ALEA, Lat. Am. J. Probab. Math. Stat. **22**, 925–952 (2025)
DOI: 10.30757/ALEA.v22-37

Markov chain entropy games and the geometry of their Nash equilibria

Michael C.H. Choi and Geoffrey Wolfer

Department of Statistics and Data Science, National University of Singapore, Singapore
E-mail address: mchchoi@nus.edu.sg
URL: http://mchchoi.github.io

Center for Data Science, Waseda University, Tokyo, Japan
E-mail address: geo.wolfer@aoni.waseda.jp
URL: https://geo-wolfer.gitlab.io

Abstract. We introduce and study a two-player zero-sum game between a probabilist and Nature defined by a convex function f, a finite collection $\mathcal{B}$ of Markov generators (or its convex hull), and a target distribution π. The probabilist selects a mixed strategy $\mu \in \mathcal{P}(\mathcal{B})$, the set of probability measures on $\mathcal{B}$, while Nature adopts a pure strategy and selects a π-reversible Markov generator M. The probabilist receives a payoff equal to the f-divergence $D_f(M\|L)$, where L is drawn according to μ. We prove that this game always admits a mixed strategy Nash equilibrium and satisfies a minimax identity. In contrast, a pure strategy equilibrium may fail to exist. We develop a projected subgradient method to compute approximate mixed strategy equilibria with provable convergence guarantees. Connections to information centroids, Chebyshev centers, and Bayes risk are discussed. This paper extends earlier minimax results on f-divergences such as (Haussler, 1997; Haussler and Opper, 1997; Gushchin and Zhdanov, 2006) to the context of Markov generators.

1. Introduction

This paper studies a two-player zero-sum game between a probabilist and Nature, defined by parameters $(f, \mathcal{B}, \pi)$. These will be introduced formally in Sections 2 and 3, but we outline them here to motivate the setup. The function f is convex and satisfies standard regularity conditions; $\mathcal{B}$ is either a finite set $\{L_i\}_{i=1}^n$ or its convex hull, with each L_i a Markov generator on a finite state space $\mathcal{X}$; and π is a given target distribution on $\mathcal{X}$. We analyze two versions of the game: one with pure strategies, and one with mixed strategies. In the mixed strategy setting, the probabilist selects

Received by the editors October 14th, 2024; accepted July 31st, 2025.

2010 *Mathematics Subject Classification.* 49J35, 60J27, 60J28, 62B10, 62C20, 90C47, 91A05, 91A68, 94A17, 94A29.

Key words and phrases. Markov chains, f-divergence, information geometry, information centroid, saddle point, Nash equilibrium, minimax theorem, Chebyshev center, Bayes risk, subgradient, algorithmic game theory.

Michael Choi acknowledges the financial support of the projects A-8001061-00-00, NUSREC-HPC-00001, NUSREC-CLD-00001, A-0000178-01-00, A-0000178-02-00 and A-8003574-00-00 at National University of Singapore. Part of this research was supported when Geoffrey Wolfer was part of the Special Postdoctoral Researcher Program (SPDR) of RIKEN and by the Japan Society for the Promotion of Science KAKENHI under Grant 23K13024.

a probability measure $\mu \in \mathcal{P}(\mathcal{B})$ and samples $L \in \mathcal{B}$ according to μ, while Nature adopts a pure strategy and chooses a reversible Markov generator $M \in \mathcal{L}(\pi)$. The probabilist receives a payoff of $D_f(M\|L)$, the f-divergence from L to M. In the pure strategy variant, both players choose deterministically from their respective strategy sets.

Games of a similar flavor appear in the literature under the name "statistician versus Nature" problems (Haussler, 1997; Haussler and Opper, 1997; Gushchin and Zhdanov, 2006), where the statistician pays a loss based on divergence between the estimator and the ground truth distribution chosen by Nature. Our game reverses this direction: the probabilist receives a gain from Nature. This role reversal is motivated by differences between reversible and non-reversible Markov chains as follows:

Reversibility is fundamental in many physical systems—quantum and classical mechanics are both time-reversible (Schrödinger, 1931). Accordingly, we model Nature as selecting from the set of π-reversible generators. This also reflects common practice in modern Markov chain Monte Carlo methods, where algorithms such as Metropolis-Hastings and overdamped Langevin diffusions are π-reversible.

In contrast, non-reversible chains have attracted recent interest for their potential to accelerate convergence. Algorithms of this type appear in various works (e.g. Hwang et al. (1993, 2005); Rosenthal and Rosenthal (2015); Bierkens (2016); Diaconis et al. (2000); Rey-Bellet and Spiliopoulos (2016); Gustafson (1998); Kamatani and Song (2023)). We therefore let the probabilist choose from a possibly non-reversible inventory of Markov generators $\mathcal{B}$. The quantity $D_f(M\|L)$ can be interpreted as an information-theoretic advantage or performance gain due to non-reversibility.

While the above explanations justify the role of the probabilist and Nature in the entropy game, from a mathematical point of view however, these two roles can be safely interchanged, or one may wish to replace the term "probabilist" by "Player A" and "Nature" by "Player B" throughout the entire manuscript.

This setting raises several natural questions: Does a Nash equilibrium exist (Karlin and Peres, 2017; Maschler et al., 2020)? Is it unique? Can it be computed efficiently? What is the game's value? Is there a sequential analogue?

While we do not aim to answer all of these questions, this paper focuses on several foundational aspects: the existence and uniqueness of an equilibrium, cases where equilibria may fail to exist, and an optimization algorithm to approximate them efficiently. We also explore connections with the Bayes risk and geometric notions such as information centroids and Chebyshev centers.

Further motivation comes from a question raised by Laurent Miclo about whether the Metropolis-type generators

$$P_{-\infty}(x,y) := \min\{L(x,y), L_\pi(x,y)\}, \quad P_\infty(x,y) := \max\{L(x,y), L_\pi(x,y)\}, \quad x \neq y,$$

can be interpreted game-theoretically. Here, L_π is the π-dual of L, and diagonal terms are set so that each row sums to zero (see Section 2). While $P_{-\infty}$ corresponds to the classical continuous-time Metropolis-Hastings generator, P_∞ has been studied in Diaconis and Miclo (2009); Choi (2020); Choi and Huang (2020). Since both involve pointwise minimum or maximum operations, a game-theoretic interpretation is natural. As we show in Example 3.12, when D_f is taken to be the total variation distance, both $P_{-\infty}$, P_∞, and their convex combinations can be understood as part of a mixed strategy Nash equilibrium in the entropy game. This gives a possible answer to Miclo's question.

We summarize our main contributions as follows:

(1) We introduce the two-person Markov chain entropy games of the probabilist against Nature. This naturally generalizes the game of the statistican against Nature (Haussler, 1997; Haussler and Opper, 1997; Gushchin and Zhdanov, 2006), where instead of Markov generators, the game in these references involves probability measures.

We draw a perhaps surprising connection between reversiblizations of Markov generators, widely studied in modern theory (Miclo, 1997; Fill, 1991; Paulin, 2015; Choi, 2020), game-theoretic concepts, and geometry.

(2) We introduce the notion of a weighted information centroid in the context of Markov generators and highlight its central role in analyzing Nash equilibria of entropy games. This naturally generalizes the notion of an information centroid of Markov chains introduced by Choi and Wolfer (2024). Our analysis shows that the (mixed strategy) Nash equilibrium is intimately related to the notion of Chebyshev center, which can be interpreted as a specific weighted information centroid. This important observation allows us to analyze the existence and uniqueness of a Nash equilibrium in the game.

(3) We propose and analyze a simple projected subgradient algorithm to find an approximate Nash equilibrium with a provable convergence rate of $\mathcal{O}(1/\sqrt{t})$. A central question in game theory, in particular algorithmic game theory, lies in developing efficient algorithms for (approximate) computation of the Nash equilibrium. To this end, we propose a simple and easy-to-implement projected subgradient algorithm which takes advantage of the information geometry of the underlying Markov generators.

The rest of this paper is organized as follows. In Section 2, we introduce and recall various notions, in particular the notion of weighted information centroid, which plays a central role in our subsequent investigation. We proceed to present the main results of the paper, where we first address fundamental questions concerning Chebyshev centers, weighted information centroids and minimax values in Section 3, followed by a pure strategy game-theoretic analysis in Section 3.1. The corresponding setting of mixed strategy game is presented in Section 3.2, and several simple yet illustrative examples are given in Section 3.1.1 and Section 3.2.1. The design and analysis of a novel projected subgradient algorithm to find an approximate mixed strategy Nash equilibrium is stated in Section 3.3. Finally, the proofs of the main results are presented in Section 4.

1.1. *Notations.* In this subsection we introduce some commonly used notations throughout the manuscript. For $a, b \in \mathbb{Z}$, we write $[\![a, b]\!] := \{a, a+1, \ldots, b-1, b\}$ and $[\![n]\!] := [\![1, n]\!]$ for $n \in \mathbb{N}$. We denote by $\mathbf{0}$ the all-zeros matrix on the finite state space $\mathcal{X}$. For a given function $g(n)$, we say that it is $\mathcal{O}(h(n))$ if there exist constants $C > 0$ and n_0 such that $g(n) \leq Ch(n)$ for all $n \geq n_0$.

2. Preliminaries

2.1. *Weighted information centroids.* Consider a convex function $f : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying $f(1) = 0$, and a given positive discrete distribution π on a finite state space $\mathcal{X}$, where $\mathbb{R}_+$ is the set of non-negative real numbers. Define $\mathcal{L}$ as the set of Markov infinitesimal generators on $\mathcal{X}$. These generators correspond to $\mathcal{X} \times \mathcal{X}$ matrices having non-negative off-diagonal elements and zero row sums. A generator L is said to be π-stationary if $\pi L = 0$. Moreover, $\mathcal{L}(\pi) \subset \mathcal{L}$ is the subset of π-reversible generators, where we recall that a generator L is said to be π-reversible if it satisfies the detailed balance condition, that is, for all $x \neq y \in \mathcal{X}$, we have $\pi(x)L(x, y) = \pi(y)L(y, x)$.

In view of Jansen and Kurt (2014, Proposition 1.2), we define the π-dual of a generator $L \in \mathcal{L}$ to be L_π. For $x \neq y$, the off-diagonal elements of L_π are given by

$$L_\pi(x, y) = \frac{\pi(y)}{\pi(x)} L(y, x),$$

and the diagonal ones ensure zero row sums for all rows. In the special case when L has π as its unique stationary distribution, we then have $L_\pi = L^*$, the adjoint of L in $\ell^2(\pi)$ or its time-reversal. Here the space $\ell^2(\pi)$ is the standard weighted ℓ^2 Hilbert space endowed with the inner product

$\langle \cdot, \cdot \rangle_\pi$ given by, for any functions $g, h : \mathcal{X} \to \mathbb{R}$,

$$\langle g, h \rangle_\pi := \sum_{x \in \mathcal{X}} g(x) h(x) \pi(x).$$

Following Diaconis and Miclo (2009), for a given target π and Markov infinitesimal generators $M, L \in \mathcal{L}$, the f-divergence from L to M with respect to π is defined as

$$D_f(M||L) = \sum_{x \in \mathcal{X}} \pi(x) \sum_{y \in \mathcal{X} \setminus \{x\}} L(x,y) f\left(\frac{M(x,y)}{L(x,y)} \right), \tag{2.1}$$

where the convention of $0 f(a/0) := 0$ for $a \geq 0$ applies. If f^* is the convex conjugate of f, defined by $f^*(t) = t f(1/t)$ for $t > 0$, then

$$D_f(M||L) = D_{f^*}(L||M),$$

and $f^*(1) = 0$. In particular, when f is self-conjugate, that is $f^* = f$, the f-divergence in (2.1) is symmetric.

For a general generator L not necessarily with π as its stationary distribution, we consider projecting L onto $\mathcal{L}(\pi)$ under f-divergence D_f. Following Wolfer and Watanabe (2021); Choi and Wolfer (2024), the notions of f-projection and f^*-projection are defined as:

$$M^f = M^f(L, \pi) := \operatorname*{arg\,min}_{M \in \mathcal{L}(\pi)} D_f(M||L), \quad M^{f^*} = M^{f^*}(L, \pi) = \operatorname*{arg\,min}_{M \in \mathcal{L}(\pi)} D_f(L||M). \tag{2.2}$$

To state our main results, we now introduce the notion of a weighted information centroid of Markov chains. It generalizes the notion of an (unweighted) information centroid of Markov chains introduced in Choi and Wolfer (2024). Let $\mathbf{w} = (w_1, \ldots, w_n) \in \mathbb{R}_+^n$ be a weight vector in the probability simplex $\mathcal{S}_n$ of n elements, that is,

$$\mathcal{S}_n := \left\{ \mathbf{w} = (w_1, \ldots, w_n) \in \mathbb{R}_+^n ; \ \sum_{i=1}^n w_i = 1 \right\}. \tag{2.3}$$

Note that we denote the simplex by $\mathcal{S}_n$ instead of $\mathcal{P}(\llbracket n \rrbracket)$ throughout the manuscript. Given a sequence of Markov generators $\{L_i\}_{i=1}^n$, where $L_i \in \mathcal{L}$ for $i \in \llbracket n \rrbracket$, we define the notions of $\mathbf{w}$-weighted f-projection centroid and f^*-projection centroid to be respectively

$$M_n^f = M_n^f(\mathbf{w}, \{L_i\}_{i=1}^n, \pi) := \operatorname*{arg\,min}_{M \in \mathcal{L}(\pi)} \sum_{i=1}^n w_i D_f(M||L_i),$$

$$M_n^{f^*} = M_n^{f^*}(\mathbf{w}, \{L_i\}_{i=1}^n, \pi) = \operatorname*{arg\,min}_{M \in \mathcal{L}(\pi)} \sum_{i=1}^n w_i D_f(L_i||M).$$

We consider two important special cases: first in the case with $n = 1$, the above notions reduce to $M_1^f = M^f$ and $M_1^{f^*} = M^{f^*}$ respectively as introduced in (2.2). In the second special case, let $n \in \mathbb{N}$ and $\{\mathbf{e}_i\}_{i=1}^n$ be the standard unit vectors. It is immediate that $M_n^f(\mathbf{e}_i, \{L_i\}_{i=1}^n, \pi) = M^f(L_i, \pi)$ and $M_n^{f^*}(\mathbf{e}_i, \{L_i\}_{i=1}^n, \pi) = M^{f^*}(L_i, \pi)$.

