

Random Walk Approach to Regret Minimization

Authored by:

Hariharan Narayanan
Alexander Rakhlin

Abstract

We propose a computationally efficient random walk on a convex body which rapidly mixes to a time-varying Gibbs distribution. In the setting of online convex optimization and repeated games, the algorithm yields low regret and presents a novel efficient method for implementing mixture forecasting strategies.

1 Paper Body

This paper brings together two topics: online convex optimization and sampling from logconcave distributions over convex bodies. Online convex optimization has been a recent focus of research [30, 25], for it presents an abstraction that unifies and generalizes a number of existing results in online learning. Techniques from the theory of optimization (in particular, Fenchel and minimax duality) have proven to be key for understanding the rates of growth of regret [25, 1]. Deterministic regularization methods [3, 25] have emerged as natural black-box algorithms for regret minimization, and the choice of the regularization function turned out to play a pivotal role in limited-feedback problems [3]. In particular, the authors of [3] demonstrated the role of self-concordant regularization functions and the Dikin ellipsoid for minimizing regret. The latter gives a handle on the local geometry of the convex set, crucial for linear optimization with limited feedback. Random walks in a convex body gained much attention following the breakthrough paper of Dyer, Frieze and Kannan [9], who exhibited a polynomial time randomized algorithm for estimating the volume of a convex body. It is known that the problem of computing this volume by a deterministic algorithm is $\#P$ -hard. Over the two decades following [9], the polynomial dependence of volume computation on the dimension n has been drastically decreased from $O(n^3)$ to $O(n^4)$ [17]. The development was accomplished through the study of several geometric random walks: the Ball Walk and Hit-and-Run (see [26] for a survey). The driving force behind such results are the isoperimetric inequalities which can be extended from uniform to general logconcave distributions. In particular, computing the volume of a convex body can be seen as a special case of integration of a logconcave function,

and there has been a number of major results on mixing time for sampling from logconcave distributions [17, 18]. Connections to optimization have been established in [12, 18], among others. More recently, a novel random walk, called the Dikin Walk has been proposed in [19, 13]. By exploiting the local geometry of the set, this random walk is shown to mix rapidly, and offers a number of advantages over the other random walks. While the aim of online convex optimization is different from that of sampling from logconcave distributions, the fact that the two communities recognized the importance of the Dikin ellipsoid is remarkable. In this paper we build a bridge between the two topics. We show that the problem of online convex optimization can be solved by sampling from logconcave distributions, and that the Dikin Walk can be adapted to mix rapidly to a certain time-varying distribution. In fact, it mixes fast enough that for linear cost functions only one step of the guided Dikin Walk is necessary per round of the repeated game. This is surprisingly similar to the sufficiency of one Damped Newton step of Algorithm 2 in [3], due to locally quadratic convergence ensured by the self-concordant regularizer. 1

The time-varying Gibbs distributions from which we sample are closely related to Mixture Forecasters and Bayesian Model Averaging methods (see [7, Section 11.10] as well as [29, 28, 4, 10]). To the best of our knowledge, the method presented in this paper is the first provably computationally efficient approach to solving a class of problems which involves integrating over continuous sets of decisions. From the Bayesian point of view, our algorithm is an efficient procedure for sampling from posterior distributions, and can be used for settings outside of regret minimization. Prior work: The closest to our work is the result of [11] for Universal Portfolios. Unlike our onestep Markov chain, the algorithm of [11] works with a discretization of the probability simplex and requires a number of steps which has adverse dependence on the time horizon and accuracy. This seems unavoidable with the Grid Walk. In [2], it was shown that the Weighted Average Forecaster [15, 27] on a prohibitively large class of experts is optimal in terms of regret for a certain multitask problem, yet computationally inefficient. A Markov chain has been proposed with the required stationary distribution, but no mixing time bounds have been derived. In [8], the authors faced a similar problem whereby a near-optimal regret can be achieved by the Weighted Average Forecaster on a prohibitively large discretization of the set of decisions. Sampling from time-varying Markov chains has been investigated in the context of network dynamics [24], and has been examined from the point of view of linear stochastic approximation in reinforcement learning [14]. Beyond [11], we are not aware of any results to date where a provably rapidly mixing walk is used to solve regret minimization problems. It is worth emphasizing that without the Dikin Walk [19], the one-step mixing results of this paper seem out of reach. In particular, when sampling from exponential distributions, the known bounds for the conductance of the Ball Walk and Hit-and-Run are not scale-independent. In order to obtain $O(\sqrt{T})$ regret, one has to be able to sample the target distribution with an error that is $O(1/\sqrt{T})$. As a consequence of the deterioration of the bounds on the conductance as the scale tends to zero, the number of steps necessary per round would tend to

infinity as T tends to infinity.

