&\therefore \log \left[\prod_{i=2}^{n+k-1} \mathbb{P}(Z_i = z_i | Z_{i-1} = z_{i-1})\right] \\
&\quad \leq (n+k-1) \sum_{o,p} \nu_{o,p}^{(k)} \log \xi_{o,p} - (k-1)|\widetilde{\chi}|^2 \log(\max_{o,p} \xi_{o,p}) - (n+k-1) \sum_{o} \bar{\nu}_o^{(k)} \log \bar{\xi}_o,
\end{aligned}$$

and

$$\begin{aligned}
&\log \left[\prod_{i=2}^{n+k-1} \mathbb{P}(Z_i = z_i | Z_{i-1} = z_{i-1})\right] \\
&\quad \geq (n+k-1) \sum_{o,p} \nu_{o,p}^{(k)} \log \xi_{o,p} + (k-1)|\widetilde{\chi}| \log(\min_{o} \bar{\xi}_o) - (n+k-1) \sum_{o} \bar{\nu}_o^{(k)} \log \bar{\xi}_o.
\end{aligned}$$

[1] A looser bound of $(n+1)^{|\widetilde{\chi}|}$ also suffices.

Observe that since the chain is ergodic, $\min_o \xi_o > 0$ and $\max_{o,p} \xi_{o,p} < 1$. Combining, we now have

$$(n+k-1)\sum_{o,p}\nu^{(k)}_{o,p}\log\xi_{o,p} - (k-1)|\widetilde{\chi}|\log\bar{\xi}_o - (n+k-1)\sum_o \bar{\nu}^{(k)}_o\log\bar{\xi}_o \leq$$
$$\log\left[\prod_{i=2}^{n+k-1}\mathbb{P}(Z_i = z_i|Z_{i-1}=z_{i-1})\right] \leq$$
$$(n+k-1)\sum_{o,p}\nu^{(k)}_{o,p}\log\xi_{o,p} + (k-1)|\widetilde{\chi}|^2\log\xi_{o,p} - (n+k-1)\sum_o \bar{\nu}^{(k)}_o\log\bar{\xi}_o$$

and since $|\widetilde{\chi}| \leq |\widetilde{\chi}|^2$ we can further refine the result into

$$(n+k-1)\sum_{o,p}\nu^{(k)}_{o,p}\log\xi_{o,p} - \Omega((k-1)|\widetilde{\chi}|^2) - (n+k-1)\sum_o \bar{\nu}^{(k)}_o\log\bar{\xi}_o \leq$$
$$\log\left[\prod_{i=2}^{n+k-1}\mathbb{P}(Z_i = z_i|Z_{i-1}=z_{i-1})\right]$$
$$\leq (n+k-1)\sum_{o,p}\nu^{(k)}_{o,p}\log\xi_{o,p} + O((k-1)|\widetilde{\chi}|^2)\log\xi_{o,p} - (n+k-1)\sum_o \bar{\nu}^{(k)}_o\log\bar{\xi}_o.$$

Finally, since $\mathbb{P}\left(Z_{n-1}=z_{n-1}, Z_n = z_n, \ldots, Z_{n+k-1}=z_{n+k-1}\right) > 0$, $\frac{\mathbb{P}(Z_1=z_1)}{\prod_{i=1}^k \mathbb{P}(Z_i=z_i|Z_{i-1}=z_{i-1})}$ remains bounded above and away from 0.

Returning to Equation (D.3), we now need to bound from above and below $\sum_{\{z_1,\ldots,z_{n+k-1}:L^{(2)}_{n+k-1}=\boldsymbol{\nu}^{(k)}\}} 1$. Define $D(\boldsymbol{\nu}) := H(\boldsymbol{\nu}) - H(\bar{\boldsymbol{\nu}})$ where $\bar{\boldsymbol{\nu}} = (\boldsymbol{\nu}_i)_{i\in\widetilde{\chi}}$

Lemma D.1. *Let $\boldsymbol{\nu}$ be such that $L^{(2)}_{n+k}(z_1,\ldots,z_n) = \boldsymbol{\nu}^{(k)}$. Then,* $(2(n+k-1))^{-|\widetilde{\chi}|^2}\exp((n+k-1)D(\boldsymbol{\nu}^{(k)})) \leq |\{z_1,\ldots,z_{n+k-1} : L^{(2)}_{n+k} = \boldsymbol{\nu}^{(k)}\}| \leq (n+k-1)\exp((n+k-1)D(\boldsymbol{\nu}^{(k)}))$

This lemma is proved in Section D.1. We define the cross entropy functional as,

Definition D.2 (Cross-entropy functional)**.** Let $\mathcal{X}$ be a finite set and let $\boldsymbol{\nu}, \boldsymbol{\mu} \in \mathcal{M}_c(\mathcal{X})$. Define

$$\mathsf{J}(\boldsymbol{\nu},\boldsymbol{\mu}) := -\sum_{x\in\mathcal{X}}\nu_x\log\mu_x.$$

Then, With $\boldsymbol{\nu}^{(k)} \in \mathcal{M}_c(\widetilde{\chi}^2)$ and $\xi \in \mathcal{M}_c(\widetilde{\chi}^2)$, and their marginals $\bar{\nu}^{(k)}$ and $\bar{\xi}$, we have

$$(n+k-1)\sum_{o,p\in\widetilde{\chi}}\nu^{(k)}_{o,p}\log\xi_{o,p} \;-\; (n+k-1)\sum_{o\in\widetilde{\chi}}\bar{\nu}^{(k)}_o\log\bar{\xi}_o$$
$$= -(n+k-1)\Big(\mathsf{J}(\nu^{(k)},\xi) - \mathsf{J}(\bar{\nu}^{(k)},\bar{\xi})\Big).$$

Then,

$$\log\left(\mathbb{P}\left((Z_1,\ldots,Z_{n-1}) = (z_1,\ldots,z_{n-1})\right)\right)$$
$$= -(n+k-1)\Big(\mathsf{J}(\nu^{(k)},\xi) - \mathsf{J}(\bar{\nu}^{(k)},\bar{\xi})\Big) \pm \Theta(k|\widetilde{\chi}|^2)$$

Step III: Fix $\boldsymbol{\nu}^{(k)} \in \mathcal{M}_c(\widetilde{\chi}^2)$ and a sample path $(z_1,\ldots,z_{n+k-1})$ such that $L^{(2)}_{n+k} = \boldsymbol{\nu}^{(k)}$. Then

$$\log\mathbb{P}\big((Z_1,\ldots,Z_{n-1}) = (z_1,\ldots,z_{n-1})\big)$$
$$= -(n+k-1)\Big(\mathsf{J}(\boldsymbol{\nu}^{(k)},\xi) - \mathsf{J}(\bar{\boldsymbol{\nu}}^{(k)},\bar{\xi})\Big) \pm \Theta(k|\widetilde{\chi}|^2) + O(1),$$

where the $O(1)$ term comes from the bounded prefactor $\frac{\mathbb{P}(Z_1=z_1)}{\prod_{i=1}^k \mathbb{P}(Z_i=z_i|Z_{i-1}=z_{i-1})}$.

Recall that $D(\boldsymbol{\nu}) := H(\boldsymbol{\nu}) - H(\bar{\boldsymbol{\nu}})$. By Lemma D.1,

$$(2(n+k-1))^{-|\widetilde{\chi}|^2}\exp\big((n+k-1)D(\boldsymbol{\nu}^{(k)})\big) \;\leq\; \big|\{z_1,\ldots,z_{n+k-1} : L^{(2)}_{n+k} = \boldsymbol{\nu}^{(k)}\}\big|$$
$$\leq\; (n+k-1)\exp\big((n+k-1)D(\boldsymbol{\nu}^{(k)})\big).$$

Combining the decomposition

$$\mathbb{P}\big(L^{(2)}_{n+k} = \boldsymbol{\nu}^{(k)}\big) = \sum_{\{z_1,\ldots,z_{n+k-1}: L^{(2)}_{n+k} = \boldsymbol{\nu}^{(k)}\}} \mathbb{P}\big((Z_1,\ldots,Z_{n-1}) = (z_1,\ldots,z_{n-1})\big)$$

with the previous two displays yields

$$\frac{\log \mathbb{P}\big(L^{(2)}_{n+k} = \boldsymbol{\nu}^{(k)}\big)}{n+k-1} \le D(\boldsymbol{\nu}^{(k)}) - \mathsf{J}(\boldsymbol{\nu}^{(k)}, \xi) + \mathsf{J}(\bar{\boldsymbol{\nu}}^{(k)}, \bar{\xi}) + O\left(\frac{\log(n+k)}{n+k}\right) + O\left(\frac{k|\widetilde{\chi}|^2}{n}\right).$$