Our first main result in this Section establishes existence and uniqueness of f and f^*-projection centroids under strict convexity of f, and its proof is given in Section 4.1.

Theorem 2.1 (Existence and uniqueness of weighted f and f^*-projection centroids)**.** *Suppose we are given a sequence of Markov generators $\{L_i\}_{i=1}^n$, where $L_i \in \mathcal{L}$ for $i \in \llbracket n \rrbracket$, a f-divergence D_f generated by a strictly convex f which is assumed to have a derivative at 1 given by $f'(1) = 0$, and a weight vector $\mathbf{w} \in \mathcal{S}_n$. We further assume that there exists at least one $L_i \neq \mathbf{0}$ and at least one*

$w_i > 0$ when $L_i \neq \mathbf{0}$. A weighted f-projection of D_f (resp. f^*-projection of D_{f^*}) that minimizes the mapping

$$\mathcal{L}(\pi) \ni M \mapsto \sum_{i=1}^n w_i D_f(M||L_i) \quad \left(\text{resp. } = \sum_{i=1}^n w_i D_{f^*}(L_i||M)\right)$$

uniquely exists and is denoted by M_n^f.

Remark 2.1. By replacing f by f^* in the main result of Theorem 2.1, we can analogously define $M_n^{f^*}$. Precisely, a weighted f^*-projection of D_f (resp. f-projection of D_{f^*}) that minimizes the mapping

$$\mathcal{L}(\pi) \ni M \mapsto \sum_{i=1}^n w_i D_f(L_i||M) \quad \left(\text{resp. } = \sum_{i=1}^n w_i D_{f^*}(M||L_i)\right)$$

uniquely exists and is denoted by $M_n^{f^*}$.

Remark 2.2 (On the convexity of f). We emphasize that throughout this manuscript, unless otherwise specified, f is assumed to be a convex function rather than a strictly convex function. This yields the following interesting consequences. Suppose that f is a convex function with $f(s) = 0$ for some $s > 0, s \neq 1$. For a given $L \in \mathcal{L}$, we thus have

$$D_f(sL||L) = 0,$$

and hence the two generators sL and L are indistinguishable with respect to D_f. From a transition semigroup perspective, this is justifiable since $P^t := e^{sLt}$, the transition semigroup generated by sL, is merely a time-change of $Q^t := e^{Lt}$, the transition semigroup generated by L, for $t \geq 0$.

Remark 2.3. We discuss the importance of the additional assumptions in Theorem 2.1 here. In the first case, if $L_i = \mathbf{0}$ for all $i \in [\![n]\!]$, then obviously for any $M \in \mathcal{L}(\pi)$ we have $D_f(M||L_i) = 0$ and hence

$$\sum_{i=1}^n w_i D_f(M||L_i) = 0.$$

As such, this shows the existence of weighted f-projection centroids which are not unique even if f is strictly convex. In the second case, suppose that there exists at least one $L_i \neq \mathbf{0}$ and $w_i = 0$ whenever $L_i \neq \mathbf{0}$. This again leads to, for any $M \in \mathcal{L}(\pi)$,

$$\sum_{i=1}^n w_i D_f(M||L_i) = 0.$$

We thus exclude the above two degenerate cases in Theorem 2.1.

Remark 2.4. We stress that in Theorem 2.1, the given positive distribution π is arbitrary and at the same time the generators L_i need not admit π as a stationary distribution or even be irreducible in the first place. The purpose of Theorem 2.1 is to consider the joint projection of these generators L_i onto the space $\mathcal{L}(\pi)$.

In the second result of this Section, we explicitly calculate the weighted f and f^*-projection centroids M_n^f and $M_n^{f^*}$ under some common f-divergences. The proof is deferred to Section 4.2. For these common choices of f, as shown in Choi and Wolfer (2024), the power mean reversibilizations P_p appear naturally as the corresponding f or f^*-projection. For $x \neq y \in \mathcal{X}$ and $p \in \mathbb{R} \setminus \{0\}$, the off-diagonal entries of P_p are given by

$$P_p(x,y) := \left(\frac{L(x,y)^p + L_\pi(x,y)^p}{2}\right)^{1/p}, \tag{2.4}$$

where diagonal entries ensure zero row sums. The limiting cases are defined to be

$$P_0(x,y) := \lim_{p\to 0} P_p(x,y) = \sqrt{L(x,y)L_\pi(x,y)},$$
$$P_\infty(x,y) := \lim_{p\to \infty} P_p(x,y) = \max\{L(x,y), L_\pi(x,y)\}, \tag{2.5}$$
$$P_{-\infty}(x,y) := \lim_{p\to -\infty} P_p(x,y) = \min\{L(x,y), L_\pi(x,y)\}. \tag{2.6}$$

Note that in the special case when the weights are given by $w_i = 1/n$ for all $i \in [\![n]\!]$, we recover the results in Choi and Wolfer (2024), as expected.

Theorem 2.2 (Examples of weighted f and f^*-projection centroids)**.** *Given a sequence of Markov generators $(L_i)_{i=1}^n$, where $L_i \in \mathcal{L}$ for $i \in [\![n]\!]$ and a weight vector $\mathbf{w} \in \mathcal{S}_n$. We further assume that there exists at least one $L_i \neq \mathbf{0}$ and at least one $w_i > 0$ when $L_i \neq \mathbf{0}$. Recall the power mean reversiblizations P_p in (2.4).*

(1) *(weighted f and f^*-projection centroids under α-divergence) Let $f(t) = \frac{t^\alpha - \alpha t - (1-\alpha)}{\alpha(\alpha-1)}$ for $\alpha \in \mathbb{R}\backslash\{0,1\}$. The unique f-projection centroid M_n^f is given by, for $x \neq y \in \mathcal{X}$,*

$$M_n^f(x,y) = \left(\sum_{i=1}^n w_i \left(M^f(L_i,\pi)(x,y)\right)^{1-\alpha}\right)^{1/(1-\alpha)},$$

while the unique f^*-projection centroid $M_n^{f^*}$ is given by, for $x \neq y \in \mathcal{X}$,

$$M_n^{f^*}(x,y) = \left(\sum_{i=1}^n w_i \left(M^{f^*}(L_i,\pi)(x,y)\right)^{\alpha}\right)^{1/\alpha},$$

where M^f, M^{f^*} are respectively the $P_{1-\alpha}, P_\alpha$-reversiblization.

(2) *(weighted f and f^*-projection centroids under squared Hellinger distance) Let $f(t) = (\sqrt{t}-1)^2$. The unique f-projection centroid M_n^f is given by, for $x \neq y \in \mathcal{X}$,*

$$M_n^f(x,y) = \left(\sum_{i=1}^n w_i \sqrt{M^f(L_i,\pi)(x,y)}\right)^2, \tag{2.7}$$

while the unique f^*-projection centroid $M_n^{f^*}$ is given by, for $x \neq y \in \mathcal{X}$,

$$M_n^{f^*}(x,y) = \left(\sum_{i=1}^n w_i \sqrt{M^{f^*}(L_i,\pi)(x,y)}\right)^2,$$

where $M^{f^*} = M^f$ is the $P_{1/2}$-reversiblization.

(3) *(f and f^*-projection centroids under Kullback-Leibler divergence) Let $f(t) = t\ln t - t + 1$. The unique f-projection centroid M_n^f is given by, for $x \neq y \in \mathcal{X}$,*

$$M_n^f(x,y) = \prod_{i=1}^n \left(M^f(L_i,\pi)(x,y)\right)^{w_i},$$

where $0^0 := 0$ in the expression above, while the unique f^*-projection centroid $M_n^{f^*}$ is given by, for $x \neq y \in \mathcal{X}$,

$$M_n^{f^*}(x,y) = \sum_{i=1}^n w_i M^{f^*}(L_i,\pi)(x,y),$$

where M^f, M^{f^*} are respectively the P_0, P_1-reversiblization.

Remark 2.5. Item (2) can be considered as a special case of item (1) by taking $\alpha = 1/2$. Another important special case of α-divergence is the χ^2-divergence where we take $\alpha = 2$.

2.2. *Two-person pure strategy and mixed strategy games.* This subsection reviews some classical game-theoretic notions in the context of the probabilist against Nature game. Much of the exposition in this subsection are drawn from content in Boyd and Vandenberghe (2004, Section 5.2.5 and 5.4.3) and Karlin and Peres (2017, Chapter 2). Consider the following two-person non-cooperative zero-sum pure strategy game with respect to the parameters $(f, \mathcal{B}, \pi)$. If Nature chooses $M \in \mathcal{L}(\pi)$, the *strategy set of Nature*, while the probabilist chooses $L \in \mathcal{B}$, the *strategy set of the probabilist*, then Nature pays an amount $D_f(M||L)$ to the probabilist. As a result, Nature wants to minimize D_f while the probabilist seeks to maximize D_f. The quantity $D_f(M||L)$ is known as the *payoff function* of the game, and it is said to be a *pure strategy* game as both players are only allowed to choose deterministically from their respective strategy sets. On the other hand, a game is said to be a *mixed strategy* game if the players are allowed to choose randomly within their respective strategy sets. In the sequel, we shall consider $\mathcal{B}$, the strategy set of the probabilist, to be either $\{L_i\}_{i=1}^n$ or its convex hull $\operatorname{conv}((L_i)_{i=1}^n)$. From now on we shall call this a two-person pure strategy game.

Suppose that Nature chooses its strategy M first, followed by the probabilist, who has the knowledge of the choice of Nature. The probabilist thus seeks to choose $L \in \mathcal{B}$ to maximize the payoff $D_f(M||L)$. The resulting payoff is $\sup_{L\in\mathcal{B}} D_f(M||L)$, which depends on M, the choice of Nature. Nature assumes that the probabilist will choose this strategy, and hence Nature seeks to minimize the worst-case payoff, that is, Nature chooses a strategy in the set

$$\arg\min_{M\in\mathcal{L}(\pi)} \sup_{L\in\mathcal{B}} D_f(M||L),$$

known as a *minimax strategy of Nature*, which yields

$$\overline{v}^f = \overline{v}^f(\mathcal{B}, \pi) := \inf_{M\in\mathcal{L}(\pi)} \sup_{L\in\mathcal{B}} D_f(M||L),$$

the payoff from Nature to the probabilist. $\overline{v}^f$ is also known as the *minimax value* of the game.

Now, suppose that the order of play is reversed: the probabilist chooses $L \in \mathcal{B}$ first, and Nature, with the knowledge of L, picks from the strategy set $\mathcal{L}(\pi)$. Following a similar argument as before, if both players follow the optimal strategy, the probabilist chooses a strategy in the set

$$\arg\sup_{L\in\mathcal{B}} \inf_{M\in\mathcal{L}(\pi)} D_f(M||L),$$

known as a *maximin strategy of the probabilist*, which yields

$$\underline{v}^f = \underline{v}^f(\mathcal{B}, \pi) := \sup_{L\in\mathcal{B}} \inf_{M\in\mathcal{L}(\pi)} D_f(M||L),$$

the payoff from Nature to the probabilist. $\underline{v}^f$ is known as the *maximin value* of the game.

If the minimax equality holds, that is, the minimax and maximin value coincides with

$$\overline{v}^f = \underline{v}^f =: v = v(f, \mathcal{B}, \pi),$$

then v is the *value of the pure strategy game*. Let M^f and L^f be respectively a minimax strategy of Nature and a maximin strategy of the probabilist. The pair (M^f, L^f) is called a *pure strategy Nash equilibrium* with respect to $(f, \mathcal{B}, \pi)$. M^f is said to be an optimal strategy of Nature while L^f is said to be an optimal strategy of the probabilist.

The above pure strategy game can be generalized to a situation where the probabilist uses a mixed strategy while Nature maintains a pure strategy. From now on we shall refer this to a two-person mixed strategy game with respect to the parameters $(f, \mathcal{B}, \pi)$. In this mixed strategy game, the probabilist first chooses a *prior measure* $\mu \in \mathcal{P}(\mathcal{B})$ and picks a $L \in \mathcal{B}$ at random according to μ. On the other hand Nature, knowing μ but not knowing L, follows a pure strategy to choose $M \in \mathcal{L}(\pi)$. Afterwards, Nature pays an amount $D_f(M||L)$ to the probabilist. This serves as a natural generalization of the game of a statistican against Nature proposed and developed in Haussler (1997); Haussler and Opper (1997); Gushchin and Zhdanov (2006).

With these notions in mind, we now introduce the minimax and maximin value of this game:

Definition 2.3 (Minimax and maximin values in the two-person mixed strategy game)**.** Consider the two-person mixed strategy game. The *minimax value* and the *maximin value* of this game are defined respectively to be

$$\overline{V}^f = \overline{V}^f(\mathcal{B}, \pi) := \inf_{M \in \mathcal{L}(\pi)} \sup_{\mu \in \mathcal{P}(\mathcal{B})} \int_{\mathcal{B}} D_f(M||L)\, \mu(dL), \tag{2.8}$$

$$\underline{V}^f = \underline{V}^f(\mathcal{B}, \pi) := \sup_{\mu \in \mathcal{P}(\mathcal{B})} \inf_{M \in \mathcal{L}(\pi)} \int_{\mathcal{B}} D_f(M||L)\, \mu(dL). \tag{2.9}$$

Note that $\overline{V}^f \geq \underline{V}^f$. We can define analogously $\overline{V}^{f^*}, \underline{V}^{f^*}$ by replacing f by f^* above.