2

Main Results

Let $K \subset \mathbb{R}^n$ be a convex compact set and let F be a set of convex functions from K to \mathbb{R} . Online convex optimization is defined as a repeated T -round game between the player (the algorithm) and Nature (adversary) [30, 25]. From the outset we assume that Nature is oblivious (see [7]), i.e. the individual sequence of decisions $x_1, \dots, x_T \in F$ can be fixed before the game. We are interested in randomized algorithms, and hence we consider the following online learning model: on round t , the player chooses a distribution (or, a mixed strategy) μ_t supported on K and plays a random $X_t \sim \mu_t$. Nature then reveals the cost function $\ell_t \in F$. The goal of the player is to control expected regret (see Lemma 1) with respect to a randomized strategy defined by a fixed distribution $\mu \in \mathcal{P}$ for some collection of distributions \mathcal{P} . If \mathcal{P} contains Dirac delta distributions, the comparator term is indeed the best fixed decision $x^* \in K$ chosen in hindsight. A procedure which guarantees sublinear growth of regret for any distribution $\mu \in \mathcal{P}$ will be called Hannan consistent with respect to \mathcal{P} . We now state a natural procedure for updating distributions μ_t which guarantees Hannan consistency for a wide range of problems. This procedure is similar to the Mixture Forecaster used in the t prediction context [29, 28, 4, 10]. Denote the cumulative cost functions by $L_t(x) = \sum_{s=1}^t \ell_s(x)$, with $L_0(x) = 0$, and let $\eta > 0$ be a learning rate. Let $q_0(x)$ be some prior probability distribution supported on K . Define the following sequence of functions $q_t(x) = q_0(x) \exp\{-\eta L_t(x)\}$,

$$\mu_t \propto \{1, \dots, T\}$$

for every $x \in K$. Define the probability distribution μ_t over K at time t to have density $\mu_t(x) \propto q_0(x) e^{-\eta L_t(x)}$ where $Z_t = \int_K q_t(x) dx$.

(1)

(2)

Let $D(p \parallel q)$ stand for the Kullback-Leibler divergence between distributions p and q . The following lemma gives an equality for expected regret with respect to a fixed randomized strategy. It bears 1 Due to its simplicity, the lemma has likely appeared in the literature, yet we could not locate a reference for this form with equality and in the context of online convex optimization. The closest results appear in [28, 10], [7, p. 326] in the context of prediction, and in [4] in the context of density estimation with exponential families.

2

striking similarity to upper bounds on regret in terms of Bregman divergences for the Follow the Regularized Leader and Mirror Descent methods [23, 5], [7, Thorem 11.1]. Lemma 1. Let X_t be a random variable distributed according to μ_t , for all $t \in \{1, \dots, T\}$, as defined in (2). Let U be a random variable with distribution μ . The expected regret is $\mathbb{E} \sum_{t=1}^T \ell_t(X_t) - \mathbb{E} \sum_{t=1}^T \ell_t(U) = \eta \sum_{t=1}^T (D(\mu_t \parallel \mu_0) - D(\mu_t \parallel \mu)) + \eta \sum_{t=1}^T D(\mu_t \parallel \mu)$.

$t=1$

$t=1$

$t=1$

Specializing to the case $l(x) \in [0, 1]$ over K , $\sum_{t=1}^T \mathbb{E} l(X_t) - \min_{x \in K} \sum_{t=1}^T l(x) \leq \frac{1}{8} \log T + \frac{1}{8}$.