Define the conditional KL divergence

$$\mathsf{D}(\boldsymbol{\nu}\|\boldsymbol{\mu}) := \sum_{o\in\widetilde{\chi}} \bar{\boldsymbol{\nu}}_o \sum_{p\in\widetilde{\chi}} \frac{\boldsymbol{\nu}_{o,p}}{\bar{\boldsymbol{\nu}}_o} \log\left(\frac{\boldsymbol{\nu}_{o,p}/\bar{\boldsymbol{\nu}}_o}{\boldsymbol{\mu}_{o,p}/\bar{\boldsymbol{\mu}}_o}\right),$$

with the convention that terms with $\bar{\boldsymbol{\nu}}_o = 0$ contribute 0. Then

$$D(\boldsymbol{\nu}) - \mathsf{J}(\boldsymbol{\nu}, \boldsymbol{\mu}) + \mathsf{J}(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\mu}}) = -\mathsf{D}(\boldsymbol{\nu}\|\boldsymbol{\mu}),$$

and hence

$$\begin{aligned}
\frac{1}{n+k-1}\log \mathbb{P}\big(L^{(2)}_{n+k} = \boldsymbol{\nu}^{(k)}\big) &\le -\mathsf{D}(\boldsymbol{\nu}^{(k)}\|\xi) + O\left(\frac{\log(n+k)}{n+k}\right) + O\left(\frac{k|\widetilde{\chi}|^2}{n}\right) \\
&\le -\mathsf{D}(\boldsymbol{\nu}^{(k)}\|\xi) + O\left(\frac{\log(n+k)}{n+k}\right) + O\left(\frac{\kappa|\widetilde{\chi}|^2}{n}\right) \\
&\le -\inf_{\boldsymbol{\nu}\in\mathcal{O}} \mathsf{D}(\boldsymbol{\nu}\|\xi) + O\left(\frac{\log(n)}{n}\right) + O\left(\frac{\kappa|\widetilde{\chi}|^2}{n}\right) \qquad \text{(D.5)}
\end{aligned}$$

where the last inequality follows from the facts that (i) $k \in \{1,\ldots,\kappa\}$ and that $n > \kappa$, and (ii) $\mathsf{D}(\boldsymbol{\nu}^{(k)}\|\xi) > \inf_{\boldsymbol{\nu}\in\mathcal{O}} \mathsf{D}(\boldsymbol{\nu}\|\xi)\ \forall\ \boldsymbol{\nu}^{(k)} \in \mathcal{O}$. Substituting this back into Equation (D.2) we get,

$$\tfrac{1}{n}\log \mathbb{P}\left(L^{(2)}_n \in \mathcal{O}\right) \le -\inf_{\boldsymbol{\nu}\in\mathcal{O}} \mathsf{D}(\boldsymbol{\nu}\|\xi) + O\left(\frac{\log n}{n}\right) + O\left(\frac{\kappa|\widetilde{\chi}|^2}{n}\right)$$

where we have used the fact that $\frac{n+k-1}{n} = 1 + O(\frac{\kappa}{n})$ if $\inf_{\boldsymbol{\nu}\in\bar{\mathcal{O}}} \mathsf{D}(\boldsymbol{\nu}\|\xi) < \infty$ and both sides are trivially ∞ otherwise. We now do the lower bound.

Step IV: Let $\mathcal{O} \subseteq \mathcal{M}_c(\widetilde{\chi}^2)$ be open and fix any $\boldsymbol{\nu} \in \mathcal{O}$. Assume first that $\boldsymbol{\nu}$ is an interior point of $\mathcal{O}$. Then there exists $\varepsilon > 0$ such that the open ball $B(\boldsymbol{\nu}, \varepsilon) \subseteq \mathcal{O}$ (in the Euclidean topology on $\mathbb{R}^{|\widetilde{\chi}|^2}$). The main argument here is that $\cup_{n,k,\boldsymbol{\nu}} \mathcal{E}(n,k,\boldsymbol{\nu})$ is countably dense set in $\mathcal{O}$. Because rational points are dense in $\mathcal{M}_c(\widetilde{\chi}^2)$ and the balance constraints are linear, there exists a sequence of integers $m_\ell \to \infty$ and measures $\boldsymbol{\nu}_\ell \in \mathcal{M}_c(\widetilde{\chi}^2)$ with $m_\ell\boldsymbol{\nu}_\ell$ integer-valued such that $\boldsymbol{\nu}_\ell \in B(\boldsymbol{\nu}, \varepsilon)$ and $\boldsymbol{\nu}_\ell \to \boldsymbol{\nu}$.

For each ℓ, it follows from the facts of the shortest path augmentation described in step 1 that at each state, the augmented measure can differ from the original measure by at most $O(\frac{1}{m_\ell})$, and thus there exists $k_\ell \in \{1,\ldots,\kappa\}$ and an augmented empirical measure $\boldsymbol{\nu}_\ell^{(k_\ell)} \in \mathcal{E}(m_\ell, k_\ell, \boldsymbol{\nu}_\ell)$ satisfying

$$\left\|\boldsymbol{\nu}_\ell^{(k_\ell)} - \boldsymbol{\nu}_\ell\right\|_1 \le \frac{C}{m_\ell},$$

for a constant C depending only on κ and $|\widetilde{\chi}|$. In particular, for all large ℓ we have $\boldsymbol{\nu}_\ell^{(k_\ell)} \in B(\boldsymbol{\nu}, \varepsilon) \subseteq \mathcal{O}$. Therefore, for all large ℓ,

$$\mathbb{P}\big(L^{(2)}_{m_\ell+k_\ell} \in \mathcal{O}\big) \ \ge\ \mathbb{P}\big(L^{(2)}_{m_\ell+k_\ell} = \boldsymbol{\nu}_\ell^{(k_\ell)}\big),$$

and hence

$$\frac{1}{m_\ell}\log \mathbb{P}\big(L^{(2)}_{m_\ell+k_\ell} \in \mathcal{O}\big) \ge \frac{1}{m_\ell}\log \mathbb{P}\big(L^{(2)}_{m_\ell+k_\ell} = \boldsymbol{\nu}_\ell^{(k_\ell)}\big).$$

Following steps similar to Equation (D.5) we have

$$\frac{1}{m_\ell + k_\ell - 1} \log \mathbb{P}\big(L^{(2)}_{m_\ell+k_\ell} = \boldsymbol{\nu}^{(k_\ell)}_\ell\big) \geq -\mathsf{D}(\boldsymbol{\nu}^{(k_\ell)}_\ell \| \xi) - \Omega\left(\frac{\log(m_\ell + k_\ell)}{m_\ell + k_\ell}\right) - \Omega\left(\frac{k_\ell |\widetilde{\chi}|^2}{m_\ell + k_\ell}\right),$$

where $m_\ell \to \infty$ as $\ell \to \infty$ uniformly over $k_\ell \leq \kappa$ and $\frac{m_\ell}{n_\ell} \xrightarrow{\ell\to\infty} 1$. Since $k_\ell \leq \kappa$ is bounded, $\frac{m_\ell + k_\ell - 1}{m_\ell} \to 1$, and therefore

$$\liminf_{\ell\to\infty} \frac{1}{m_\ell} \log \mathbb{P}\big(L^{(2)}_{m_\ell+k_\ell} \in \mathcal{O}\big) \geq -\lim_{\ell\to\infty} \mathsf{D}(\boldsymbol{\nu}^{(k_\ell)}_\ell \| \xi).$$

Finally, by continuity of $\mathsf{D}(\cdot\|\xi)$ on $\mathcal{M}_c(\widetilde{\chi}^2)$ (with the convention $0 \log 0 = 0$), and $\boldsymbol{\nu}^{(k_\ell)}_\ell \to \boldsymbol{\nu}$, we conclude

$$\liminf_{\ell\to\infty} \frac{1}{m_\ell} \log \mathbb{P}\big(L^{(2)}_{m_\ell+k_\ell} \in \mathcal{O}\big) \geq -\mathsf{D}(\boldsymbol{\nu}\|\xi).$$

Since the above inequality holds for every single $\boldsymbol{\nu} \in \mathcal{O}^\circ$, we conclude that

$$\liminf_{\ell\to\infty} \frac{1}{m_\ell} \log \mathbb{P}\big(L^{(2)}_{m_\ell+k_\ell} \in \mathcal{O}\big) \geq -\inf_{\boldsymbol{\nu}\in\mathcal{O}^\circ} \mathsf{D}(\boldsymbol{\nu}\|\xi).$$

Observing that

$$M_{pair}(o,p) := \mathbb{P}(Z_{i+1} = p \mid Z_i = o), \qquad o,p \in \widetilde{\chi}.$$

and

$$\mathsf{D}(\cdot\|\xi) = I(\cdot)$$

completes the proof.