Next, we define the notions of Bayes risk and Bayes strategy of Nature:

Definition 2.4 (Bayes risk and Bayes strategy)**.** Given a $\mu \in \mathcal{P}(\mathcal{B})$, we say that a Markov generator $M_\mu \in \mathcal{L}(\pi)$ is a *Bayes strategy* with respect to μ if the mapping $M \mapsto \int_{\mathcal{B}} D_f(M||L)\, \mu(dL)$ attains its infimum at M_μ, that is,

$$\inf_{M \in \mathcal{L}(\pi)} \int_{\mathcal{B}} D_f(M||L)\, \mu(dL) = \int_{\mathcal{B}} D_f(M_\mu||L)\, \mu(dL).$$

The minimum value $\int_{\mathcal{B}} D_f(M_\mu||L)\, \mu(dL)$ is said to be the *Bayes risk* with respect to μ,

Finally, we give the notions of minimax and maximin strategy of this game:

Definition 2.5 (Minimax and maximin strategies in the two-person mixed strategy game)**.** Consider the two-person mixed strategy game. A strategy in the set

$$\arg\inf_{M \in \mathcal{L}(\pi)} \sup_{\mu \in \mathcal{P}(\mathcal{B})} \int_{\mathcal{B}} D_f(M||L)\, \mu(dL)$$

is known as a *minimax strategy of Nature*, while a strategy in the set

$$\arg\sup_{\mu \in \mathcal{P}(\mathcal{B})} \inf_{M \in \mathcal{L}(\pi)} \int_{\mathcal{B}} D_f(M||L)\, \mu(dL)$$

is known as a *maximin strategy of the probabilist*.

If the minimax equality holds, that is,

$$\overline{V}^f(\mathcal{B}, \pi) = \underline{V}^f(\mathcal{B}, \pi) =: V = V(f, \mathcal{B}, \pi),$$

then we say that V is the *value of the mixed strategy game*. The pair of minimax strategy of Nature and maximin strategy of the probabilist is a *mixed strategy Nash equilibrium* with respect to the parameters $(f, \mathcal{B}, \pi)$.

3. Main results

In this Section, we gradually introduce and state the main results of this paper. We first state two well-known and useful max-min inequalities for D_f and D_{f^*}. Let $\mathcal{A}, \mathcal{B} \subseteq \mathcal{L}$. We have

$$\sup_{L \in \mathcal{B}} \inf_{M \in \mathcal{A}} D_f(M||L) \leq \inf_{M \in \mathcal{A}} \sup_{L \in \mathcal{B}} D_f(M||L), \tag{3.1}$$

$$\sup_{L \in \mathcal{B}} \inf_{M \in \mathcal{A}} D_f(L||M) \leq \inf_{M \in \mathcal{A}} \sup_{L \in \mathcal{B}} D_f(L||M). \tag{3.2}$$

In the following, we are primarily interested in projection onto the space $\mathcal{L}(\pi)$, and hence we shall take $\mathcal{A} = \mathcal{L}(\pi)$. We now illustrate the relationship between various minimax and maximin values and saddle point:

Definition 3.1 (Minimax and maximin values, saddle point property and saddle point)**.** Let $\mathcal{B} \subseteq \mathcal{L}$. Recall that

$$\overline{v}^f = \overline{v}^f(\mathcal{B}, \pi) = \inf_{M \in \mathcal{L}(\pi)} \sup_{L \in \mathcal{B}} D_f(M||L),$$
$$\underline{v}^f = \underline{v}^f(\mathcal{B}, \pi) = \sup_{L \in \mathcal{B}} \inf_{M \in \mathcal{L}(\pi)} D_f(M||L).$$

We say that $\overline{v}^f$ is the minimax value with respect to $(f, \mathcal{B}, \pi)$, and similarly $\underline{v}^f$ is the maximin value with respect to $(f, \mathcal{B}, \pi)$. In view of (3.1) and (3.2), we have

$$\overline{v}^f \geq \underline{v}^f, \quad \overline{v}^{f^*} \geq \underline{v}^{f^*}.$$

If the equality holds, that is,

$$\overline{v}^f = \underline{v}^f,$$

we say that the saddle point property is satisfied with respect to $(f, \mathcal{B}, \pi)$. In this case, suppose that the pair $(M^f, L^f) \in \mathcal{L}(\pi) \times \mathcal{L}$ attains the optimal value, that is,

$$\overline{v}^f = D_f(M^f||L^f),$$

then we say that (M^f, L^f) is a saddle point with respect to $(f, \mathcal{B}, \pi)$.

Let us now assume that we are given a set of Markov generators $\{L_i\}_{i=1}^n$, where $L_i \in \mathcal{L}$ for $i \in [\![n]\!]$, and in the sequel we shall investigate either $\mathcal{B} = \{L_i\}_{i=1}^n$ or the convex hull of $\{L_i\}_{i=1}^n$. Consider the minimax problem

$$\inf_{M \in \mathcal{L}(\pi)} \max_{i \in [\![n]\!]} D_f(M||L_i). \tag{3.3}$$

As $L \mapsto D_f(M||L)$ is convex for fixed M and pointwise maximum of convex functions preserves convexity (Boyd and Vandenberghe, 2004, Section 3.2.3), the mapping $M \mapsto \max_{i \in [\![n]\!]} D_f(M||L_i)$ is thus convex. As such the outer minimization is a convex minimization problem over the convex set $\mathcal{L}(\pi)$. Inspired by the reformulation of an analogous minimax problem in the context of probability measures in Candan (2020, equation (3)), the minimax problem (3.3) can be equivalently casted as the following constrained convex minimization:

$$\begin{aligned} \min_{M \in \mathcal{L}(\pi), r} \quad & r \\ \text{s.t.} \quad & D_f(M||L_i) \leq r, \text{ for all } i \in [\![n]\!]. \end{aligned} \tag{3.4}$$

This problem can be interpreted geometrically as the Chebyshev center problem in the context of Markov chains. For a given $M \in \mathcal{L}(\pi)$ and $r \geq 0$, the set $\{L \in \mathcal{L};\ D_f(M||L) \leq r\}$ can be understood as the ball of Markov generators within radius r of the center M. The constraint in (3.4) thus entails that this ball encloses the set $\{L_i\}_{i=1}^n$. As the minimization is with respect to the radius r, a so-called Chebyshev center is a center of the minimum radius ball that contains $\{L_i\}_{i=1}^n$, and its optimal value is referred as the Chebyshev radius. Precisely we define the Chebyshev center and radius as follows:

Definition 3.2 (Chebyshev center and radius)**.** Consider the minimax problem (3.3) and its reformulation (3.4). An optimizer

$$\arg\min_{M \in \mathcal{L}(\pi)} \max_{i \in [\![n]\!]} D_f(M||L_i)$$

is called a Chebyshev center with respect to $(f, \{L_i\}_{i=1}^n, \pi)$, and the minimax value $\overline{v}^f$ is called the Chebyshev radius with respect to $(f, \{L_i\}_{i=1}^n, \pi)$.

For the problem (3.4), we denote the Lagrangian function $\mathsf{L} : \mathbb{R}_+ \times \mathcal{L}(\pi) \times \mathbb{R}_+^n \to \mathbb{R}$ to be

$$\mathsf{L}(r, M, \mathbf{w}) := r + \sum_{i=1}^n w_i (D_f(M||L_i) - r),$$

where $\mathbf{w}$ is the associated Lagrange multiplier. Differentiating L with respect to r and setting it to zero gives $\mathbf{w} \in \mathcal{S}_n$, the probability simplex that we introduce in (2.3). The dual problem of (3.4) is now written as

$$\begin{aligned}\max_{\mathbf{w}\in\mathbb{R}_+^n} \min_{r\geq 0, M\in\mathcal{L}(\pi)} \mathsf{L}(r, M, \mathbf{w}) &= \max_{\mathbf{w}\in\mathcal{S}_n} \min_{M\in\mathcal{L}(\pi)} \sum_{i=1}^n w_i D_f(M||L_i) \\ &= \max_{\mathbf{w}\in\mathcal{S}_n} \sum_{i=1}^n w_i D_f(M_n^f||L_i), \end{aligned} \tag{3.5}$$

where we recall that $M_n^f = M_n^f(\mathbf{w}, \{L_i\}_{i=1}^n, \pi)$ is the $\mathbf{w}$-weighted information centroid as in Theorem 2.1 in the second equality. Note that by weak duality we thus have $\overline{v}^f \geq \max_{\mathbf{w}\in\mathcal{S}_n} \sum_{i=1}^n w_i D_f(M_n^f||L_i)$. Our main result below shows that in fact the strong duality holds and connects various notions introduced earlier on. The proof is stated in Section 4.3.

Theorem 3.3.

(1) *(Strong duality holds for (3.4)) The strong duality holds for (3.4) and there exists* $\mathbf{w}^f = \mathbf{w}^f(\{L_i\}_{i=1}^n, \pi) = (w_i^f)_{i=1}^n \in \mathcal{S}_n$ *such that the optimal value for the dual problem (3.5) is attained at* $\mathbf{w}^f$*, that is,*

$$\overline{v}^f = \sum_{i=1}^n w_i^f D_f(M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi)||L_i).$$

The primal problem (3.4) optimal value is attained at $(M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi), \overline{v}^f)$*, and the optimal value of the problem (3.3) is attained at* $(M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi), j)$*, where* $j \in [\![n]\!]$ *is such that* $w_j^f > 0$*.*

(2) *(Chebyshev center is a weighted information centroid* M_n^f*) A given pair* $(M, r) \in \mathcal{L}(\pi) \times \mathbb{R}_+$ *minimizes the primal problem (3.4) and a given* $\mathbf{w} \in \mathcal{S}_n$ *maximizes the dual problem (3.5) if and only if they satisfy the following two conditions:*

(a) *(Complementary slackness) For* $i \in [\![n]\!]$*, we have*

$$D_f(M||L_i) \begin{cases} = r & \text{if } w_i > 0, \\ \leq r & \text{if } w_i = 0. \end{cases}$$

Furthermore this implies $r = \overline{v}^f$*.*

(b) $M = M_n^f(\mathbf{w}, \{L_i\}_{i=1}^n, \pi) \in \arg\min_{M\in\mathcal{L}(\pi)} \mathsf{L}(r, M, \mathbf{w})$.

(3) *(Concavity of the Lagrangian dual) The mapping*

$$\mathcal{S}_n \ni \mathbf{w} = (w_i)_{i=1}^n \mapsto \sum_{i=1}^n w_i D_f(M_n^f(\mathbf{w}, \{L_i\}_{i=1}^n, \pi)||L_i)$$

is concave.

(4) *(Uniqueness of Chebyshev center under strict convexity of* f*) Suppose that the parameters* $(f, \{L_i\}_{i=1}^n, \pi)$ *satisfy the assumptions in Theorem 2.1. If both* $\mathbf{w}_1, \mathbf{w}_2 \in \mathcal{S}_n$ *maximize the dual problem (3.5), we have*

$$M_n^f(\mathbf{w}_1, \{L_i\}_{i=1}^n, \pi) = M_n^f(\mathbf{w}_2, \{L_i\}_{i=1}^n, \pi).$$

In other words, the Chebyshev center is unique.

(5) *(Characterization of* $\overline{v}^f = \underline{v}^f$*) The saddle point property with respect to* $(f, \{L_i\}_{i=1}^n, \pi)$ *(Definition 3.1) holds, that is,*

$$\overline{v}^f = \underline{v}^f$$

if and only if there exists $j \in [\![n]\!]$ *and* $\mathbf{w}^f = (w_i^f)_{i=1}^n$ *a maximizer of the dual problem* (3.5) *with* $w_j^f > 0$ *such that*

$$M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi) \in \operatorname*{arg\,min}_{M \in \mathcal{L}(\pi)} D_f(M||L_j).$$

In particular, if the parameters $(f, \{L_i\}_{i=1}^n, \pi)$ *satisfy the assumptions in Theorem 2.1, then the above statement is equivalent to*

$$M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi) = M^f(L_j, \pi).$$

(6) *(Saddle point) If the saddle point property with respect to* $(f, \{L_i\}_{i=1}^n, \pi)$ *holds, then*

$$(M^f(L_l, \pi), L_l)$$

is a saddle point with respect to $(f, \{L_i\}_{i=1}^n, \pi)$*, where* $l = \arg\max_{i \in [\![n]\!]} D_f(M^f(L_i, \pi)||L_i)$.

Remark 3.1. The concavity of the Lagrangian dual in item (3) plays an important role in developing a projected subgradient algorithm for solving (3.5) in Section 3.3.