Before proceeding, let us make a few remarks. First, if the divergence between the comparator μ distribution p_U and the prior μ_0 is bounded, the result yields $O(\sqrt{T})$ rates of regret growth for bounded losses by choosing μ appropriately. To bound the divergence between a continuous initial μ_0 and a point comparator at some x^* , the analysis can be carried out in two stages: comparison to a “small-covariance” Gaussian centered at x^* , followed by an observation that the loss of the “small-covariance” Gaussian strategy is not very different from the loss of the deterministic strategy x^* . This analysis can be found in [7, p. 326] and gives a near-optimal $O(\sqrt{T \log T})$ regret bound. We also note that for linear cost functions, the notion of expected regret coincides with regret for deterministic strategies. Third, we note that if the prior is of the form $q_0(x) \propto \exp\{R(x)\}$ for some convex function R , then $q_t(x) \propto \exp\{(\eta_t l(x) + R(x))\}$, bearing similarity to the objective function of the Follow the Regularized Leader algorithm [23, 3]. In general, we can encode prior knowledge in q_0 . For instance, if the cost functions are linear and the set K is a convex hull of N vertices (e.g. probability simplex), then the minimum loss is attained at one of the vertices, and a uniform prior on the vertices yields the Weighted Average Forecaster with the usual $\log N$ dependence [7]. Finally, we note that in online convex optimization, one of the difficulties is the issue of projections back to the set K . This issue does not arise when dealing with distributions, but instead translates into the difficulty of sampling. We find these parallels between sampling and optimization intriguing. We defer the easy proof of Lemma 1 to p. 8. Having a bound on regret, a natural question is whether there exists a computationally efficient algorithm for playing X_t according to the mixed strategy given in (2). The main result of this paper is that for linear Lipschitz cost functions the guided random walk (Algorithm 1 below) produces a sequence of points $X_1, \dots, X_T \in K$ with respective distributions μ_0, \dots, μ_T such that μ_i is close to μ_i for all $0 \leq i \leq T-1$. Moreover, X_i is obtained from X_{i-1} with only one random step. The step requires sampling from a Gaussian distribution with covariance given by the Hessian of the self-concordant barrier and can be implemented efficiently whenever the Hessian can be computed. The computation time exactly matches [3, Algorithm 2]: it is the same as time spent inverting a Hessian matrix, which is $O(n^3)$ or less. Let us now discuss our assumptions. First, the analysis of the random walk is carried out only for linear cost functions with a bounded Lipschitz constant. An analysis for general convex functions might be possible, but for the sake of brevity we restrict ourselves to the linear case. Note that convex cost functions can be linearized and a standard argument shows that regret for linearized functions can only be larger than that for the convex functions [30]. The second assumption is that q_0 does not depend on T and has a bounded L_2 norm with respect to the uniform distribution on K . This means that q_0 can be not only uniform, but, for instance, of the form $q_0(x) \propto \exp\{R(x)\}$. Theorem 2. Suppose $l_t : K \rightarrow [0, 1]$ are linear functions with Lipschitz constant 1 and the prior q_0 is of bounded

L_2 norm with respect to uniform distribution on K . Then the one-step random walk (Algorithm 1) produces a sequence X_1, \dots, X_T with distributions μ_0, \dots, μ_T such that for all i , $\|d_i(x) - d_{i+1}(x)\| \leq C n^{-1/2}$, $x \in K$

where d_i are defined in (2), ϵ is the parameter of self-concordance, and C is an absolute constant. Therefore, regret of Algorithm 1 is within $O(\epsilon T)$ from the ideal procedure of Lemma 1. In [3]

particular, by choosing ϵ appropriately, for an absolute constant C , $\forall T \geq 1$, $\mathbb{E} \sum_{t=1}^T \ell_t(X_t) \leq \sum_{t=1}^T \ell_t(U) + C \sqrt{n} \sum_{t=1}^T D(p_U \| \mu_t)$.

(3)

$t=1$

Proof. The statement follows directly from Lemma 1, Theorem 9, and an observation that for bounded losses ℓ_t , $\mathbb{E} \sum_{t=1}^T \ell_t(X_t) \leq \sum_{t=1}^T \mathbb{E} \ell_t(X_t) + \sum_{t=1}^T \mathbb{E} \|d_t(X_t) - d_{t+1}(X_t)\| \leq \sum_{t=1}^T \mathbb{E} \ell_t(X_t) + C n^{-1/2} \sum_{t=1}^T \mathbb{E} \|d_t(X_t) - d_{t+1}(X_t)\|$.

3

Sampling from a time-varying Gibbs distribution

Sketch of the Analysis The sufficiency of only one step of the random walk is made possible by the fact that the distributions μ_{t+1} and μ_t are close, and thus μ_{t+1} is a (very) warm start for μ_t . The reduction in distance between the distributions after a single step is due to a general fact (Lovász-Simonovits [16]) which we state in Theorem 6. The majority of the work goes into lower bounding the conductance of the random walk by a quantity independent of T (Lemma 5). Since the random walk of Algorithm 1 takes advantage of the local geometry of the set, the conductance is lower bounded by (a) proving an isoperimetric inequality (Theorem 3) for the Riemannian metric (which states that the measure of the gap between two well-separated sets is large) and (b) by proving that for close-by (in the Riemannian metric) points, their transition functions are not too different (Lemma 4). Section 3 is organized as follows. In Section 3.1, the main building blocks for proving mixing time are stated, and their proofs appear later in Section 4. In Section 3.2, we use the mixing result of Section 3.1 to show that Algorithm 1 indeed closely tracks the distributions μ_t (Theorem 9). 3.1

Bounding Mixing Time

In the remainder of this paper, C will denote a universal constant that may change from line to line. For any function F on the interior $\text{int}(K)$ having continuous derivatives of order k , for vectors $h_1, \dots, h_k \in \mathbb{R}^n$ and $x \in \text{int}(K)$, for $k \geq 1$, we recursively define