D.1 Proof of Lemma D.1

Proof. Consider the directed multigraph on $\widetilde{\chi}$ with $n^{(k)}_{o,p}$ directed edges from o to p. Because $\boldsymbol{\nu}^{(k)} \in \mathcal{M}_c(\widetilde{\chi}^2)$ is balanced, we have the flow-balance

$$\sum_p n^{(k)}_{o,p} = \sum_p n^{(k)}_{p,o} = \bar{n}^{(k)}_o, \ \forall\, o,$$

so the graph is Eulerian. The number of cyclic sample paths with these edge-multiplicities can be bounded in terms of multinomial allocations of outgoing edges; standard counting arguments give for $o,p \in \widetilde{\chi}$,

$$\frac{\prod_{o\in\widetilde{\chi}}(\bar{n}^{(k)}_o - 1)!}{\prod_{o,p\in\widetilde{\chi}} n^{(k)}_{o,p}!} \ \leq \sum_{\{z_1,\ldots,z_{n+k-1}:L^{(2)}_{n+k}=\boldsymbol{\nu}^{(k)}\}} 1 \ \leq \ (n+k-1)\frac{\prod_{o\in\widetilde{\chi}} \bar{n}^{(k)}_o!}{\prod_{o,p\in\widetilde{\chi}} n^{(k)}_{o,p}!}. \tag{D.6}$$

Our main result in proving this lemma will be the bound on the multinomial coefficient using entropy. Let $\boldsymbol{n}^{(k)}$ be the distribution of empirical frequencies given by $(n^{(k)}_{o,p})_{o,p\in\widetilde{\chi}}$ and $\bar{\boldsymbol{n}}^{(k)}$ be the same with $(\bar{n}^{(k)}_o)_{o\in\chi}$. Then, a combinations of Lemmas 1, and 2 in Csiszar [1998], and the techniques in the proof of Lemma 1 in Vidyasagar [2014] show that the multinomial coefficient

$$\frac{\bar{n}^{(k)}_o!}{\prod_{p\in\widetilde{\chi}} n^{(k)}_{o,p}!}$$

satisfies, for all $n + k \geq |\widetilde{\chi}|$, and $\bar{\boldsymbol{n}}^{(k)}_o = (n^{(k)}_{o,p}/\bar{n}^{(k)}_o)_{p\in\widetilde{\chi}}$

$$(2(n+k-1))^{-|\widetilde{\chi}|^2} \exp\big((n+k-1)H(\bar{\boldsymbol{n}}^{(k)}_o)\big) \ \leq \ \frac{\bar{n}^{(k)}_o!}{\prod_{p\in\widetilde{\chi}} n^{(k)}_{o,p}!} \ \leq \ \exp\big((n+k-1)H(\bar{\boldsymbol{n}}^{(k)}_o)\big).$$

Now, substituting the previous bounds into Equation (D.6) we then get,

$$\begin{aligned}
&(2(n+k-1))^{-|\widetilde{\chi}|^2} \exp\big((n+k-1)D(\boldsymbol{\nu}^{(k)})\big) \\
&\leq \sum_{\{z_1,\ldots,z_{n+k-1}:L^{(2)}_{n+k}=\boldsymbol{\nu}^{(k)}\}} 1 \leq (n+k-1)\exp\big((n+k-1)D(\boldsymbol{\nu}^{(k)})\big)
\end{aligned}$$

□

E A Proof of LDP for $L_n^{(2)}$ using Gärtner–Ellis theorem

To formalise the proofs, we require some results, starting with Perron–Frobenius Theorem [Dembo and Zeitouni, 2010, Theorem 3.1.1].

Theorem E.1 (Perron–Frobenius). *For a fixed index set Σ, let $B = \{B(i,j)\}_{i,j=1}^{|\Sigma|}$ be an irreducible (see Definition C.1) nonnegative matrix. Then B possesses an eigenvalue ρ (called the Perron–Frobenius eigenvalue) such that:*

1. *ρ is a positive real number $\rho > 0$.*
2. *For any eigenvalue λ of B, $|\lambda| \le \rho$.*
3. *There exist left and right eigenvectors corresponding to the eigenvalue ρ that have strictly positive coordinates.*
4. *The left and right eigenvectors μ, θ corresponding to the eigenvalue ρ are unique up to a constant multiple.*
5. *For every $i \in \Sigma$ and every $\sigma = (\sigma_1, \dots, \sigma_{|\Sigma|})$ such that $\sigma_j > 0$ for all j,*

$$\lim_{n\to\infty} \frac{1}{n} \log \left(\sum_{j=1}^{|\Sigma|} B^n(i,j)\, \sigma_j \right) = \lim_{n\to\infty} \frac{1}{n} \log \left(\sum_{j=1}^{|\Sigma|} \sigma_j\, B^n(j,i) \right) = \log \rho. \tag{E.1}$$

Now we state **Gärtner–Ellis theorem**.

Theorem E.2 (Theorem 2.3.6, Dembo and Zeitouni [2010]). *Let $(Z_n)_{n\ge1}$ be a sequence of random vectors in $\mathbb{R}^d$. For each n, define the logarithmic moment generating function (log-mgf)*

$$\Lambda_n(\lambda) := \log \mathbb{E}\left[e^{\langle \lambda, Z_n \rangle} \right], \qquad \lambda \in \mathbb{R}^d.$$

Assume that the following limit exists (possibly as $+\infty$):

$$\Lambda(\lambda) := \lim_{n\to\infty} \frac{1}{n} \Lambda_n(n\lambda), \qquad \lambda \in \mathbb{R}^d. \tag{E.2}$$

Suppose further that

1. *Λ is lower semicontinuous, and $\liminf_{x\to\lambda} \Lambda(x) \ge \Lambda(\lambda)$ and essentially smooth (as defined in Definition 2.3.5 Dembo and Zeitouni [2010]).*
2. *the effective domain $\mathcal{D}_\Lambda := \{\lambda \in \mathbb{R}^d : \Lambda(\lambda) < \infty\}$ has non-empty interior.*

Then (Z_n) satisfies a Large Deviation Principle (LDP) with good convex rate function given by the Legendre–Fenchel transform of Λ:

$$I(x) := \sup_{\lambda \in \mathbb{R}^d} \{ \langle \lambda, x \rangle - \Lambda(\lambda) \}, \qquad x \in \mathbb{R}^d.$$

That is, for any Borel set $A \subseteq \mathbb{R}^d$,

$$-\inf_{x \in A^\circ} I(x) \;\le\; \liminf_{n\to\infty} \tfrac{1}{n} \log \mathbb{P}(Z_n \in A) \;\le\; \limsup_{n\to\infty} \tfrac{1}{n} \log \mathbb{P}(Z_n \in A) \;\le\; -\inf_{x \in \overline{A}} I(x).$$

Theorem E.3 (Contraction Principle Dembo and Zeitouni [2010]). *Let $(\mathcal{X}, \tau)$ and $(\mathcal{Y}, \tau')$ be Hausdorff topological spaces, and let $f : \mathcal{X} \to \mathcal{Y}$ be continuous. Suppose that $\{X_n\}_{n\ge1}$ satisfies a large deviation principle on $\mathcal{X}$ with good rate function $I : \mathcal{X} \to [0, \infty]$. Then the sequence $\{f(X_n)\}_{n\ge1}$ satisfies a large deviation principle on $\mathcal{Y}$ with good rate function*

$$J(y) := \inf\{\, I(x) : x \in \mathcal{X},\ f(x) = y\}, \qquad y \in \mathcal{Y}.$$

Let $(X_t, a_t)_{t\ge0}$ be a finite-state irreducible Markov chain with state space $\widetilde{\chi} = \chi \times \mathbb{I}$ and define the pair process

$$Z_i^{(2)} := \big((X_i, a_i), (X_{i+1}, a_{i+1})\big), \qquad i \ge 1,$$

with finite state space $\widetilde{\chi}^2$. Observe that $Z_i^{(2)}$ is Markovian with transition kernel

$$M_{pair}^{(2)}\big((y,b)\times(y',b'),(x,a)\times(x',a')\big) = \mathbb{I}_{(y',b')}((x,a))\,\mathbb{P}\left(Z_{i+1}=(x',a')\mid Z_i=(x,a)\right).$$

where $\mathbb{P}\left(Z_{i+1}=(x',a')\mid Z_i=(x,a)\right) := \mathbb{P}(X_{i+1}=x'\mid X_i=x,a_i=a)\,\mathbb{P}(a_{i+1}=a'\mid X_{i+1}=x)$. For simplicity, we have assumed that $M_{pair}(i,j)>0,\ \forall\, i,j\in\widetilde{\chi}$, which easily implies that the paired process $Z_t^{(2)}$ is irreducible. The empirical transition measure of $Z_i^{(2)}$ is

$$L_n^{(2)}(o,p) \;=\; \frac{1}{n}\sum_{i=1}^{n}\mathbb{1}[Z_i=o,Z_{i+1}=p], \qquad s_1,s_2\in\widetilde{\chi}.$$

Let $M_{pair}^{(2)}$ denote the transition matrix of the chain $(Z_i^{(2)})$ and, for $\lambda = \Big(\lambda(y,b,y',b')\Big)_{(y,b,y',b')\in\widetilde{\chi}^2} \in \mathbb{R}^{d\times k\times d\times k}(=\mathbb{R}^{|\widetilde{\chi}|^2})$, define the tilted matrix

$$M_{pair,\lambda}^{(2)}(s,t) := M_{pair}^{(2)}(s,t)\,e^{\lambda_t}, \qquad s,t\in\widetilde{\chi}^2.$$

We further define $\mathcal{M}_c(\widetilde{\chi}^2)$ for the simplex of probability measures on $\widetilde{\chi}^2$.