In Theorem 3.3, we have been investigating the case where $\mathcal{B} = \{L_i\}_{i=1}^n$ in Definition 3.1, which can be generalized to the convex hull of these generators. Precisely, we shall take

$$\mathcal{B} = \operatorname{conv}((L_i)_{i=1}^n) := \left\{ \sum_{i=1}^n \alpha_i L_i;\ (\alpha_i)_{i=1}^n \in \mathcal{S}_n, L_i \in \mathcal{L} \text{ for all } i \in [\![n]\!] \right\}$$

and consider minimax problem of the form

$$\inf_{M \in \mathcal{L}(\pi)} \sup_{L \in \operatorname{conv}((L_i)_{i=1}^n)} D_f(M||L) = \overline{v}^f(\operatorname{conv}((L_i)_{i=1}^n), \pi). \tag{3.6}$$

An important family of Markov generators that can be written as a convex hull is the family of continuized doubly stochastic Markov generators that we denote by $\mathcal{D}$. We say that $L \in \mathcal{D} \subseteq \mathcal{L}$ if all the row and column sums of L are 0 and $L = P - I$ where P is a Markov matrix and I is the identity matrix on $\mathcal{X}$, and it can be shown that the discrete uniform distribution on $\mathcal{X}$ is the stationary distribution of such L. Let P_i be a permutation matrix on $\mathcal{X}$, then by the Birkhoff-von Neumann theorem (Horn and Johnson, 2013, Theorem 8.7.2) we have

$$\mathcal{D} = \operatorname{conv}((P_i - I)_{i=1}^{|\mathcal{X}|!}). \tag{3.7}$$

Another family of Markov generators that can be casted under this framework is the set of uniformizable (see Keilson (1979, Chapter 2)) and μ-reversible generators, which are used in finite truncation of countably infinite Markov chains in queueing theory (van Dijk et al., 2018), see Example 3.15 below.

Our next proposition shows that the analysis in this case of convex hull of generators can be reduced to the setup of Theorem 3.3. The proof can be found in Section 4.4.

Proposition 3.4 (Reduction to the finite case)**.**

$$\overline{v}^f(\operatorname{conv}((L_i)_{i=1}^n), \pi) = \overline{v}^f(\{L_i\}_{i=1}^n, \pi),$$
$$\underline{v}^f(\operatorname{conv}((L_i)_{i=1}^n), \pi) = \underline{v}^f(\{L_i\}_{i=1}^n, \pi).$$

The next result collects and combines both Theorem 3.3 and Proposition 3.4. Its proof is omitted as it is simply restating these two results in a single Corollary.

Corollary 3.5.

(1) *(Attainment of the minimax problem* (3.6)*) There exists* $\mathbf{w}^f = \mathbf{w}^f(\{L_i\}_{i=1}^n, \pi) = (w_i^f)_{i=1}^n \in \mathcal{S}_n$ *such that the problem* (3.6) *optimal value is attained at* $(M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi), L_j)$*, where* $j \in [\![n]\!]$ *is such that* $w_j^f > 0$.

(2) *(Characterization of $\overline{v}^f = \underline{v}^f$) The saddle point property with respect to $(f, \mathrm{conv}((L_i)_{i=1}^n), \pi)$ (Definition 3.1) holds, that is,*

$$\overline{v}^f(\mathrm{conv}((L_i)_{i=1}^n), \pi) = \underline{v}^f(\mathrm{conv}((L_i)_{i=1}^n), \pi)$$

if and only if there exists $j \in [\![n]\!]$ and $\mathbf{w}^f = (w_i^f)_{i=1}^n$ a maximizer of the dual problem (3.5) with $w_j^f > 0$ such that

$$M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi) \in \operatorname*{arg\,min}_{M \in \mathcal{L}(\pi)} D_f(M||L_j).$$

In particular, if the parameters $(f, \{L_i\}_{i=1}^n, \pi)$ satisfy the assumptions in Theorem 2.1, then the above statement is equivalent to

$$M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi) = M^f(L_j, \pi).$$

(3) *(Saddle point) If the saddle point property with respect to $(f, \mathrm{conv}((L_i)_{i=1}^n), \pi)$ holds, then*

$$(M^f(L_l, \pi), L_l)$$

is a saddle point with respect to $(f, \mathrm{conv}((L_i)_{i=1}^n), \pi)$, where $l = \operatorname*{arg\,max}_{i \in [\![n]\!]} D_f(M^f(L_i, \pi), L_i)$.

In particular, the above results hold for the set of continuized doubly stochastic Markov generators $\mathcal{D}$ (3.7) with $L_i = P_i$ and $n = |\mathcal{X}|!$.

3.1. *Results on the pure strategy game.* In this subsection, we provide a game-theoretic interpretation of the max-min inequalities and the saddle point property (Definition 3.1) in the context of the two-person pure strategy game as introduced in Section 2.2.

The max-min inequalities in (3.1) and (3.2) state that it is advantageous for a player to play second, or more precisely, to know the strategy of the opponent. The difference $\overline{v}^f - \underline{v}^f \geq 0$ can be interpreted as the advantage conferred to a player in knowing the opponent's strategy. If the minimax equality holds, that is, if $\overline{v}^f = \underline{v}^f$, then there is no advantage in playing second or in knowing the strategy of the opponent.

The main result below characterizes the existence and uniqueness of a Nash equilibrium in this pure strategy game. The proof is deferred to Section 4.5.

Corollary 3.6. *Consider the two-person pure strategy game as discussed in Section 2.2. Let $\mathcal{B}$ be either $\{L_i\}_{i=1}^n$ or its convex hull $\mathrm{conv}((L_i)_{i=1}^n)$.*

(1) *(Characterization of pure strategy Nash equilibrium) A pure strategy Nash equilibrium with respect to $(f, \mathcal{B}, \pi)$ exists if and only if there exists $j \in [\![n]\!]$ and $\mathbf{w}^f = (w_i^f)_{i=1}^n$ a maximizer of the dual problem (3.5) with $w_j^f > 0$ such that*

$$M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi) \in \operatorname*{arg\,min}_{M \in \mathcal{L}(\pi)} D_f(M||L_j).$$

In particular, if the parameters $(f, \{L_i\}_{i=1}^n, \pi)$ satisfy the assumptions in Theorem 2.1, then the above statement is equivalent to

$$M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi) = M^f(L_j, \pi).$$

(2) *(Uniqueness of pure strategy Nash equilibrium) Suppose that the parameters $(f, \{L_i\}_{i=1}^n, \pi)$ satisfy the assumptions in Theorem 2.1 and a pure strategy Nash equilibrium exists, which is given by*

$$(M^f(L_l, \pi), L_l),$$

where $l \in \arg\max_{i \in [\![n]\!]} D_f(M^f(L_i, \pi), L_i)$. It is unique if and only if the index

$$l = \operatorname*{arg\,max}_{i \in [\![n]\!]} D_f(M^f(L_i, \pi), L_i)$$

is unique.

3.1.1. *Examples.* In this subsection, we give a few simple yet illustrative examples to demonstrate the theory that we have developed thus far. We shall see that, depending on the parameters of the game $(f, \{L_i\}_{i=1}^n, \pi)$, the associated pure strategy Nash equilibrium or saddle point may or may not exist.

Example 3.7 (An example with multiple pure strategy Nash equilibria (or saddle points)). In the first example, we take $n = 2$ and consider two generators with $L_1 = L$ and $L_2 = L_\pi$, where $L \in \mathcal{L}$ and recall that L_π is the π-dual of L.

In view of the bisection property (Choi and Wolfer, 2024), we have $D_f(M||L_1) = D_f(M||L_2)$, and hence for any weight $\mathbf{w} \in \mathcal{S}_2$, we see that

$$M_2^f(\mathbf{w}, \{L_i\}_{i=1}^2, \pi) \in \operatorname*{arg\,min}_{M \in \mathcal{L}(\pi)} D_f(M||L_i)$$

for $i = 1, 2$. Thus, a pure strategy Nash equilibrium or saddle point, with respect to $(f, \{L_i\}_{i=1}^2, \pi)$, exists, according to Corollary 3.6.

Now, we further assume that the parameters $(f, \{L_i\}_{i=1}^n, \pi)$ satisfy the assumptions in Theorem 2.1. Using the bisection property again, the two pure strategy Nash equilibria are

$$(M^f(L, \pi), L), \quad (M^f(L, \pi), L_\pi).$$

Example 3.8 (An example with a unique pure strategy Nash equilibrium (or saddle point)). In the second example, we again take $n = 2$ and consider two generators with $L_1, L_2 \in \mathcal{L}$. We further assume that D_f is the α-divergence, and L_1, L_2 are chosen to satisfy the condition that

$$D_f(M^f(L_1, \pi)||L_1) > D_f(M^f(L_1, \pi)||L_2). \tag{3.8}$$

Applying the Pythagorean identity of the α-divergence (Choi and Wolfer, 2024) to the right hand side above, we thus have

$$\begin{aligned} D_f(M^f(L_1, \pi)||L_1) > D_f(M^f(L_1, \pi)||L_2) &= D_f(M^f(L_2, \pi)||L_2) + D_f(M^f(L_1, \pi)||M^f(L_2, \pi)) \\ &\geq D_f(M^f(L_2, \pi)||L_2). \end{aligned} \tag{3.9}$$

Now, we claim that in this example, the pair $(M^f(L_1, \pi), D_f(M^f(L_1, \pi)||L_1)) \in \mathcal{L}(\pi) \times \mathbb{R}_+$ minimizes the primal problem (3.4) and $\mathbf{w} = (1, 0) \in \mathcal{S}_2$ maximizes the dual problem (3.5). This readily follows from item (2) in Theorem 3.3, where the condition (3.8) ensures that the complementary slackness holds.

Using item (1) in Corollary 3.6, we see that a Nash equilibrium exists, and in view of (3.9) and item (2) in Corollary 3.6, the unique pure strategy Nash equilibrium or saddle point is given by

$$(M^f(L_1, \pi), L_1).$$

Example 3.9 (An example where a pure strategy Nash equilibrium (or saddle point) fails to exist). In the final example, we again specialize into $n = 2$ and take two Markov generators with $L_1, L_2 \in \mathcal{L}$. We further assume that D_f is the α-divergence, and L_1, L_2 are chosen to satisfy the conditions that $M^f(L_1, \pi) \neq M^f(L_2, \pi)$ and

$$D_f(M^f(L_1, \pi)||L_1) = D_f(M^f(L_2, \pi)||L_2). \tag{3.10}$$

Suppose for contradiction that a pure strategy Nash equilibrium exists. In view of Corollary 3.6 and f is strictly convex under α-divergence, the only possible candidates for the maximizer of the dual problem is either $\mathbf{w}^f = (1, 0)$ or $\mathbf{w}^f = (0, 1)$.

In the former case, we thus have $D_f(M^f(L_1,\pi)||L_1) = r$ and

$$\begin{aligned} r &= D_f(M^f(L_1,\pi)||L_1) \\ &\geq D_f(M^f(L_1,\pi)||L_2) \\ &= D_f(M^f(L_2,\pi)||L_2) + D_f(M^f(L_1,\pi)||M^f(L_2,\pi)) \\ &= D_f(M^f(L_1,\pi)||L_1) + D_f(M^f(L_1,\pi)||M^f(L_2,\pi)), \end{aligned}$$

where we use the Pythagorean identity (Choi and Wolfer, 2024) in the second equality and (3.10) in the last equality. This implies that $D_f(M^f(L_1,\pi)||M^f(L_2,\pi)) = 0$ and hence $M^f(L_1,\pi) = M^f(L_2,\pi)$, which contradicts the assumption. Analogously we can handle the case of $\mathbf{w}^f = (0,1)$.

As a result, we see that a pure strategy Nash equilibrium fails to exist. For further discussions on this example, please refer to Example 3.13.

3.2. *Results on the mixed strategy game.* This subsection concerns the two-person mixed strategy game as introduced in Section 2.2. Before we state the main results of this subsection, we give an important proposition that connects various minimax and maximin values we have introduced thus far, namely $\overline{v}^f, \overline{V}^f, \underline{V}^f$ and the dual problem (3.5). The proof is deferred to Section 4.6.

Proposition 3.10. *For any $\mathcal{B} \subseteq \mathcal{L}$, we have*

$$\overline{v}^f(\mathcal{B},\pi) \geq \overline{V}^f(\mathcal{B},\pi) \geq \underline{V}^f(\mathcal{B},\pi), \tag{3.11}$$

where we recall that $\overline{v}^f$ is introduced in Definition 3.1 while $\overline{V}^f, \underline{V}^f$ are defined in Definition 2.3. In particular, if $\mathcal{B}$ is either $\{L_i\}_{i=1}^n$ or its convex hull $\mathrm{conv}((L_i)_{i=1}^n)$, this leads to

$$\overline{v}^f(\mathcal{B},\pi) \geq \overline{V}^f(\mathcal{B},\pi) \geq \underline{V}^f(\mathcal{B},\pi) \geq \max_{\mathbf{w}\in\mathcal{S}_n} \sum_{i=1}^n w_i D_f(M_n^f||L_i). \tag{3.12}$$

In the setting of (3.12), by Theorem 3.3 item (1), the strong duality holds and hence we have

$$\overline{v}^f(\mathcal{B},\pi) = \max_{\mathbf{w}\in\mathcal{S}_n} \sum_{i=1}^n w_i D_f(M_n^f||L_i).$$

This forces a minimax equality, that is,

$$\overline{V}^f(\mathcal{B},\pi) = \underline{V}^f(\mathcal{B},\pi).$$

In other words, all inequalities are in fact equalities in (3.12). We collect this result and a few others in Corollary 3.5 into the following main results:

Theorem 3.11. *Consider the two-person mixed strategy game with respect to the parameters $(f,\mathcal{B},\pi)$, where $\mathcal{B}$ is either $\{L_i\}_{i=1}^n$ or its convex hull $\mathrm{conv}((L_i)_{i=1}^n)$. Recall that there exists $\mathbf{w}^f = \mathbf{w}^f(\{L_i\}_{i=1}^n,\pi) = (w_i^f)_{i=1}^n \in \mathcal{S}_n$ such that the optimal value for the dual problem (3.5) is attained at $\mathbf{w}^f$ by Theorem 3.3.*

(1) *(A mixed strategy Nash equilibrium always exists) The minimax equality holds, that is,*

$$\overline{V}^f(\mathcal{B},\pi) = \underline{V}^f(\mathcal{B},\pi) = \sum_{i=1}^n w_i^f D_f(M_n^f(\mathbf{w}^f,\{L_i\}_{i=1}^n,\pi)||L_i) =: V,$$

and hence V is the value of the mixed strategy game. The pair $(M_n^f(\mathbf{w}^f,\{L_i\}_{i=1}^n,\pi),\mathbf{w}^f) \in \mathcal{L}(\pi)\times\mathcal{P}(\mathcal{B})$ is a mixed strategy Nash equilibrium with respect to the parameters $(f,\mathcal{B},\pi)$.

(2) *(Game-theoretic interpretation of Chebyshev center and Chebyshev radius) Recall the definition of Chebyshev center and Chebyshev radius in Definition 3.2. A minimax strategy of Nature is given by* $M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi)$*, a weighted information centroid, which is also a Chebyshev center. In view of the complementary slackness in Theorem 3.3, we have*

$$V = D_f(M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi)||L_l),$$

where the index l *satisfies* $w_l^f > 0$*. In words, the Chebyshev radius is the value of the game* V*. If the parameters* $(f, \{L_i\}_{i=1}^n, \pi)$ *satisfy the assumptions in Theorem 2.1, then* $M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi)$ *is the unique minimax strategy.*

(3) *(Bayes risk with respect to* $\mathbf{w}^f$ *is value of the game)* $\mathbf{w}^f$ *is a maximin strategy of the probabilist.* $M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi)$ *is a Bayes strategy with respect to* $\mathbf{w}^f$*, and the Bayes risk with respect to* $\mathbf{w}^f$ *is the value of the game* V*.*

Remark 3.2. It is interesting to note that a mixed strategy Nash equilibrium always exists in the two-person mixed strategy game (Theorem 3.11) while a pure strategy Nash equilibrium may or may not exist in the two-person pure strategy game (Corollary 3.6). For an example of the latter case we recall Example 3.9 and another Example 3.13 below. This is analogous to the classical two-person zero-sum game in the game theory literature where a pure strategy Nash equilibrium may not exist, see Karlin and Peres (2017, Chapter 2).

Remark 3.3 (Game-theoretic consequences of the mixed strategy game). As a mixed strategy Nash equilibrium always exists, there is no advantage for a player (the probabilist or Nature) to play second, or more generally, to know the strategy of the opponent in the mixed strategy game.

Remark 3.4 (Another proof of $\overline{V}^f(\mathcal{B}, \pi) = \underline{V}^f(\mathcal{B}, \pi)$ via the Sion's minimax theorem). One common strategy in establishing minimax results, in the context of source coding and information theory, relies on the Sion's minimax theorem, see for instance van Erven and Harremoës (2014, Theorem 35). For given $\mu \in \mathcal{P}(\mathcal{B})$ and $M \in \mathcal{L}(\pi)$, the mapping

$$(M, \mu) \mapsto \int_{\mathcal{B}} D_f(M||L)\, \mu(dL)$$

is clearly concave in μ and convex in M. As we are minimizing over $M \in \mathcal{L}(\pi)$, the set of π-reversible generator $\mathcal{L}(\pi)$ is convex yet unbounded, and hence it is not a compact subset of $\mathcal{L}$. On the other hand, as $\mathcal{B}$ is either a finite set of Markov generators $\{L_i\}_{i=1}^n$ or its convex hull $\mathrm{conv}((L_i)_{i=1}^n)$, the set $\mathcal{P}(\mathcal{B})$ is thus compact. As a result, the Sion's minimax theorem is readily applicable in this setting which yields

$$\overline{V}^f(\mathcal{B}, \pi) = \underline{V}^f(\mathcal{B}, \pi).$$

The proof that we presented in (3.12) relies on the Lagrangian duality theory as in Theorem 3.3, which naturally gives fine properties concerning the mixed strategy Nash equilibrium and related important objects such as $\mathbf{w}^f$ and connects with earlier sections of the manuscript. These properties do not seem to be immediate consequences or corollaries from the Sion's minimax theorem.

3.2.1. *Examples.* In this subsection, similar to Section 3.1.1, we detail four simple examples to demonstrate the theory concerning the two-person mixed strategy game.

Example 3.12 (Answering a question of Laurent Miclo: an example with the existence of multiple mixed strategy Nash equilibria). In the first example, we continue the setting discussed in Example 3.7. That is, we consider two ($n = 2$) generators with $L_1 = L$ and $L_2 = L_\pi$, where $L \in \mathcal{L}$.

Using the bisection property of D_f (Choi and Wolfer, 2024), we have $D_f(M||L_1) = D_f(M||L_2)$, and for any weight $\mathbf{w} \in \mathcal{S}_2$, we see that

$$M_2^f(\mathbf{w}, \{L_i\}_{i=1}^2, \pi) \in \operatorname*{arg\,min}_{M \in \mathcal{L}(\pi)} D_f(M||L_i)$$

for $i = 1, 2$. In other words, any pair of the form $(M^f(L_1, \pi), \mathbf{w})$ is a mixed strategy Nash equilibrium with respect to the parameters of this game. Note that this game also has multiple pure strategy Nash equilibria as shown in Example 3.7.

In the special case of considering $f(t) = |t - 1|$, the f-divergence is the total variation distance. We also recall $P_{-\infty}, P_\infty$ are introduced in (2.6) and (2.5) respectively. It is shown in Choi and Huang (2020) that any convex combinations of $P_{-\infty}$ and P_∞ minimize

$$\{aP_{-\infty} + (1-a)P_\infty;\ a \in [0,1]\} \subseteq \underset{M \in \mathcal{L}(\pi)}{\arg\min}\, D_f(M||L_i).$$

As such, based on our earlier analysis, $(P_\infty, \mathbf{w})$ and $(P_{-\infty}, \mathbf{w})$ are mixed strategy Nash equilibria of the game under the total variation distance. This gives a possible answer to a question of Laurent Miclo on game-theoretic interpretations of $P_{-\infty}, P_\infty$.

If f is strictly convex with $f'(1) = 0$, for instance the α-divergence, then $M^f(L_1, \pi)$ is unique, and hence it is the unique minimax strategy of Nature in the mixed strategy game, while any $\mathbf{w} \in \mathcal{S}_2$ is a maximin strategy of the probabilist. $M^f(L_1, \pi)$ is also the unique Bayes strategy with respect to any $\mathbf{w}$. The value of the game is given by $V = D_f(M^f(L_1, \pi)||L_1)$.

Note that $M^f(L_1, \pi)$ is also the unique minimax strategy of Nature in the pure strategy game.

Example 3.13 (An example where a pure strategy Nash equilibrium fails to exist and a mixed strategy Nash equilibrium (always) exists). In the second example, we continue our investigation in Example 3.9 and take $n = 2$ along with two Markov generators where $L_1, L_2 \in \mathcal{L}$. We choose D_f to be the α-divergence, and L_1, L_2 are assumed to satisfy the conditions that $M^f(L_1, \pi) \neq M^f(L_2, \pi)$ and

$$D_f(M^f(L_1, \pi)||L_1) = D_f(M^f(L_2, \pi)||L_2). \tag{3.13}$$

Recall that in Example 3.9, we have already shown that a pure strategy Nash equilibrium cannot exist with these choices of parameters, and furthermore a maximizer of the dual problem (3.5), which exists, cannot be $\mathbf{w}^f = (1, 0)$ and $\mathbf{w}^f = (0, 1)$.

We thus come to the conclusion that $w_i^f > 0$ for $i = 1, 2$. The weighted information centroid $M_2^f(\mathbf{w}^f, \{L_i\}_{i=1}^2, \pi)$ is the unique minimax strategy of Nature, and is also the unique Bayes strategy with respect to $\mathbf{w}^f$.

Figure 3.1 demonstrates a possible visualization of this example.

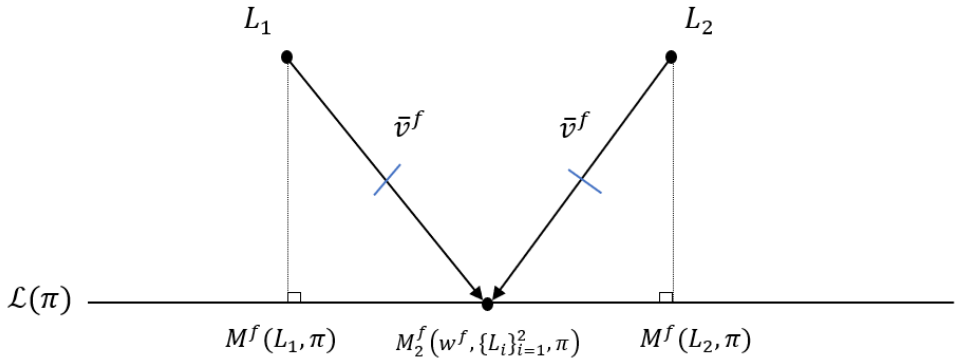

FIGURE 3.1. A visualization of Example 3.13. Note that the mixed strategy Nash equilibrium is $(M_2^f(\mathbf{w}^f, \{L_i\}_{i=1}^2, \pi), \mathbf{w}^f)$, while a pure strategy Nash equilibrium does not exist. The complementary slackness in Theorem 3.3 ensures that $D_f(M_2^f(\mathbf{w}^f, \{L_i\}_{i=1}^2, \pi)||L_1) = D_f(M_2^f(\mathbf{w}^f, \{L_i\}_{i=1}^2, \pi)||L_2) = \overline{v}^f = V$, the value of the mixed strategy game. (3.13) also holds in this figure.

Example 3.14 (An example where a pure strategy Nash equilibrium and mixed strategy Nash equilibrium coincides). In this example, we take $n = 3$ Markov generators $L_i \in \mathcal{L}$ for $i \in [\![3]\!]$, and they are chosen in a special way such that

$$D_f(M^f(L_2, \pi)||L_1) = D_f(M^f(L_2, \pi)||L_3) = D_f(M^f(L_2, \pi)||L_2).$$

At the same time, we require that, for $j \in \{1, 3\}$,

$$D_f(M^f(L_2, \pi)||L_2) > D_f(M^f(L_j, \pi)||L_j).$$

With these parameter choices of the game, using Theorem 3.3 and 3.11 we see that we can take $\mathbf{w}^f = (0, 1, 0)$, and $(M^f(L_2, \pi), \mathbf{w}^f)$ is a mixed strategy Nash equilibrium while $(M^f(L_2, \pi), L_2)$ is a pure strategy Nash equilibrium. Figure 3.2 provides a visualization in this setting.

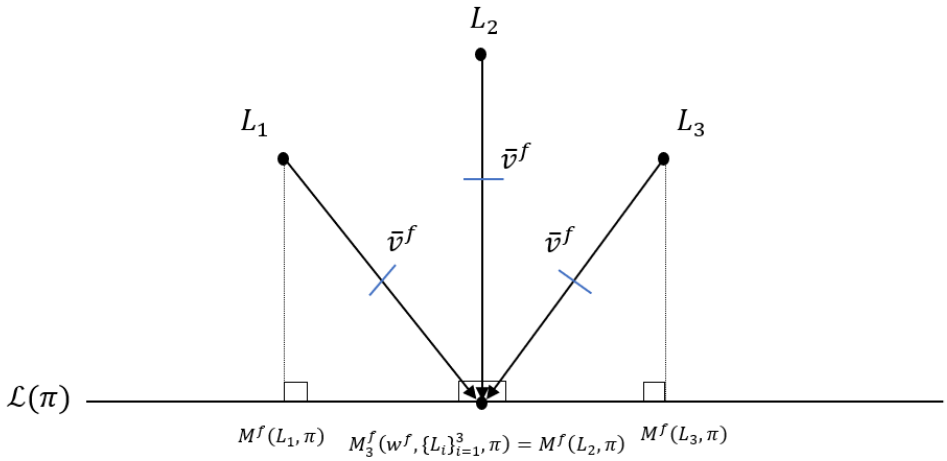

FIGURE 3.2. A visualization of Example 3.14.

Example 3.15 (An example of uniformizable and μ-reversible generators). Recall that we have been considering settings where $\mathcal{B}$ is either $\{L_i\}_{i=1}^n$ or its convex hull $\mathrm{conv}((L_i)_{i=1}^n)$. This covers for instance $\mathcal{D}$, the set of doubly stochastic Markov generators, as discussed in (3.7). The aim of this example is to show that this also includes the set of uniformizable and μ-reversible generators, where $\mu \neq \pi$.

Let $L \in \mathcal{L}(\mu)$. L is said to be λ-uniformizable if $\max_{x \in \mathcal{X}} |L(x, x)| \leq \lambda$, and we denote by $\mathcal{L}_\lambda(\mu) \subset \mathcal{L}(\mu)$ to be the set of such Markov generators. Without loss of generality, in this example we let $m = |\mathcal{X}|$ and label the state space $\mathcal{X} = \{1, 2, \ldots, m\}$. For a given $\lambda > 0$, we now claim that

$$\mathcal{L}_\lambda(\mu) \subseteq \mathrm{conv}(\{L_{x,y}\}_{x<y, x,y \in \mathcal{X}} \cup \{\mathbf{0}\}),$$

where for $x < y \in \mathcal{X}$, $L_{x,y}(x, y) := \lambda \dfrac{m(m-1)}{2}$ while $L_{x,y}(y, x) := \lambda \dfrac{m(m-1)}{2} \dfrac{\mu(x)}{\mu(y)}$ and is zero otherwise for other off-diagonal entries of $L_{x,y}$. We also recall that $\mathbf{0}$ is the all-zero Markov generator. Let $L \in \mathcal{L}_\lambda(\mu)$. Making use of the μ-reversibility, we write that

$$L = \sum_{x<y} w_{x,y} L_{x,y} + \left(1 - \sum_{x<y} w_{x,y}\right) \mathbf{0},$$

where $w_{x,y} := \dfrac{L(x, y)}{\lambda \dfrac{m(m-1)}{2}}$.

As a consequence, various results that have been stated are readily applicable to the set $\mathcal{L}_\lambda(\mu)$.