Proposition E.4. *$\{L_n^{(2)}\}_{n\geq 1}$ satisfies a Large Deviation Principle (LDP) on the simplex of probability measures on $\widetilde{\chi}^2$ with good convex rate function*

$$I(\mu) \;=\; \sup_{\lambda\in\mathbb{R}^{|\widetilde{\chi}|^2}}\Big\{\langle\lambda,\mu\rangle - \log\rho(M_{pair,\lambda}^{(2)})\Big\},$$

where $\rho(\cdot)$ denotes the Perron–Frobenius eigenvalue.

Proof. We wish to apply Gärtner–Ellis theorem on $L_n^{(2)}$ for which, we require the corresponding log-mgf $\Lambda_n^{(L_n)}$ and its asymptotic counterpart $\Lambda^{(L_n)}$. Observe that

We define the averaged probability and expectation operators by

$$\mathbb{P}_\sigma(\cdot) := \sum_{z\in\widetilde{\chi}^2}\sigma(\{z\})\,\mathbb{P}_z(\cdot), \qquad \mathbb{E}_\sigma(\cdot) := \sum_{z\in\widetilde{\chi}^2}\sigma(\{z\})\,\mathbb{E}_z(\cdot), \tag{E.3}$$

that is, the law and expectation when the initial state is drawn according to the distribution σ. Furthermore, we set

$$M_{pair}^{(2)}(\sigma,z_1) := \sum_{z\in\widetilde{\chi}^2}\sigma(\{z\})M_{pair}^{(2)}(z,z_1). \tag{E.4}$$

Now observe that

$$\Lambda_n^{(L_n^{(2)})}(\lambda) = \log\mathbb{E}_\sigma\left[\exp\left(\tfrac{1}{n}\sum_{i=1}^{n}\lambda_s\mathbb{1}_{\{Z_i^{(2)}=s\}}\right)\right] = \log\mathbb{E}_\sigma\left[\exp\left(\tfrac{1}{n}\sum_{i=1}^{n}\lambda_{Z_i^{(2)}}\right)\right].$$

w

So that (following Equation (E.2)), $\Lambda_n^{(L_n^{(2)})}(n\lambda)$ becomes

$$\begin{aligned}
\Lambda_n^{(L_n^{(2)})}(n\lambda) &= \log\left(\sum_{z_1^{(2)},\ldots,z_n^{(2)}\in\widetilde{\chi}^{(2)}}\mathbb{P}_\sigma(Z_1^{(2)}=z_1,\ldots,Z_n^{(2)}=z_n)\,e^{\sum_{i=1}^n\lambda_{z_i}}\right)\\
&= \log\left(\sum_{z_1,\ldots,z_n}M_{pair}^{(2)}(\sigma,z_1)\prod_{i=1}^{n-1}M_{pair}^{(2)}(z_i,z_{i+1})\prod_{i=1}^{n}e^{\lambda_{z_i}}\right)\\
&\overset{(i)}{=} \log\left(\sum_{z_1,\ldots,z_n}M_{pair,\lambda}^{(2)}(\sigma,z_1)\prod_{i=1}^{n-1}M_{pair,\lambda}^{(2)}(z_i,z_{i+1})\right)\\
&\overset{(ii)}{=} \log\left(\sum_{z_n}^{|\tilde{S}|}\big(M_{pair,\lambda}^{(2)}\big)^n(\sigma,z_n)\right),
\end{aligned}$$

where (i) follows from $M^{(2)}_{pair,\lambda}(s,t) = M^{(2)}_{pair}(s,t)e^{\lambda_t}$ and (ii) follows from matrix multiplication. By item 5 of Theorem E.1 (choosing σ to be a tensor of all 1's) for the finite irreducible nonnegative matrix $M_{pair,\lambda}$,

$$\Lambda(\lambda) := \lim_{n\to\infty} \frac{1}{n}\Lambda_n(n\lambda) = \log \rho(M^{(2)}_{pair,\lambda}). \tag{E.5}$$

As $\rho(M^{(2)}_{pair,\lambda}) > 0$ and is differentiable in λ (simple dominant eigenvalue), $\Lambda(\lambda)$ is finite, convex and differentiable (see Lancaster [1969], Theorem 7.7.1). Hence the Gärtner–Ellis theorem applies and yields the LDP for $L^{(2)}_n$ with good convex rate function

$$I(\mu) = \sup_{\lambda\in\mathbb{R}^{|\widetilde{\chi}|^2}} \Big\{\langle\lambda,\mu\rangle - \Lambda(\lambda)\Big\} = \sup_{\lambda} \Big\{ \sum_{s\in\widetilde{\chi}^2} \lambda_s\mu(s) - \log\rho(M^{(2)}_{pair,\lambda})\Big\}.$$

This completes the proof. □

F Equivalence of Rate Functions

We then provide an alternative characterization of $I(\mu)$. We use the notation $v \succ 0$ to denote vectors with strictly positive entries.

Proposition F.1. *Define, for $q \in \mathcal{M}_c(\widetilde{\chi}^2)$,*

$$J(q) := \sup_{u\succ 0} \sum_{s\in\widetilde{\chi}} q_s \log \frac{u_s}{(uM^{(2)}_{pair})_s},$$

and set $J(q) = \infty$ for $q \notin \mathcal{M}_c(\widetilde{\chi}^2)$. Then, under the conditions of Proposition E.4

$$I(q) = J(q) \quad \textit{for every } q \in \mathcal{M}_c(\widetilde{\chi}^2).$$

Proof. Fix $q \in \mathcal{M}_c(\widetilde{\chi}^2)$. We prove $I(q) = J(q)$ by two inequalities.

(1) $I(q) \geq J(q)$. Let $u \succ 0$ be arbitrary and define $\lambda \in \mathbb{R}^{|\widetilde{\chi}|^2}$ by

$$\lambda_s := \log \frac{u_s}{(uM^{(2)}_{pair})_s}, \qquad s \in \widetilde{\chi}^2.$$

With this choice,

$$(uM^{(2)}_{pair,\lambda})_t = \sum_s u_s M^{(2)}_{pair}(s,t)e^{\lambda_t} = u_t,$$

so $uM^{(2)}_{pair,\lambda} = u$ which implies $u\left(M^{(2)}_{pair}\right)^n_\lambda = u$ and hence $\rho\left(M^{(2)}_{pair,\lambda}\right) = 1$ by part 5 in Theorem E.1. Therefore

$$\sum_s \lambda_s q_s - \log\rho(M^{(2)}_{pair,\lambda}) = \sum_s q_s \log\frac{u_s}{(uM^{(2)}_{pair})_s}.$$

Taking supremum over all λ on the left (the definition of $I(q)$) yields

$$I(q) \geq \sum_s q_s \log\frac{u_s}{(uM^{(2)}_{pair})_s}.$$

Since $u \succ 0$ was arbitrary, $I(q) \geq \sup_{u\succ 0}\sum_s q_s \log\frac{u_s}{(uM^{(2)}_{pair})_s} = J(q)$.

(2) $I(q) \leq J(q)$. Let $\lambda \in \mathbb{R}^{|\widetilde{\chi}|^2}$ be arbitrary and let $r \succ 0$ be a strictly positive right eigenvector of M_λ corresponding to its Perron–Frobenius eigenvalue $\rho := \rho(M^{(2)}_{pair,\lambda})$, so

$$rM^{(2)}_{pair,\lambda} = \rho r.$$

Equivalently,

$$e^{\lambda_s}(rM^{(2)}_{pair})_s = \rho r_s, \qquad s \in \widetilde{\chi}^2,$$

and it follows that

$$\lambda_s = \log \frac{\rho r_s}{(rM^{(2)}_{pair})_s} = \log r_s - \log(rM^{(2)}_{pair})_s + \log \rho.$$

We can see that

$$\sum_s \lambda_s q_s - \log \rho = \sum_s q_s \big(\log r_s - \log(rM^{(2)}_{pair})_s \big) = \sum_s q_s \log \frac{r_s}{(rM^{(2)}_{pair})_s}.$$

But the right-hand side is one admissible value in the supremum defining $J(q)$ (take $u = r$). Thus for this λ we have

$$\sum_s \lambda_s q_s - \log \rho(M^{(2)}_{pair,\lambda}) \le J(q).$$

Taking supremum over all λ gives $I(q) \le J(q)$. Combining (1) and (2) yields $I(q) = J(q)$ for every $q \in \mathcal{M}_c(\widetilde{\chi}^2)$, as required.