3.3. *A projected subgradient algorithm to find an approximate mixed strategy Nash equilibrium (or Chebyshev center).* The aim of this subsection is to develop a simple and easy-to-implement projected subgradient method to find an approximate mixed strategy Nash equilibrium. Throughout this subsection, we consider the two-person mixed strategy game introduced in Section 2.2 with respect to the parameters $(f, \mathcal{B}, \pi)$, where $\mathcal{B}$ is either $\{L_i\}_{i=1}^n$ or its convex hull $\mathrm{conv}((L_i)_{i=1}^n)$. Recall that $\mathbf{w}^f$ is a maximin strategy of the probabilist, and the pair $(M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi), \mathbf{w}^f)$ is a mixed strategy Nash equilibrium. To find such an equilibrium algorithmically, one crucial step amounts to solving the corresponding dual problem (3.5). If the weighted information centroid admits a closed-form expression, for instance in the case of α-divergence or the examples given in Theorem 2.2, then making use of $\mathbf{w}^f$ a mixed strategy Nash equilibrium is found.

Instead of solving the maximization problem in (3.5), we reformulate it into the following equivalent minimization problem by multiplying by -1, in order for us to consider the subgradient (rather than supergradient) method in the optimization literature:

$$\min_{\mathbf{w}\in\mathcal{S}_n} h(\mathbf{w})$$

where

$$h(\mathbf{w}) := -\sum_{i=1}^n w_i D_f(M_n^f(\mathbf{w}, \{L_j\}_{j=1}^n, \pi)||L_i) \tag{3.14}$$

Note that the dependence of h on $(f, \mathcal{B}, \pi)$ is suppressed to avoid notational burden in the definition above. In view of Theorem 3.3 item (3), the function h is convex.

The algorithm that we propose is novel from two perspectives. First, no algorithm has been developed in the context of the game of the statistician against Nature (Haussler, 1997; Haussler and Opper, 1997; Gushchin and Zhdanov, 2006) to compute the Nash equilibrium, while in this paper we propose a simple subgradient algorithm to approximately find the mixed strategy equilibrium. Second, our algorithm harnesses on the centroid structure and the subgradient of h, which has not been observed in the literature of Chebyshev center computation of probability measures (Eldar et al., 2008; Candan, 2020).

In the first main result of this Section, we identify a subgradient of h, and its proof is deferred to Section 4.7:

Theorem 3.16 (Subgradient of h and an upper bound of its ℓ^2-norm)**.** *A subgradient of h at $\mathbf{v} \in \mathcal{S}_n$ is given by $\mathbf{g} = \mathbf{g}(\mathbf{v}) = (g_1, g_2, \ldots, g_n) \in \mathbb{R}^n$, where for $i \in [\![n]\!]$, we have*

$$g_i = D_f(M_n^f(\mathbf{v}, \{L_j\}_{j=1}^n, \pi)||L_n) - D_f(M_n^f(\mathbf{v}, \{L_j\}_{j=1}^n, \pi)||L_i). \tag{3.15}$$

That is, $\mathbf{g}$ satisfies, for $\mathbf{w}, \mathbf{v} \in \mathcal{S}_n$,

$$h(\mathbf{w}) \geq h(\mathbf{v}) + \sum_{i=1}^n g_i(w_i - v_i).$$

Moreover, the ℓ^2-norm of $\mathbf{g}(\mathbf{v})$ is bounded above by

$$\|\mathbf{g}(\mathbf{v})\|_2^2 := \sum_{i=1}^n g_i^2 \leq n\left(|\mathcal{X}| \sup_{\mathbf{v}\in\mathcal{S}_n;\ i\in[\![n]\!];\ L_i(x,y)>0} L_i(x,y) f\left(\frac{M_n^f(\mathbf{v}, \{L_j\}_{j=1}^n, \pi)(x,y)}{L_i(x,y)}\right)\right)^2 =: B. \tag{3.16}$$

Remark 3.5. The choice of L_n in the first term of g_i in (3.15) is in fact completely arbitrary, and we shall see in the proof that it can be replaced by any L_l with $l \in [\![n]\!]$. We thus obtain at least n subgradients of h at a given point $\mathbf{v} \in \mathcal{S}_n$.

Remark 3.6. The upper bound B plays an important role in determining the convergence rate of the projected subgradient algorithm, see Theorem 3.17 and its proof below.

We proceed to develop and analyze a projected subgradient algorithm to solve the dual problem (3.5). Suppose that the number of iterations to be ran is t.

For $i = 1, 2, \ldots, t$, we first update the weights of the current iteration by its subgradient, that is, we set

$$\mathbf{v}^{(i)} = \mathbf{w}^{(i-1)} - \eta \mathbf{g}(\mathbf{w}^{(i-1)}),$$

where $\eta > 0$ is the constant stepsize of the algorithm while we recall that $\mathbf{g}$ is a subgradient of h as in (3.15). In the second step, we project $\mathbf{v}^{(i)}$ onto the probability simplex $\mathcal{S}_n$ by computing

$$\mathbf{w}^{(i)} = \arg\min_{\mathbf{w} \in \mathcal{S}_n} \left\| \mathbf{w} - \mathbf{v}^{(i)} \right\|_2^2.$$

This can be done by applying existing efficient projection algorithms (see e.g. Condat (2016)), and we do not further investigate these projection algorithms in this manuscript. We then repeat the above two steps for $i = 1, 2, \ldots, t$. The output of the algorithm is the sequence $(\mathbf{w}^{(i)})_{i=1}^t$. Precisely, the algorithm is stated in Algorithm 1.

Algorithm 1: A projected subgradient algorithm to find an approximate mixed strategy Nash equilibrium

Input : Initial weight value $\mathbf{w}^{(0)} \in \mathcal{S}_n$, function f (and D_f), set $\{L_i\}_{i=1}^n$, target distribution π, stepsize $\eta > 0$ and number of iterations t.

Output: The sequence $(\mathbf{w}^{(i)})_{i=1}^t$.

for $i = 1, 2, \ldots, t$ **do**

  `// Update via subgradient`

  $\mathbf{v}^{(i)} = \mathbf{w}^{(i-1)} - \eta \mathbf{g}(\mathbf{w}^{(i-1)})$;

  `// Projection onto` $\mathcal{S}_n$

  $\mathbf{w}^{(i)} = \arg\min_{\mathbf{w} \in \mathcal{S}_n} \left\| \mathbf{w} - \mathbf{v}^{(i)} \right\|_2^2$.

Before we proceed to analyze the convergence of Algorithm 1, we first build some intuition about why it might possibly work. Suppose that we are in the hypothetical setting where the optimal weights are all positive, that is, $w_i^f > 0$ for all $i \in [\![n]\!]$. Recall that by the complementary slackness condition in Theorem 3.3, in this setting the subgradient of h at $\mathbf{w}^f$ is zero since

$$g_i(\mathbf{w}^f) = D_f(M_n^f(\mathbf{w}^f, \{L_j\}_{j=1}^n, \pi) || L_n) - D_f(M_n^f(\mathbf{w}^f, \{L_j\}_{j=1}^n, \pi) || L_i) = 0.$$

As a result, once Algorithm 1 reaches $\mathbf{w}^f$, it stays put and no longer moves again, which is a desirable behaviour as an optimal value has been reached. On the other hand, if at the current iteration the subgradient is $g_i > 0$ at some i, then the weights are updated along the subgradient direction followed by projection onto $\mathcal{S}_n$.

The success of Algorithm 1 relies crucially on computation of the subgradient $\mathbf{g}(\mathbf{w})$ of h at $\mathbf{w}$, which further depends on the expression $M_n^f(\mathbf{w}, \{L_j\}_{j=1}^n, \pi)$. For example, it is known in the case of α-divergence or the examples given in Theorem 2.2, and hence the algorithm can be readily applied in these settings. Note that in the implementation of Algorithm 1, we assume that the centroid $M_n^f(\mathbf{w}, \{L_j\}_{j=1}^n, \pi)$ is accessible and we do not discuss algorithms to compute these centroids numerically.

In this paper, we do not consider the many possible variants of the projected subgradient algorithm in Algorithm 1, for instance algorithms with changing or adaptive stepsize, or algorithms with stochastic subgradients.

Our second main result of this subsection gives a convergence rate of $\mathcal{O}(1/\sqrt{t})$ of Algorithm 1 with an appropriate choice of stepsize. The proof can be found in Section 4.8. We also write

$$\overline{\mathbf{w}}^t := \frac{1}{t}\sum_{i=1}^{t} \mathbf{w}^{(i)},$$

the arithmetic average up to iteration t of the outputs of Algorithm 1.

Theorem 3.17. *Consider Algorithm 1 with its output $(\mathbf{w}^{(i)})_{i=1}^t$ and the notations therein. We have*

$$h\left(\overline{\mathbf{w}}^t\right) - h\left(\mathbf{w}^f\right) \leq \frac{n}{2\eta t} + \frac{\eta B}{2}, \tag{3.17}$$

where we recall that h is the function defined in (3.14) and B is introduced in (3.16). In particular, if we take the constant stepsize to be

$$\eta = \sqrt{\frac{n}{tB}},$$

we thus have

$$h\left(\overline{\mathbf{w}}^t\right) - h\left(\mathbf{w}^f\right) \leq \frac{1}{2}\sqrt{\frac{nB}{t}} + \frac{1}{2}\sqrt{\frac{nB}{t}} = \sqrt{\frac{nB}{t}}.$$

Given an arbitrary $\varepsilon > 0$, if we further choose

$$t = \left\lceil \frac{nB}{\varepsilon^2} \right\rceil,$$

then we reach an ε-close value to $h\left(\mathbf{w}^f\right)$ in the sense that

$$h\left(\overline{\mathbf{w}}^t\right) - h\left(\mathbf{w}^f\right) \leq \varepsilon.$$

Before we introduce the final main result of this Section, we shall fix a few notations. We shall consider the set of continuized generators that does not stay at the same state in one step, that is, for $i \in [\![n]\!]$ we consider

$$L_i := P_i - I, \tag{3.18}$$

where I is the identity matrix on $\mathcal{X}$ and P_i is a transition matrix with $P_i(x,x) = 0$ for all $x \in \mathcal{X}$. For two Markov generators $L, M \in \mathcal{L}$, by choosing $f(t) = |t-1|/2$, the total variation distance between L, M is given by

$$D_{\mathrm{TV}}(L||M) := D_f(L||M) = \frac{1}{2}\sum_{x\in\mathcal{X}} \pi(x) \sum_{y\in\mathcal{X}\setminus\{x\}} |L(x,y) - M(x,y)| = D_{\mathrm{TV}}(M||L).$$

For the family of $\{L_i\}_{i=1}^n$ satisfying (3.18), we quantify the convergence of $M_n^f(\overline{\mathbf{w}}^t, \{L_j\}_{j=1}^n, \pi)$ towards $M_n^f(\mathbf{w}^f, \{L_j\}_{j=1}^n, \pi)$ in terms of the total variation distance D_{TV} and a strictly convex f. The proof can be found in Section 4.9.

Theorem 3.18 ($\mathcal{O}\left(1/\sqrt{t}\right)$ convergence rate of the weighted information centroid generated by Algorithm 1)**.** *Consider Algorithm 1 applied to the family $\{L_j\}_{j=1}^n$ satisfying (3.18) under a strictly convex f with its output $(\mathbf{w}^{(i)})_{i=1}^t$ and the notations therein, and the stepsize is chosen to be*

$$\eta = \sqrt{\frac{n}{tB}}.$$

We have

$$D_{\mathrm{TV}}(M_n^f(\overline{\mathbf{w}}^t, \{L_j\}_{j=1}^n, \pi)||M_n^f(\mathbf{w}^f, \{L_j\}_{j=1}^n, \pi)) = \mathcal{O}\left(\frac{1}{\sqrt{t}}\right), \tag{3.19}$$

where the constant in front on the right hand side depends on all the parameters except t, that is, the constant depends on $(f, \{L_j\}_{j=1}^n, \pi, n, \mathcal{X})$.

4. Proofs of the main results

4.1. *Proof of Theorem 2.1.* The proof is a generalization of Diaconis and Miclo (2009, Proposition 1.5) and Choi and Wolfer (2024, Theorem 3.9) to the setting of weighted information centroid. Pick an arbitrary total ordering on $\mathcal{X}$ with strict inequality being denoted by $\prec$. For $i \in [\![n]\!]$, we write

$$a = a(x,y) = \pi(x)M(x,y), \quad a' = a'(y,x) = \pi(y)M(y,x),$$
$$\beta_i = \beta_i(x,y) = \pi(x)L_i(x,y), \quad \beta_i' = \beta_i'(y,x) = \pi(y)L_i(y,x).$$

We note that $M \in \mathcal{L}(\pi)$ gives $a = a'$. Using this, we see that

$$\begin{aligned}
\sum_{i=1}^n w_i D_f(M||L_i) &= \sum_{i=1}^n \sum_{x\prec y} w_i\pi(x)L_i(x,y) f\left(\frac{M(x,y)}{L_i(x,y)}\right) + w_i\pi(y)L_i(y,x) f\left(\frac{M(y,x)}{L_i(y,x)}\right) \\
&= \sum_{i=1}^n \sum_{x\prec y} w_i\beta_i f\left(\frac{a}{\beta_i}\right) + w_i\beta_i' f\left(\frac{a}{\beta_i'}\right) \\
&= \sum_{x\prec y} \sum_{i=1}^n w_i\beta_i f\left(\frac{a}{\beta_i}\right) + w_i\beta_i' f\left(\frac{a}{\beta_i'}\right) \\
&= \sum_{x\prec y} \sum_{\{i;\ w_i>0 \text{ and } \beta_i>0 \text{ or } \beta_i'>0\}} w_i\beta_i f\left(\frac{a}{\beta_i}\right) + w_i\beta_i' f\left(\frac{a}{\beta_i'}\right) \\
&=: \sum_{x\prec y} \Phi_{\mathbf{w},\beta_1,\dots,\beta_n,\beta_1',\dots,\beta_n'}(a).
\end{aligned}$$

To minimize with respect to M, we are led to minimize the summand above $\phi := \Phi_{\mathbf{w},\beta_1,\dots,\beta_n,\beta_1',\dots,\beta_n'} : \mathbb{R}_+ \to \mathbb{R}_+$, where $\mathbf{w} \in \mathcal{S}_n$ and $(\beta_1,\dots,\beta_n,\beta_1',\dots,\beta_n') \in \mathbb{R}_+^{2n}$ are assumed to be fixed. The summation in the expression of ϕ is non-empty in view of the assumptions in the Theorem.

As ϕ and f are convex, we denote by ϕ_+' and f_+' to be their right derivative respectively. It suffices to show the existence of $a_* > 0$ such that for all $a \in \mathbb{R}_+$,

$$\phi_+'(a) = \begin{cases} < 0, & \text{if} \quad a < a_*, \\ > 0, & \text{if} \quad a > a_*. \end{cases} \tag{4.1}$$

Now, we compute that for all $a \in \mathbb{R}_+$,

$$\begin{aligned}
\phi_+'(a) = &\sum_{\{i;\ \beta_i>0 \text{ and } \beta_i'>0\}} w_i f_+'\left(\frac{a}{\beta_i}\right) + w_i f_+'\left(\frac{a}{\beta_i'}\right) \\
&+ \sum_{\{i;\ \beta_i>0 \text{ and } \beta_i'=0\}} w_i f_+'\left(\frac{a}{\beta_i}\right) + \sum_{\{i;\ \beta_i=0 \text{ and } \beta_i'>0\}} w_i f_+'\left(\frac{a}{\beta_i'}\right).
\end{aligned}$$

As $f'(1) = 0$ and f is strictly convex, for sufficiently small $a > 0$ $\phi_+'(a) < 0$ while for sufficiently large $a > 0$ $\phi_+'(a) > 0$ and ϕ_+' is increasing, we conclude that there exists a unique $a_* > 0$ such that (4.1) is satisfied.

If we replace f by f^* in the above proof, and by noting that f^* is also a strictly convex function with $f^*(1) = f^{*\prime}(1) = 0$, the existence and uniqueness of $M_n^{f^*}$ is shown.

4.2. *Proof of Theorem 2.2.* We shall only prove (2.7) as the rest follows exactly the same computation procedure with different choices of f. Pick an arbitrary total ordering on $\mathcal{X}$ with strict

inequality being denoted by $\prec$. For $i = 1, \dots, n$, we also write

$$a = a(x,y) = \pi(x)M(x,y), \quad a' = a'(y,x) = \pi(y)M(y,x),$$
$$\beta_i = \beta_i(x,y) = \pi(x)L_i(x,y), \quad \beta_i' = \beta_i'(y,x) = \pi(y)L_i(y,x).$$

The π-reversibility of M yields $a = a'$, which leads to

$$\begin{aligned}\sum_{i=1}^n w_i D_f(M||L_i) &= \sum_{i=1}^n \sum_{x\prec y} w_i \left(\pi(x)L_i(x,y) f\left(\frac{M(x,y)}{L_i(x,y)}\right) + \pi(y)L_i(y,x) f\left(\frac{M(y,x)}{L_i(y,x)}\right)\right) \\ &= \sum_{i=1}^n \sum_{x\prec y} w_i \left(a - 2\sqrt{a\beta_i} + \beta_i + a' - 2\sqrt{a'\beta_i'} + \beta_i'\right) \\ &= \sum_{x\prec y} \sum_{i=1}^n w_i \left(2a - 2\sqrt{a\beta_i} - 2\sqrt{a\beta_i'} + \beta_i + \beta_i'\right).\end{aligned}$$

Next, we aim to minimize each term in the sum, leading us to minimize the strictly convex function of a:

$$a \mapsto \sum_{i=1}^n w_i \left(2a - 2\sqrt{a\beta_i} - 2\sqrt{a\beta_i'}\right).$$

Differentiating the above expression with respect to a yields

$$M_n^f(x,y) = \left(\sum_{i=1}^n w_i \sqrt{M^f(L_i,\pi)(x,y)}\right)^2.$$

4.3. *Proof of Theorem 3.3.* We first prove item (1) and (2). To see that the strong duality holds for (3.4) and the dual optimal is attained, we shall show that the Slater's constraint qualification (Beck (2017, Theorem A.1) or Boyd and Vandenberghe (2004, Section 5.2.3)) is verified, which amounts to proving that the constraints in (3.4) are strictly feasible. We take

$$\begin{aligned} M &= M^f(L_1,\pi), \\ r &= \max_{i\in[\![n]\!]} D_f(M||L_i) + 1 > D_f(M||L_l),\end{aligned}$$

for all $l \in [\![n]\!]$, and hence the pair (M, r) is strictly feasible. This proves the first part of item (1). As the strong duality holds, using the optimality conditions under strong duality (Beck, 2017, Theorem A.2), it proves item (2). In particular, the complementary slackness conditions entail that for all $i \in [\![n]\!]$,

$$w_i(D_f(M||L_i) - r) = 0,$$

which is equivalent to

$$D_f(M||L_i) \begin{cases} = r & \text{if } w_i > 0, \\ \le r & \text{if } w_i = 0. \end{cases}$$

We return to show the second part of item (1). By the first part and item (2), we have

$$\overline{v}^f = \sum_{i=1}^n w_i^f D_f(M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi)||L_i) = D_f(M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi)||L_l),$$

where $l \in [\![n]\!]$ is such that $w_l^f > 0$.

Next, we prove item (3). As the Lagrangian dual is concave (Boyd and Vandenberghe, 2004, Section 5.1.2), in our case this amounts to

$$\mathbf{w} \mapsto \min_{r\geq 0, M\in\mathcal{L}(\pi)} \mathsf{L}(r, M, \mathbf{w}) = \sum_{i=1}^{n} w_i D_f(M_n^f || L_i)$$

is concave.

We proceed to prove item (4). Using item (1), we see that the pair $(M_n^f(\mathbf{w}_1, \{L_i\}_{i=1}^n, \pi), \overline{v}^f)$ minimizes the primal problem and $\mathbf{w}_2 = (w_{2,i})_{i=1}^n$ maximizes the dual problem, and hence by item (2), this gives

$$M_n^f(\mathbf{w}_1, \{L_i\}_{i=1}^n, \pi) \in \arg\min_{M\in\mathcal{L}(\pi)} \mathsf{L}(r, M, \mathbf{w}_2) = \arg\min_{M\in\mathcal{L}(\pi)} \sum_{i=1}^{n} w_{2,i} D_f(M||L_i).$$

As f is strictly convex, the minimization problem on the right hand side admits a unique minimizer $M_n^f(\mathbf{w}_2, \{L_i\}_{i=1}^n, \pi)$ (Theorem 2.1), and hence $M_n^f(\mathbf{w}_1, \{L_i\}_{i=1}^n, \pi) = M_n^f(\mathbf{w}_2, \{L_i\}_{i=1}^n, \pi)$.

Next, we prove item (5). We first show sufficiency. Using item (2), we note that

$$\overline{v}^f = D_f(M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi)||L_j) = \min_{M\in\mathcal{L}(\pi)} D_f(M||L_j) \leq \max_{l\in[\![n]\!]} \min_{M\in\mathcal{L}(\pi)} D_f(M||L_l) = \underline{v}^f.$$

This gives $\overline{v}^f = \underline{v}^f$ since $\overline{v}^f \geq \underline{v}^f$. To show necessity, let

$$j = \arg\max_{l\in[\![n]\!]} \min_{M\in\mathcal{L}(\pi)} D_f(M||L_l),$$

and we take $\mathbf{w}^f$ to be the standard unit vector at the j-th coordinate, that is, $w_j^f = 1$ and 0 otherwise. Since $\overline{v}^f = \underline{v}^f$, this choice of $\mathbf{w}^f$ maximizes the dual problem (3.5) and

$$M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi) \in \arg\min_{M\in\mathcal{L}(\pi)} D_f(M||L_j)$$

clearly holds.

Finally, we prove item (6). As the saddle point property holds we have

$$\overline{v}^f = \max_{i\in[\![n]\!]} D_f(M^f(L_i, \pi), L_i).$$

The stated pair in the theorem statement is thus a saddle point.

4.4. *Proof of Proposition 3.4.* First, as $D_f(\cdot||\cdot)$ is jointly convex in its arguments, and maximization of a convex function over a convex hull occur at an extreme point, we see that

$$\sup_{L\in\mathrm{conv}((L_i)_{i=1}^n)} D_f(M||L) = \max_{i\in[\![n]\!]} D_f(M||L_i).$$

Taking inf on both sides over $M \in \mathcal{L}(\pi)$ gives the desired result for $\overline{v}^f$.

For $\underline{v}^f$, we note that the mapping

$$L \mapsto \inf_{M\in\mathcal{L}(\pi)} D_f(M||L)$$

is convex as $D_f(\cdot||\cdot)$ is jointly convex and partial minimization of convex function preserves convexity (Boyd and Vandenberghe, 2004, Section 3.2.5). Using again the property that maximization of a convex function over a convex hull occur at an extreme point, we have

$$\underline{v}^f(\mathrm{conv}((L_i)_{i=1}^n), \pi) = \sup_{L\in\mathrm{conv}((L_i)_{i=1}^n)} \inf_{M\in\mathcal{L}(\pi)} D_f(M||L) = \max_{i\in[\![n]\!]} \inf_{M\in\mathcal{L}(\pi)} D_f(M||L_i) = \underline{v}^f(\{L_i\}_{i=1}^n, \pi).$$

4.5. *Proof of Corollary 3.6.* For the first item, it is simply a restatement of the saddle point result Corollary 3.5 in the context of the two-person game and Nash equilibrium.

For the second item, as the index l is unique, using the fact that f is strictly convex $M^f(L_l, \pi)$ is unique, and we thus have a unique saddle point and hence Nash equilibrium. For the other direction, suppose that $l_1, l_2 \in \arg\max_{i\in[\![n]\!]} D_f(M^f(L_i,\pi), L_i)$, then both $(M^f(L_{l_1},\pi), L_{l_1})$ and $(M^f(L_{l_2},\pi), L_{l_2})$ are saddle points or Nash equilibria.

4.6. *Proof of Proposition 3.10.* We first prove the inequalities in (3.11). The inequality $\overline{V}^f \geq \underline{V}^f$ is obvious. To see the first inequality, we note that

$$\int_{\mathcal{B}} D_f(M||L)\,\mu(dL) \leq \sup_{L\in\mathcal{B}} D_f(M||L).$$

The desired result follows by taking $\sup_{\mu\in\mathcal{P}(\mathcal{B})}$ then $\inf_{M\in\mathcal{L}(\pi)}$.

Next, we prove the inequalities in (3.12), and it only remains to show the rightmost inequality. We observe that

$$\begin{aligned}
\underline{V}^f(\mathcal{B},\pi) &= \sup_{\mu\in\mathcal{P}(\mathcal{B})} \inf_{M\in\mathcal{L}(\pi)} \int_{\mathcal{B}} D_f(M||L)\,\mu(dL) \\
&\geq \max_{\mathbf{w}\in\mathcal{S}_n} \inf_{M\in\mathcal{L}(\pi)} \sum_{i=1}^n w_i D_f(M||L_i) \\
&= \max_{\mathbf{w}\in\mathcal{S}_n} \sum_{i=1}^n w_i D_f(M_n^f||L_i),
\end{aligned}$$

which completes the proof.

4.7. *Proof of Theorem 3.16.* First, by the definition of weighted information centroid, we have

$$\sum_{i=1}^n w_i D_f(M_n^f(\mathbf{w},\{L_j\}_{j=1}^n,\pi)||L_i) \leq \sum_{i=1}^n w_i D_f(M_n^f(\mathbf{v},\{L_j\}_{j=1}^n,\pi)||L_i).$$

Multiplying both sides by -1 followed by substracting $h(\mathbf{v})$, we obtain

$$\begin{aligned}
&h(\mathbf{w}) - h(\mathbf{v}) \\
&\quad= -\sum_{i=1}^n w_i D_f(M_n^f(\mathbf{w},\{L_j\}_{j=1}^n,\pi)||L_i) + \sum_{i=1}^n v_i D_f(M_n^f(\mathbf{v},\{L_j\}_{j=1}^n,\pi)||L_i) \\
&\quad\geq \sum_{i=1}^n -D_f(M_n^f(\mathbf{v},\{L_j\}_{j=1}^n,\pi)||L_i)(w_i - v_i) \\
&\quad= \sum_{i=1}^n -D_f(M_n^f(\mathbf{v},\{L_j\}_{j=1}^n,\pi)||L_i)(w_i - v_i) + \sum_{i=1}^n D_f(M_n^f(\mathbf{v},\{L_j\}_{j=1}^n,\pi)||L_n)(w_i - v_i) \\
&\quad= \sum_{i=1}^n g_i(w_i - v_i),
\end{aligned}$$

where the second equality follows from $\mathbf{w}, \mathbf{v} \in \mathcal{S}_n$.