Let us now consider the case when $q \notin \mathcal{M}_c(\widetilde{\chi}^2)$. For every $n \in \mathbb{N}$ and every realization of the Markov chain, the empirical measure satisfies

$$L^{(2)}_n \in \mathcal{M}_c(\widetilde{\chi}^2),$$

where $\mathcal{M}_c(\widetilde{\chi}^2)$ denotes the probability simplex on $\widetilde{\chi}^2$. Note that $\mathcal{M}_c(\widetilde{\chi}^2)$ is a closed subset of $\mathbb{R}^{|\widetilde{\chi}|^2}$, since $\mathbb{R}^{|\widetilde{\chi}|^2}_{\{\ge 0\}}$ is closed in $\mathbb{R}^{|\widetilde{\chi}|^2}$, the affine hyperplane

$$\Big\{ x \in \mathbb{R}^{|\widetilde{\chi}|^2} : \sum_s x_s = 1 \Big\}$$

is also closed, and therefore their intersection, namely $\mathcal{M}_c(\widetilde{\chi}^2)$, is closed as well. Consequently,

$$\mathbb{P}\left(L^{(2)}_n \in \mathcal{M}_c(\widetilde{\chi}^2)^c \right) = 0 \quad \text{for all } n,$$

and therefore

$$-\infty = \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(L^{(2)}_n \in \mathcal{M}_c(\widetilde{\chi}^2)^c \right).$$

By the lower bound in the large deviation principle, this implies for the open set $\mathcal{M}_c(\widetilde{S})^c$,

$$-\infty \ \ge \ - \inf_{q \notin \mathcal{M}_c(\widetilde{\chi}^2)} I(q),$$

which forces

$$I(q) = \infty \quad \text{for all } q \notin \mathcal{M}_c(\widetilde{\chi}^2).$$

□

The analysis so far provides us with a form of $I(q)$ in which it is still an optimization problem, and we wish to provide a representation in the form of relative entropy, which is more elegant. Recall the definition of shift invariance to denote $q \in \mathcal{M}_c(\widetilde{\chi}^2)$ such that $q_1 = q_2$, where

$$q_1(i) \ = \ \sum_{j \in \widetilde{\chi}} q(i,j) \qquad \text{and} \qquad q_2(i) \ = \ \sum_{j \in \widetilde{\chi}} q(j,i).$$

In the following theorems we assume, without loss of generality, that M_{pair} is strictly positive. Indeed, an analogous result holds if one replaces $\widetilde{\chi}$ by $\{(i,j) : M_{pair}(i,j) > 0\}$.

Proposition F.2. *For every probability measure q on $\widetilde{\chi}^2$,*

$$I(q) = \begin{cases} \displaystyle\sum_{i \in \widetilde{\chi}} q_1(i) \, \mathrm{KL}\left(q_f(\cdot \mid i) \parallel M_{pair}(i,\cdot) \right), & \text{if } q \in \mathcal{M}_c(\widetilde{\chi}), \\ +\infty, & \text{otherwise}. \end{cases}$$

Here $q_f(\cdot \mid i) = q(i,\cdot)/q_1(i)$ whenever $q_1(i) > 0$, and

$$\mathrm{KL}\left(q_f(\cdot \mid i) \parallel M_{pair}(i,\cdot) \right) = \sum_{j \in \widetilde{\chi}} q_f(j \mid i) \log \frac{q_f(j \mid i)}{M_{pair}(i,j)}.$$

Proof. We treat two cases.

Case 1: q is not shift-invariant. Then there exists $j_0 \in \widetilde{\chi}$ with $q_1(j_0) < q_2(j_0)$. For $\alpha > 0$ define the test function $u^{(\alpha)} \succ 0$ by

$$u^{(\alpha)}(i,j) = \begin{cases} e^{\alpha}, & j = j_0, \\ 1, & j \neq j_0. \end{cases}$$

Compute the objective $J(q)$ in Theorem F.1 for $u^{(\alpha)}$ and using the definition of $M^{(2)}_{pair}$ in the second equality, we have:

$$\begin{aligned} \sum_{s\in\widetilde{\chi}} q_s \log \frac{u^{(\alpha)}_s}{(u^{(\alpha)} M^{(2)}_{pair})_s} &= \sum_{i,j\in\widetilde{\chi}} q(i,j) \log \frac{u^{(\alpha)}(i,j)}{\big(\sum_k u^{(\alpha)}(k,i)\big)\, M_{pair}(i,j)} \\ &= \sum_{i,j\in\widetilde{\chi}} q(i,j) \log u^{(\alpha)}(i,j) \\ &\quad - \sum_{i,j\in\widetilde{\chi}} q(i,j) \log \big(\sum_k u^{(\alpha)}(k,i)\big) - \sum_{i,j\in\widetilde{\chi}} q(i,j) \log M_{pair}(i,j). \end{aligned}$$

Only the first two sums depend on α. The first sum equals $\alpha\, q_2(j_0)$. The second sum equals $\sum_i q_1(i) \log\big(\sum_k u^{(\alpha)}(k,i)\big)$, and for each fixed i the quantity $\sum_k u^{(\alpha)}(k,i)$ is either of order e^{α} (if some of the terms have j_0 as second coordinate) or bounded, so that the second sum contributes $\alpha\, q_1(j_0) + O(1)$. Hence the overall α-dependent part is

$$\alpha\big(q_2(j_0) - q_1(j_0)\big) + O(1).$$

Because $q_2(j_0) - q_1(j_0) > 0$, letting $\alpha \to \infty$ drives the expression to $+\infty$. Thus the supremum in $J(q)$ is $+\infty$ and $I(q) = \infty$ for non shift-invariant q.

Case 2: q is shift-invariant. Assume $q_1 = q_2$. Fix any $u \succ 0$ and write for each state $j \in \widetilde{\chi}$,

$$s_j := \sum_k u(k,j) > 0, \qquad u(i \mid j) := \frac{u(i,j)}{s_j}.$$

Also write the backward conditional $q_b(i \mid j) := q(i,j)/q_2(j)$ (well-defined when $q_2(j) > 0$). Using the definition of $(u\mathcal{M})(i,j)$ and shift-invariance we have (since $q_1 = q_2$)

$$\begin{aligned} \sum_{i,j\in\widetilde{\chi}} q(i,j)\, \log \frac{u(i,j)}{\big(\sum_k u(k,i)\big)\, M_{pair}(i,j)} &= \sum_{i,j\in\widetilde{\chi}} q(i,j)\, \log \frac{u(i,j)\, q_2(i)}{(\sum_k u(k,i))\, q_1(j)\, M_{pair}(i,j)} \\ &= \sum_{i,j\in\widetilde{\chi}} q(i,j) \log \frac{u(i \mid j)}{q_b(i \mid j)} + \sum_{i,j\in\widetilde{\chi}} q(i,j) \log \frac{q_f(j \mid i)}{M_{pair}(i,j)}. \end{aligned} \tag{F.1}$$

In the second term we recognize

$$\sum_{i,j\in\widetilde{\chi}} q(i,j) \log \frac{q_f(j \mid i)}{M_{pair}(i,j)} = \sum_i q_1(i) \mathrm{KL}\big(q_f(\cdot \mid i) \,\|\, \mathbb{P}(Z_{k+1} = . \mid Z_k = i)\big).$$

For the first term observe

$$\begin{aligned} \sum_{i,j} q(i,j)\, \log \frac{u(i \mid j)}{q_b(i \mid j)} &= \sum_j q_2(j) \sum_i q_b(i \mid j)\, \log \frac{u(i \mid j)}{q_b(i \mid j)} \\ &= -\sum_j q_2(j)\, \mathrm{KL}\big(q_b(\cdot \mid j) \,\|\, u(\cdot \mid j)\big). \end{aligned}$$

Since relative entropy is nonnegative we obtain for every $u \succ 0$ the upper bound

$$\sum_{i,j} q(i,j)\, \log \frac{u(i,j)}{\big(\sum_k u(k,i)\big)\, \mathbb{P}(Z_{k+1} = j \mid Z_k = i)} \leq \sum_i q_1(i)\, \mathrm{KL}\big(q_f(\cdot \mid i) \,\|\, \mathbb{P}(Z_{k+1} = . \mid Z_k = i)\big).$$

Taking supremum over $u \succ 0$ yields

$$I(q) \leq \sum_i q_1(i)\,\mathrm{KL}\big(q_f(\cdot \mid i) \,\|\, \mathbb{P}(Z_{k+1} = . \mid Z_k = i)\big).$$

To obtain the reverse inequality in the case when $q \succ 0$, take

$$u(i,j) = q(i,j) \quad \text{for all } i,j,$$

and then $s_j = \sum_k u(k,j) = q_2(j) = q_1(j)$, hence the objective in Theorem F.1 equals

$$\sum_i q_1(i)\,\mathrm{KL}\big(q_f(\cdot \mid i) \,\|\, \mathbb{P}(Z_{k+1} = . \mid Z_k = i)\big).$$

Therefore

$$I(q) \geq \sum_i q_1(i)\,\mathrm{KL}\big(q_f(\cdot \mid i) \,\|\, \mathbb{P}(Z_{k+1} = . \mid Z_k = i)\big),$$

and combining with the previous inequality gives equality for strictly positive q:

$$I(q) = \sum_i q_1(i)\,\mathrm{KL}\big(q_f(\cdot \mid i) \,\|\, \mathbb{P}(Z_{k+1} = . \mid Z_k = i)\big).$$

This completes the proof in the case when $q \succ 0$. Now, to deal with the case when q is not strictly positive, we can consider a sequence $u_n \succ 0$, where $u_n \to q$ and so $q_2(j)\,\mathrm{KL}\big(q_b(\cdot \mid j) \,\|\, u_n(\cdot \mid j)\big) \to 0$ for each $j \in \widetilde{\chi}$. Thus, we obtain the same conclusion for the case when q is not strictly positive. Finally, observing that

$$M_{pair}(i,j) := \mathbb{P}(Z_{k+1} = j \mid Z_k = i), \qquad o, p \in \widetilde{\chi}.$$

completes the proof. □

F.1 Proof of Proposition 4.3

Proof. Let $\chi = \{1,2\}$ and $\mathbb{I} = \{\alpha, \beta\}$. Define a deterministic policy $\pi : \chi \to \mathbb{I}$ by

$$\pi(1) = \alpha, \qquad \pi(2) = \beta,$$

and set $a_i = \pi(X_i)$ for all i.

Define the controlled transition kernels $\{M^{(l)}\}_{l \in \mathcal{A}}$ by

$$M^{(\alpha)}_{1,2} = 1, \quad M^{(\alpha)}_{1,1} = 0, \qquad M^{(\beta)}_{2,1} = 1, \quad M^{(\beta)}_{2,2} = 0,$$

and (arbitrarily) specify the remaining rows so that each $M^{(l)}$ is a stochastic matrix. Under the policy $a_i = \pi(X_i)$, the state process evolves deterministically:

$$X_1 \in \{1,2\} \quad \Rightarrow \quad X_{i+1} = 3 - X_i \ \text{ for all } i \geq 1.$$

Hence the pair process $Z_i = (X_i, a_i)$ takes values in the two-point set

$$\widetilde{\chi}_\pi := \{(1,\alpha), (2,\beta)\} \subset \widetilde{\chi},$$

and alternates deterministically:

$$(1,\alpha) \mapsto (2,\beta) \mapsto (1,\alpha) \mapsto \cdots .$$

Therefore $\{Z_i\}$ is a time-homogeneous Markov chain on $\widetilde{\chi}_\pi$ with transition matrix

$$M_{\mathrm{pair}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{(in the ordered basis } (1,\alpha), (2,\beta)),$$

which is irreducible and has period 2. In particular, there is no $\kappa \in \mathbb{N}$ such that $M^\kappa_{\mathrm{pair}} > 0$.

Recurrence: Under the policy $a_i = \pi(X_i)$, we have $Z_i \in \widetilde{\chi}_\pi$ for all i. Moreover, by construction the evolution is deterministic:

$$(1,\alpha) \mapsto (2,\beta) \mapsto (1,\alpha) \mapsto (2,\beta) \mapsto \cdots$$

and similarly if started from $(2,\beta)$. Hence, if $Z_1 = (1,\alpha)$ then

$$Z_{2m-1} = (1,\alpha), \qquad Z_{2m} = (2,\beta) \qquad \forall m \geq 1,$$

and if $Z_1 = (2,\beta)$ then

$$Z_{2m-1} = (2,\beta), \qquad Z_{2m} = (1,\alpha) \qquad \forall m \geq 1.$$

In either case, each of $(1,\alpha)$ and $(2,\beta)$ is visited at all even (respectively odd) times, and therefore is visited infinitely often with probability one.

Finally, since $\widetilde{\chi}_\pi$ is a finite closed communicating class, recurrence implies positive recurrence. Concretely, the return time to each state is identically 2, so the expected return time is finite.

No LDP: Now consider the empirical pair measure $L_n^{(2)}$ on $\widetilde{\chi}^2$. Let

$$\nu^\star \in \mathcal{M}_c(\widetilde{\chi}^2) \quad \text{be defined by} \quad \nu^\star_{(1,\alpha),(2,\beta)} = \nu^\star_{(2,\beta),(1,\alpha)} = \frac{1}{2},$$

and $\nu^\star_{o,p} = 0$ for all other $(o,p) \in \widetilde{\chi}^2$. Then for every even n,

$$L_n^{(2)} = \nu^\star \quad \text{almost surely.}$$

For odd n, the last transition count is unbalanced by 1, and one has

$$\left\| L_n^{(2)} - \nu^\star \right\|_1 = \frac{1}{n} \quad \text{almost surely.}$$

Let $\mathcal{O}$ be any open neighborhood of $\nu^\star$ in $\mathcal{M}_c(\widetilde{\chi}^2)$. Choose n_0 such that $1/n_0$ is smaller than the radius of $\mathcal{O}$ in $\|\cdot\|_1$. Then for all $n \geq n_0$,

$$\mathbb{P}(L_n^{(2)} \in \mathcal{O}) = \begin{cases} 1, & n \text{ even}, \\ 0, & n \text{ odd}, \end{cases}$$

because $L_n^{(2)} = \nu^\star$ when n is even, while $L_n^{(2)}$ stays at $\|\cdot\|_1$-distance exactly $1/n$ from $\nu^\star$ when n is odd.

Consequently,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(L_n^{(2)} \in \mathcal{O}) = 0, \qquad \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(L_n^{(2)} \in \mathcal{O}) = -\infty.$$

Thus the exponential upper and lower rates do not match, and therefore no rate function $I : \mathcal{M}_c(\widetilde{\chi}^2) \to [0,\infty]$ can satisfy the LDP bounds in Daures [2025, Definition 1.1]. Hence $\{L_n^{(2)}\}$ does not satisfy an LDP. That a types bound cannot hold is now a simple deduction. □

G The in-Ergodic Case

To formalise the in-ergodic case, we require some notations. Let $\mathcal{A}$ be the set of all *admissible measures* (defined as in Daures [2025], Definition 1.2) and fix $\mathcal{A} \cap \boldsymbol{\mu} \in \mathcal{M}_c(\widetilde{\chi})$. Let $\{C_j\}_{j\in\mathcal{J}}$ denote the collection of irreducible (communicating) classes of the Markov chain associated with $\boldsymbol{\mu}$, indexed by $\mathcal{J}$ (we point to Meyn and Tweedie [2012] for the relevant definitions). For each $j \in \mathcal{J}$, define the normalized restriction of $\boldsymbol{\mu}$ to C_j^2 by

$$\tilde{\boldsymbol{\mu}}_j \;:=\; \boldsymbol{\mu}(C_j^2)^{-1}\, \boldsymbol{\mu}\big|_{C_j^2}.$$

Then, we have the following theorem

Theorem G.1. *$L_n^{(2)}$ satisfies a LDP with the good rate function*

$$I_{red}(\boldsymbol{\mu}) \;=\; \sum_{j\in\mathcal{J}} \boldsymbol{\mu}(C_j^2)\, I(\tilde{\boldsymbol{\mu}}_j). \tag{G.1}$$

where I is the good rate function from Daures [2025, Theorem 1.3].

H Proof of Theorem G.1

Definition H.1. Let $\big(L_N^{(2)}\big)_{N\geq 1}$ be $\mathcal{M}(\widetilde{\chi}^2)$-valued random elements, where $\mathcal{M}(\widetilde{\chi}^2)$ is equipped with its Borel σ-field induced by the weak topology.

(1) **Exponential tightness.** We say $\big(L_N^{(2)}\big)$ is *exponentially tight* if for every $\alpha < \infty$ there exists a compact set $K \subset \mathcal{M}(\widetilde{\chi}^2)$ such that

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}\big(L_N^{(2)} \notin K\big) \leq -\alpha.$$

(2) **Weak/full LDP.** We say $\big(L_N^{(2)}\big)$ satisfies the *weak* LDP if there exists a lower semicontinuous function $I : \mathcal{M}(\widetilde{\chi}^2) \to [0,\infty]$ such that

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}\big(L_N^{(2)} \in K\big) \leq -\inf_{\mu\in K} I(\mu), \qquad \text{for every compact } K \subset \mathcal{M}(\widetilde{\chi}^2),$$

and

$$\liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}\big(L_N^{(2)} \in U\big) \geq -\inf_{\mu\in U} I(\mu), \qquad \text{for every open } U \subset \mathcal{M}(\widetilde{\chi}^2).$$

The function I is called the *rate function*. It is *good* if its level sets $\{\mu \in \mathcal{M}(\widetilde{\chi}^2) : I(\mu) \leq C\}$ are compact for every $C \in \mathbb{R}$.