Next, we proceed to prove (3.16). First we note that for any $i \in [\![n]\!]$, we have

$$D_f(M_n^f(\mathbf{w}, \{L_j\}_{j=1}^n, \pi)||L_i) = \sum_{x\in\mathcal{X}} \pi(x) \sum_{y\in\mathcal{X}\setminus\{x\}} L_i(x,y) f\left(\frac{M_n^f(\mathbf{w}, \{L_j\}_{j=1}^n, \pi)(x,y)}{L_i(x,y)}\right)$$
$$\leq |\mathcal{X}| \sup_{\mathbf{v}\in\mathcal{S}_n;\ i\in[\![n]\!];\ L_i(x,y)>0} L_i(x,y) f\left(\frac{M_n^f(\mathbf{v}, \{L_j\}_{j=1}^n, \pi)(x,y)}{L_i(x,y)}\right),$$

and hence

$$\|\mathbf{g}\|_2^2 = \sum_{i=1}^n g_i^2 \leq \sum_{i=1}^n \max\left\{D_f(M_n^f(\mathbf{w}, \{L_j\}_{j=1}^n, \pi)||L_n)^2, D_f(M_n^f(\mathbf{w}, \{L_j\}_{j=1}^n, \pi)||L_i)^2\right\}$$
$$\leq n \max_{l\in[\![n]\!]} D_f(M_n^f(\mathbf{v}, \{L_j\}_{j=1}^n, \pi)||L_l)^2 \leq B.$$

4.8. *Proof of Theorem 3.17.* We first prove (3.17). For $i \in [\![t]\!]$, we have

$$\left\|\mathbf{w}^{(i+1)} - \mathbf{w}^f\right\|_2^2 \leq \left\|\mathbf{v}^{(i+1)} - \mathbf{w}^f\right\|_2^2$$
$$= \left\|\mathbf{w}^{(i)} - \mathbf{w}^f - \eta\mathbf{g}(\mathbf{w}^{(i)})\right\|_2^2$$
$$= \left\|\mathbf{w}^{(i)} - \mathbf{w}^f\right\|_2^2 + \eta^2\left\|\mathbf{g}(\mathbf{w}^{(i)})\right\|_2^2 - 2\eta\mathbf{g}(\mathbf{w}^{(i)})\cdot(\mathbf{w}^{(i)} - \mathbf{w}^f)$$
$$\leq \left\|\mathbf{w}^{(i)} - \mathbf{w}^f\right\|_2^2 + \eta^2 B - 2\eta\mathbf{g}(\mathbf{w}^{(i)})\cdot(\mathbf{w}^{(i)} - \mathbf{w}^f),$$

where the first inequality follows from the fact that $\mathbf{w}^{(i+1)}$ is the projection of $\mathbf{v}^{(i+1)}$ onto the simplex $\mathcal{S}_n$, while the second inequality follows from the definition of B in (3.16). Upon rearranging and using the definition of subgradient, this leads to

$$h(\mathbf{w}^{(i)}) - h(\mathbf{w}^f) \leq \mathbf{g}(\mathbf{w}^{(i)})\cdot(\mathbf{w}^{(i)} - \mathbf{w}^f)$$
$$= \frac{1}{2\eta}\left(\left\|\mathbf{w}^{(i)} - \mathbf{w}^f\right\|_2^2 - \left\|\mathbf{w}^{(i+1)} - \mathbf{w}^f\right\|_2^2\right) + \frac{\eta B}{2}.$$

Now, we sum i from 1 to t to give

$$\sum_{i=1}^t h(\mathbf{w}^{(i)}) - h(\mathbf{w}^f) \leq \frac{1}{2\eta}\left(\left\|\mathbf{w}^{(1)} - \mathbf{w}^f\right\|_2^2 - \left\|\mathbf{w}^{(t+1)} - \mathbf{w}^f\right\|_2^2\right) + \frac{\eta B t}{2}$$
$$\leq \frac{1}{2\eta}\left\|\mathbf{w}^{(1)} - \mathbf{w}^f\right\|_2^2 + \frac{\eta B t}{2} \leq \frac{1}{2\eta}n + \frac{\eta B t}{2},$$

where in the last inequality we use the fact that $\mathbf{w}^{(1)}, \mathbf{w}^f \in \mathcal{S}_n$. Dividing both sides by t followed by the convexity of h (see (3.14)) yields

$$h\left(\frac{1}{t}\sum_{i=1}^t \mathbf{w}^{(i)}\right) - h\left(\mathbf{w}^f\right) \leq \frac{1}{t}\left(\sum_{i=1}^t h(\mathbf{w}^{(i)}) - h(\mathbf{w}^f)\right) \leq \frac{n}{2\eta t} + \frac{\eta B}{2}.$$

The right hand side as a function of the stepsize η is minimized when we take $\eta = \sqrt{\frac{n}{tB}}$.

4.9. *Proof of Theorem 3.18.* We first state the following lemma in the setting of this Theorem, which is of independent interest. It can be considered as an extension of Csiszár (1972, equation (3.25)) to finite Markov chains:

Lemma 4.1. *There exists a constant* $\infty > C = C(f, \{L_i\}_{i=1}^n, \pi, n, \mathcal{X}) > 0$ *such that*

$$\begin{aligned} &D_{\mathrm{TV}}(M_n^f(\overline{\mathbf{w}}^t, \{L_i\}_{i=1}^n, \pi)||M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi)) \\ &\quad \leq C\left(\sum_{i=1}^n w_i^f D_f(M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi)||L_i) - \sum_{i=1}^n \overline{w}_i^t D_f(M_n^f(\overline{\mathbf{w}}^t, \{L_i\}_{i=1}^n, \pi)||L_i)\right) \\ &\quad = C\left(h\left(\overline{\mathbf{w}}^t\right) - h\left(\mathbf{w}^f\right)\right). \end{aligned}$$

Once we have Lemma 4.1, the desired result is obtained, since from Theorem 3.17 with the same choice of stepsize in Algorithm 1 we know that

$$D_{\mathrm{TV}}(M_n^f(\overline{\mathbf{w}}^t, \{L_i\}_{i=1}^n, \pi)||M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi)) \leq C\left(h\left(\overline{\mathbf{w}}^t\right) - h\left(\mathbf{w}^f\right)\right) = \mathcal{O}\left(\frac{1}{\sqrt{t}}\right).$$

It thus remains to prove Lemma 4.1. In the remaining of this proof, to minimize notational overhead we write

$$M_n^f(\overline{\mathbf{w}}^t) = M_n^f(\overline{\mathbf{w}}^t, \{L_i\}_{i=1}^n, \pi), \quad M_n^f(\mathbf{w}^f) = M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi).$$

Using the strict convexity of f and (3.18), according to Csiszár (1972, equation (3.25)) we have that

$$\begin{aligned} &D_{\mathrm{TV}}(M_n^f(\overline{\mathbf{w}}^t)||M_n^f(\mathbf{w}^f)) \\ &\quad \leq C\left(\sum_{i=1}^n \overline{w}_i^t D_f(M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi)||L_i) - \sum_{i=1}^n \overline{w}_i^t D_f(M_n^f(\overline{\mathbf{w}}^t, \{L_i\}_{i=1}^n, \pi)||L_i)\right) \\ &\quad \leq C\left(\max_{i\in[\![n]\!]} D_f(M_n^f(\mathbf{w}^f, \{L_i\}_{i=1}^n, \pi)||L_i) + h\left(\overline{\mathbf{w}}^t\right)\right) \\ &\quad = C\left(-h\left(\mathbf{w}^f\right) + h\left(\overline{\mathbf{w}}^t\right)\right), \end{aligned}$$

where the last equality follows from the complementary slackness in Theorem 3.3.

Acknowledgements

The authors gratefully acknowledge the careful reading and constructive comments of a reviewer that has improved the quality and presentation of the manuscript. Michael Choi would like to thank Laurent Miclo for raising a question about possible game-theoretic interpretations of $P_{-\infty}$ and P_{∞} when he visited Toulouse School of Economics. For a possible answer to this question, the readers are referred to Example 3.12.

References

Beck, A. *First-order methods in optimization*, volume 25 of *MOS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA (2017). ISBN 978-1-611974-98-0. MR3719240.

Bierkens, J. Non-reversible Metropolis-Hastings. *Stat. Comput.*, **26** (6), 1213–1228 (2016). MR3538633.

Boyd, S. and Vandenberghe, L. *Convex optimization*. Cambridge University Press, Cambridge (2004). ISBN 0-521-83378-7. MR2061575.

Candan, Ç. Chebyshev Center Computation on Probability Simplex With α-Divergence Measure. *IEEE Signal Process. Lett.*, **27**, 1515–1519 (2020). DOI: 10.1109/LSP.2020.3018661.

Choi, M. C. H. Metropolis-Hastings reversiblizations of non-reversible Markov chains. *Stochastic Process. Appl.*, **130** (2), 1041–1073 (2020). MR4046529.

Choi, M. C. H. and Huang, L.-J. On hitting time, mixing time and geometric interpretations of Metropolis-Hastings reversiblizations. *J. Theoret. Probab.*, **33** (2), 1144–1163 (2020). MR4091579.

Choi, M. C. H. and Wolfer, G. Systematic approaches to generate reversiblizations of Markov chains. *IEEE Trans. Inform. Theory*, **70** (5), 3145–3161 (2024). MR4740773.

Condat, L. Fast projection onto the simplex and the l_1 ball. *Math. Program.*, **158** (1-2, Ser. A), 575–585 (2016). MR3511394.

Csiszár, I. A class of measures of informativity of observation channels. *Period. Math. Hungar.*, **2**, 191–213 (1972). MR335152.

Diaconis, P., Holmes, S., and Neal, R. M. Analysis of a nonreversible Markov chain sampler. *Ann. Appl. Probab.*, **10** (3), 726–752 (2000). MR1789978.

Diaconis, P. and Miclo, L. On characterizations of Metropolis type algorithms in continuous time. *ALEA Lat. Am. J. Probab. Math. Stat.*, **6**, 199–238 (2009). MR2520174.

Eldar, Y. C., Beck, A., and Teboulle, M. A minimax Chebyshev estimator for bounded error estimation. *IEEE Trans. Signal Process.*, **56** (4), 1388–1397 (2008). MR2512472.

Fill, J. A. Eigenvalue bounds on convergence to stationarity for nonreversible Markov chains, with an application to the exclusion process. *Ann. Appl. Probab.*, **1** (1), 62–87 (1991). MR1097464.

Gushchin, A. A. and Zhdanov, D. A. A minimax result for f-divergences. In *From stochastic calculus to mathematical finance*, pp. 287–294. Springer, Berlin (2006). MR2233545.

Gustafson, P. A guided walk Metropolis algorithm. *Statistics and computing*, **8** (4), 357–364 (1998). DOI: 10.1023/A:1008880707168.

Haussler, D. A general minimax result for relative entropy. *IEEE Trans. Inform. Theory*, **43** (4), 1276–1280 (1997). MR1454958.

Haussler, D. and Opper, M. Mutual information, metric entropy and cumulative relative entropy risk. *Ann. Statist.*, **25** (6), 2451–2492 (1997). MR1604481.

Horn, R. A. and Johnson, C. R. *Matrix analysis*. Cambridge University Press, Cambridge, second edition (2013). ISBN 978-0-521-54823-6. MR2978290.

Hwang, C.-R., Hwang-Ma, S.-Y., and Sheu, S. J. Accelerating Gaussian diffusions. *Ann. Appl. Probab.*, **3** (3), 897–913 (1993). MR1233633.

Hwang, C.-R., Hwang-Ma, S.-Y., and Sheu, S.-J. Accelerating diffusions. *Ann. Appl. Probab.*, **15** (2), 1433–1444 (2005). MR2134109.

Jansen, S. and Kurt, N. On the notion(s) of duality for Markov processes. *Probab. Surv.*, **11**, 59–120 (2014). MR3201861.

Kamatani, K. and Song, X. Non-reversible guided Metropolis kernel. *J. Appl. Probab.*, **60** (3), 955–981 (2023). MR4624051.

Karlin, A. R. and Peres, Y. *Game theory, alive*. American Mathematical Society, Providence, RI (2017). ISBN 978-1-4704-1982-0. MR3642417.

Keilson, J. *Markov chain models—rarity and exponentiality*, volume 28 of *Applied Mathematical Sciences*. Springer-Verlag, New York-Berlin (1979). ISBN 0-387-90405-0. MR528293.

Maschler, M., Solan, E., and Zamir, S. *Game theory*. Cambridge University Press, Cambridge (2020). ISBN 978-1-108-49345-1; 978-1-108-82514-6. MR4411376.

Miclo, L. Remarques sur l'hypercontractivité et l'évolution de l'entropie pour des chaînes de Markov finies. In *Séminaire de Probabilités, XXXI*, volume 1655 of *Lecture Notes in Math.*, pp. 136–167. Springer, Berlin (1997). MR1478724.

Paulin, D. Concentration inequalities for Markov chains by Marton couplings and spectral methods. *Electron. J. Probab.*, **20**, no. 79, 32 (2015). MR3383563.

Rey-Bellet, L. and Spiliopoulos, K. Improving the convergence of reversible samplers. *J. Stat. Phys.*, **164** (3), 472–494 (2016). MR3519206.

Rosenthal, J. S. and Rosenthal, P. Spectral bounds for certain two-factor non-reversible MCMC algorithms. *Electron. Commun. Probab.*, **20**, no. 91, 10 (2015). MR3434208.

Schrödinger, E. Über die Umkehrung der Naturgesetze. *Sitzungsber. Preuss. Akad. Wiss., Phys.-Math. Kl.*, **8** (9), 144–153 (1931).

van Erven, T. and Harremoës, P. Rényi divergence and Kullback-Leibler divergence. *IEEE Trans. Inform. Theory*, **60** (7), 3797–3820 (2014). MR3225930.

van Dijk, N. M., van Brummelen, S., and Boucherie, R. Uniformization: Basics, extensions and applications. *Performance Evaluation*, **118**, 8–32 (2018). DOI: 10.1016/j.peva.2017.09.008.

Wolfer, G. and Watanabe, S. Information geometry of reversible Markov chains. *Inf. Geom.*, **4** (2), 393–433 (2021). MR4345345.