Definition H.2 (Ruelle–Lanford functions)**.** Let $\{L_n^{(2)}\}_{n\geq 1}$ be a sequence of random elements taking values in $\mathcal{M}_c(\widetilde{\chi}^2)$, equipped with the total variation topology. For any Borel set $A \subset \mathcal{M}_c(\widetilde{\chi}^2)$, define

$$\underline{s}(A) \;:=\; \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(L_n^{(2)} \in A\right), \qquad \overline{s}(A) \;:=\; \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(L_n^{(2)} \in A\right).$$

For any $\nu \in \mathcal{M}_c(\widetilde{\chi}^2)$, define the pointwise limits

$$\underline{s}(\nu) \;:=\; \lim_{\delta\downarrow 0} \underline{s}\left(B(\nu,\delta)\right), \qquad \overline{s}(\nu) \;:=\; \lim_{\delta\downarrow 0} \overline{s}\left(B(\nu,\delta)\right),$$

where $B(\nu,\delta)$ denotes the open ball of radius δ centered at ν. We say that $\{L_n^{(2)}\}$ admits a *Ruelle–Lanford (RL) function* $s(\nu)$ if

$$\underline{s}(\nu) = \overline{s}(\nu) =: s(\nu) \qquad \text{for all } \nu \in \mathcal{M}_c(\widetilde{\chi}^2).$$

For $t \in \mathbb{N}$, a *word* (path) of length $t+1$ is $z_{1:t+1} = (z_1, \ldots, z_{t+1}) \in \widetilde{\chi}^{t+1}$.

We now recall Theorem 4.1.1 in Dembo and Zeitouni [2010]

Lemma H.3. *Assume that $\big(L_N^{(2)}\big)_{N\geq 1}$ admits a Ruelle–Lanford (RL) function s, i.e. the associated lower/upper exponents coincide: $\underline{s} = \overline{s} =: s$ on $\mathcal{M}(\widetilde{\chi}^2)$. Then s is upper semicontinuous (w.r.t. the weak topology), and $\big(L_N^{(2)}\big)_{N\geq 1}$ satisfies the weak LDP on $\mathcal{M}(\widetilde{\chi}^2)$ with rate function*

$$I(\mu) := -s(\mu), \qquad \mu \in \mathcal{M}(\widetilde{\chi}^2).$$

The proof of the following Lemma will be following the broad strategy outlined in Theorem 2.4 Daures [2025].

Lemma H.4. *The sequence $(L_n^{(2)})$ admits a Ruelle–Lanford function $s := \underline{s} = \overline{s}$ and satisfies the weak large deviation principle with rate function*

$$I := -s \qquad \text{on } \mathcal{M}_c(\widetilde{\chi}).$$

Proof. Let $\boldsymbol{\mu} \in \mathcal{M}_c(\widetilde{\chi})$. We set $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \boldsymbol{\mu}$ and $S_2 = \mathcal{S}$ in Proposition 2.4 of Daures which yields

$$\mathbb{P}\left(L_t^{(2)} \in B(\mu, 2\delta)\right) \;\geq\; C_{n,t}\, \mathbb{P}\left(L_n^{(2)} \in B(\mu,\delta)\right)^N$$

where $B(\mu, \delta)$ is the ball of radius δ around μ in the TV metric. Taking logarithms on both sides, dividing by t, and letting $t \to \infty$, we obtain

$$\underline{s}\left(B(\mu, 2\delta)\right) \;\geq\; \lim_{t\to\infty} \frac{1}{t} \log C_{n,t} \;+\; \frac{1}{n(1-4\delta)} \log \mathbb{P}\left(L_n^{(2)} \in B(\mu, \delta)\right).$$

Taking the limit superior as $n \to \infty$ we have

$$\underline{s}\left(B(\mu, 2\delta)\right) \;\geq\; \frac{1}{1-4\delta}\, \overline{s}\left(B(\mu, \delta)\right).$$

Letting $\delta \downarrow 0$ yields

$$\underline{s}(\mu) \;\geq\; \overline{s}(\mu).$$

Since the reverse inequality $\underline{s}(\mu) \leq \overline{s}(\mu)$ holds trivially, the sequence $(L_n^{(2)})$ admits a Ruelle–Lanford function $s = \underline{s} = \overline{s}$. The weak large deviation principle with rate function $-s$ then follows from Lemma H.3. □

We will now specialize to our finite case to establish that this LDP is full. For that we require the following Lemma.

Lemma H.5. *Let X be a Hausdorff space, and let $\{X_n\}_{n\geq 1}$ be X-valued random elements with laws $P_n(\,\cdot\,) \coloneqq \mathbb{P}(X_n \in \cdot)$. Suppose that $\{P_n\}$ satisfies a* weak *large deviation principle on X with speed n and (not necessarily good) rate function $I : X \to [0, \infty]$, i.e.,*

$$\liminf_{n\to\infty} \frac{1}{n} \log P_n(G) \geq -\inf_{x\in G} I(x) \qquad \textit{for every open set } G \subset X, \tag{H.1}$$

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n(K) \leq -\inf_{x\in K} I(x) \qquad \textit{for every compact set } K \subset X. \tag{H.2}$$

The Lemma is stated in parts

1. *Assume further that $\{P_n\}$ is* exponentially tight*, namely: for each $a > 0$ there exists a compact set $K_a \subset X$ such that*

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n(K_a^c) \leq -a. \tag{H.3}$$

   Then $\{P_n\}$ satisfies the full *LDP on X with speed n and the same rate function I, i.e. the upper bound* (H.2) *extends to all closed sets $F \subset X$:*

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n(F) \leq -\inf_{x\in F} I(x), \qquad \textit{for every closed set } F \subset X. \tag{H.4}$$

2. *If X is compact, then any sequence of probability measures $\{P_n\}$ on X is exponentially tight. Consequently, on a compact space, a weak LDP implies a full LDP.*

3. *let $I : X \to [0, \infty]$ be lower semicontinuous. Then I is a good rate function; that is, for every $a < \infty$ the sublevel set $\{x \in X :\, I(x) \leq a\}$ is compact.*

Proof. **Part (1):** Fix a closed set $F \subset X$ and $a > 0$. By exponential tightness, choose a compact $K_a \subset X$ satisfying (H.3). Then for each n,

$$P_n(F) \leq P_n(F \cap K_a) + P_n(K_a^c).$$

Using $\log(u + v) \leq \log 2 + \max\{\log u, \log v\}$ for $u, v \geq 0$, we obtain

$$\frac{1}{n} \log P_n(F) \leq \frac{\log 2}{n} + \max\Big\{\frac{1}{n} \log P_n(F \cap K_a),\; \frac{1}{n} \log P_n(K_a^c)\Big\}.$$

Taking $\limsup_{n\to\infty}$ and noting that $\frac{\log 2}{n} \to 0$, we get

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n(F) \leq \max\Big\{\limsup_{n\to\infty} \frac{1}{n} \log P_n(F \cap K_a),\; \limsup_{n\to\infty} \frac{1}{n} \log P_n(K_a^c)\Big\}. \tag{H.5}$$

Since $F \cap K_a$ is compact (being closed in a compact set), the weak-LDP upper bound (H.2) yields

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n(F \cap K_a) \leq - \inf_{x \in F \cap K_a} I(x),$$

and (H.3) gives

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n(K_a^c) \leq -a.$$

Substituting into (H.5) and rewriting $\max\{-A, -a\} = -\min\{A, a\}$, we obtain

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n(F) \leq - \min\Big\{ \inf_{x \in F \cap K_a} I(x),\ a \Big\}.$$

Because $F \cap K_a \subset F$, we have $\inf_{x\in F} I(x) \leq \inf_{x\in F\cap K_a} I(x)$, hence

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n(F) \leq - \min\Big\{ \inf_{x \in F} I(x),\ a \Big\}.$$

Finally, letting $a \to \infty$ and using $\min\{c, a\} \uparrow c$ for $c \in [0, \infty]$, we conclude

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n(F) \leq - \inf_{x \in F} I(x),$$

which is (H.4). Since the lower bound (H.1) is identical in the weak and full LDP definitions, the full LDP follows. We turn to part (2).

Part (2): Fix $a > 0$ and take $K_a = X$, which is compact. Then $K_a^c = \varnothing$ and hence $P_n(K_a^c) = 0$, so $\frac{1}{n} \log P_n(K_a^c) = -\infty$ for all n, implying $\limsup_{n\to\infty} \frac{1}{n} \log P_n(K_a^c) = -\infty \leq -a$. This proves exponential tightness. The final statement follows by the first part. We turn to part (3).

Part (3): Fix $a < \infty$ and define $K_a \coloneqq \{x \in X :\ I(x) \leq a\}$. Since I is lower semicontinuous, the set K_a is closed in X. Because X is compact and closed subsets of compact sets are compact, K_a is compact. Hence I is good. □

Using Lemma H.5 in conjunction with the fact that $\mathcal{M}_c(\widetilde{\chi})$ is compact on a finite state space, we have that the LDP for Z_n derived using Lemma H.3 is **full and the rate function is good**. Our next objective is to study the properties of this rate function. The following Lemma (Proposition 2.7 from Daures) establishes that the rate function is convex when restricted on irreducibility classes.

Proposition H.6. *Let $I^{(2)}$ be the rate function from Lemma H.4 (so $I^{(2)} = -s$ on its effective domain). The following properties hold.*

1. *Let $\boldsymbol{\mu} \in \mathcal{M}(\widetilde{\chi})$ and for $\mathcal{J}_\mu$ denoting the maximal irreducibility class for $\boldsymbol{\mu}$ (where we refer the readers to Meyn and Tweedie [2012] for the definition of maximal irreducibility class), let*

$$F \ \coloneqq\ \big\{\boldsymbol{\nu} \in \mathcal{M}(\widetilde{\chi}) \ :\ \mathcal{J}_{\boldsymbol{\nu}} \subset \mathcal{J}_{\boldsymbol{\mu}}\big\}.$$

   Then $I^{(2)}\,|_F$ is convex and $\forall \lambda_1 \in [0,1], \lambda_2 = 1 - \lambda_1, \boldsymbol{\mu}_1, \boldsymbol{\mu}_2 \in \mathcal{M}(\widetilde{\chi})$

$$I^{(2)}(\lambda_1 \boldsymbol{\mu}_1 + \lambda_2 \boldsymbol{\mu}_2) \ \leq\ \lambda_1 I^{(2)}(\boldsymbol{\mu}_1) + \lambda_2 I^{(2)}(\boldsymbol{\mu}_2).$$

2. *Let $\mathcal{J}_{\boldsymbol{\mu}_1} \cap \mathcal{J}_{\boldsymbol{\mu}_2} = \varnothing$, then*

$$I^{(2)}(\lambda_1 \boldsymbol{\mu}_1 + \lambda_2 \boldsymbol{\mu}_2) \ =\ \lambda_1 I^{(2)}(\boldsymbol{\mu}_1) + \lambda_2 I^{(2)}(\boldsymbol{\mu}_2).$$

3. *Let $\mu \in \mathcal{M}(\widetilde{\chi}) \setminus \mathcal{M}_c(\widetilde{\chi})$. Then*

$$I^{(2)}(\mu) = \infty.$$

4. *For every Borel set $A \subset \mathcal{M}(\widetilde{\chi})$, define*

$$\underline{s}_\infty(A) \ \coloneqq\ \sup\Big\{\underline{s}\big(A \cap \mathcal{M}(K)\big) \ :\ K \subset \widetilde{\chi} \text{ finite}\Big\},$$

   (where $\mathcal{M}(K)$ denotes the set of measures in $\mathcal{M}(\widetilde{\chi})$ supported on K). Let $\mu \in \mathcal{M}(\widetilde{\chi})$. Then the map $r \mapsto \underline{s}_\infty\big(B(\mu, r)\big)$ is nondecreasing and

$$I^{(2)}(\mu) \ =\ -\lim_{r\to 0} \underline{s}_\infty\big(B(\mu, r)\big).$$

We now demonstrate using that any such function must satisfy the following property which is a consequence of Proposition 3.1 in Daures [2025].

Proposition H.7. *Let $\{L_n^{(2)}\}$ be the empirical pair measure with Λ the scaled cumulant generating function as defined in Equation* (E.5) *and $\Lambda(\boldsymbol{\mu})$ is Legendre-Fenchel transform, and the Donsker-Varadhan (DV) entropy functional [Daures, 2025, Eq. (3.4)]. Then the good rate function $I^{(2)}$ for $\{L_n^{(2)}\}$ satisfies*

$$I^{(2)}(\boldsymbol{\mu}) = \begin{cases} J(\boldsymbol{\mu}) & \mu \text{ shift-invariant (balanced)}, \\ \\ \infty, & \text{otherwise}. \end{cases}$$

Our next objective will be to show the final form for the rate function. In other words, we will show that it decomposes into a combination of rate function over the irreducibility classes. Assume that $\widetilde{\chi}$ admits a partition into nonempty disjoint sets

$$\widetilde{\chi} = \bigsqcup_{j \in \mathcal{J}} C_j$$

such that $\boldsymbol{\mu}$ has no cross-block mass, i.e.

$$\mu_{o,p} = 0 \qquad \text{whenever } o \in C_j,\ p \in C_k,\ j \neq k.$$

Set $\alpha_j := \boldsymbol{\mu}(C_j^2)$ and for each j with $\alpha_j > 0$ define the normalized restriction

$$\tilde{\boldsymbol{\mu}}_j \ := \ \alpha_j^{-1}\, \boldsymbol{\mu}\big|_{C_j^2} \ \in \ \mathcal{P}(C_j^2).$$

Lemma H.8. *Let $\widetilde{\chi}$ be a finite set and let $\nu : \widetilde{\chi}^2 \to (0,1]$ be a reference transition kernel. For $\boldsymbol{\mu} \in \mathcal{P}(\widetilde{\chi}^2)$, write $\mu_o := \sum_{p \in \widetilde{\chi}} \mu_{o,p}$ and define*

$$R(\boldsymbol{\mu}) := \mathrm{KL}(\boldsymbol{\mu} \,\|\, \bar{\boldsymbol{\mu}} \otimes p) = \sum_{o,v \in \widetilde{\chi}} \mu_{o,v} \log \frac{\mu_{o,v}}{\mu_o\, \nu(o,v)}$$

with the usual conventions $0 \log 0 := 0$ and $R(\boldsymbol{\mu}) = +\infty$ if $\boldsymbol{\mu} \not\ll \bar{\boldsymbol{\mu}} \otimes \nu$. Then

$$R(\boldsymbol{\mu}) \ = \sum_{j \in \mathcal{J} \colon \alpha_j > 0} \alpha_j\, R(\tilde{\boldsymbol{\mu}}_j),$$

where $R(\tilde{\boldsymbol{\mu}}_j)$ is computed with the same formula as above (viewing $\tilde{\boldsymbol{\mu}}_j$ as a measure on $\widetilde{\chi}^2$ supported on C_j^2).

Proof. Fix $j \in \mathcal{J}$ and let $\alpha_j = \boldsymbol{\mu}(C_j^2)$. By the no-cross-block assumption, for $(o,p) \in C_j^2$ we have $\mu_{o,p} = \alpha_j \tilde{\mu}_{j;o,p}$. Moreover, for $o \in C_j$,

$$\mu_o = \sum_{p \in S} \mu_{o,p} = \sum_{p \in C_j} \mu_{o,p} = \alpha_j \sum_{p \in C_j} \tilde{\mu}_{j;o,p} = \alpha_j\, \tilde{\mu}_{j;o}^{(1)},$$

where $\tilde{\mu}_{j;o}^{(1)} := \sum_{p \in C_j} \tilde{\mu}_{j;o,p}$. Therefore, for $(o,p) \in C_j^2$ with $\mu_o > 0$,

$$\frac{\mu_{o,p}}{\mu_o\, \nu(o,p)} = \frac{\alpha_j \tilde{\mu}_{j;o,p}}{\alpha_j \tilde{\mu}_{j;o}^{(1)}\, \nu(o,p)} = \frac{\tilde{\mu}_{j;o,p}}{\tilde{\mu}_{j;o}^{(1)}\, \nu(o,p)}.$$

Now split the defining sum for $R(\boldsymbol{\mu})$ over the disjoint blocks:

$$\begin{aligned} R(\boldsymbol{\mu}) &= \sum_{j \in \mathcal{J}} \sum_{(o,p) \in C_j^2} \mu_{o,p} \log \frac{\mu_{o,p}}{\mu_o\, \nu(o,p)} \\ &= \sum_{j \in \mathcal{J} \colon \alpha_j > 0} \sum_{(o,p) \in C_j^2} \alpha_j \tilde{\mu}_{j;o,p} \log \frac{\tilde{\mu}_{j;o,p}}{\tilde{\mu}_{j;o}^{(1)}\, \nu(o,p)} \\ &= \sum_{j \in \mathcal{J} \colon \alpha_j > 0} \alpha_j \sum_{(o,p) \in C_j^2} \tilde{\mu}_{j;o,p} \log \frac{\tilde{\mu}_{j;o,p}}{\tilde{\mu}_{j;o}^{(1)}\, \nu(o,p)} \\ &= \sum_{j \in \mathcal{J} \colon \alpha_j > 0} \alpha_j\, R(\tilde{\boldsymbol{\mu}}_j), \end{aligned}$$

where we used $0 \log 0 = 0$ and the fact that the ratio above is unchanged by the normalization factor α_j. □

A combinations of the previous results directly yields us the proof of Theorem G.1.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The paper theoretically justifies the claims made.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: Mentioned in the conclusions.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

Justification: Present in the sections 3 and 4.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [N/A]

   Justification: No experiments are present.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [N/A]

Justification: No experiments are present.

Guidelines:

- The answer [N/A] means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [N/A]

   Justification: No experiments are present.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [N/A]

   Justification: No experiments are present.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.

- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [N/A]

   Justification: No experiments are present.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification:

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [N/A]

    Justification: This work is primarily theoretical.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.
    - If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.

- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: This work is mostly theoretical.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [N/A]

    Justification: Does not use existing assets.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.

- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: There is no new assets.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: No human subjects appear.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: No human subjects appear.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

Answer: [N/A]

Justification: LLMs are only used for writing, editing, and formatting purposes.

Guidelines:

- The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.