$D_{k+1} F(x + \sum_{i=1}^k h_i) = D_{k+1} F(x)[h_1, \dots, h_{k+1}]$, $\forall x \in \text{int}(K)$ where $D_0 F(x) := F(x)$. Let F be a self-concordant barrier of K with a parameter ϵ (see [20]). For $x, y \in K$, $\rho(x, y)$ is the distance in the Riemannian metric whose metric tensor is the Hessian of F . Thus, the metric tensor on the tangent space at x assigns to a vector v the length $\|v\|_x := \sqrt{2 D^2 F(x)[v, v]}$, and (to a pair of vectors v, w , the inner product $\langle v, w \rangle_x := D^2 F(x)[v, w]$. We have $\rho(x, y) = \inf \int \sqrt{2 D^2 F(z)} dz$ where the infimum is taken over all rectifiable paths γ from x to y . Let M be the metric space whose point set is K and metric is ρ . We assume ℓ_i are linear and $1/\epsilon$ -Lipschitz with respect to ρ . For $x \in \text{int}(K)$, let G_x denote the unique Gaussian probability density function on \mathbb{R}^n such that $\int G_x(z) dz = 1$.

* $n-x \cdot y^{-2x-1} G_x(y) \cdot \exp \left(-\frac{r}{2} V(x) \right)$, where $V(x) = \ln \det D^2 F(x)$ and $r = 1/(Cn) \cdot r_2 \cdot 2 \cdot D_k F(x)[h_1, \dots, h_k] := \lim$

Further, define the scaled cumulative cost as $st(y) := r_2 \cdot Lt(y)$. Note that shape of G_x is precisely given by the Dikin ellipsoid, which is defined as a unit ball in ℓ_2 -norm around a point x [20, 3].

The Markov chain M_t considered in this paper is such that for $x, y \in K$, one step $x \rightarrow y$ is given by Algorithm 1. A simple calculation shows that the detailed balance conditions are satisfied with respect to a stationary distribution π_t (defined in Eq. (2)). Therefore the Markov chain is reversible and has this stationary measure. The next results imply that this Markov chain is rapidly mixing. The first main ingredient is an isoperimetric inequality necessary for lower bounding conductance. Theorem 3. Let S_1 and S_2 be measurable subsets of K and π a probability measure supported on K that possesses a density whose logarithm is concave. Then, $1 - \pi((K \setminus S_1) \cap S_2) \leq \pi(S_1 \cap S_2) \cdot \pi(S_1 \setminus S_2) \cdot \pi(S_2 \setminus S_1) \cdot 2(1 + 3\pi(S_1 \cap S_2))^{-1}$.

Algorithm 1 One Step Random Walk (X_t, st) Input: current point $X_t \in K$ and scaled cumulative cost st . Output: next point $X_{t+1} \in K$

Toss a fair coin. If Heads, set $X_{t+1} := X_t$. Else,

K

X_{t+1}

Sample Z from G_{X_t} . If $Z \in K$, let $X_{t+1} := X_t$. If $Z \notin K$, let $\alpha, \tau \in \exp(st(X_t))$ with prob. $\min(1, \frac{G_{X_t}(Z)}{G_{X_t}(X_t)} \exp(\alpha(st(Z) - st(X_t))))$. $X_{t+1} := X_t$ otherwise.

X_t

Figure 1: The new point is sampled from a Gaussian distribution whose shape is defined by the local metric. Dotted lines are the unit Dikin ellipsoids.

The next Lemma relates the Riemannian metric γ to the Markov Chain. Intuitively, it says that for close-by points, their transition distributions cannot be far apart. Lemma 4. If $x, y \in K$ and $\gamma(x, y) \leq C$, then $d_T(V(\pi_x), \pi_y) \leq C \cdot \frac{1}{n}$. Theorem 3 and Lemma 4 together give a lower bound on conductance of the Markov Chain. Lemma 5 (Bound on Conductance). Let π be any exponential distribution on K . The conductance $\Phi(\pi, (K \setminus S_1) \cap S_2) := \inf_{S_1, S_2} \frac{\pi(S_1 \cap S_2)}{\pi(S_1) \pi(S_2)}$ of the Markov Chain in Algorithm 1 is bounded below by

$$\frac{1}{C} \cdot \frac{1}{n}$$

The lower bound on conductance of Lemma 5 can now be used with the following general result on the reduction of distance between distributions. Theorem 6 (Lovász-Simonovits [16]). Let π_0 be the initial distribution for a lazy reversible ergodic Markov chain whose conductance is Φ and stationary measure is π , and π_k be the distribution of π_0 at the k th step. Let $M := \sup_S \pi(S)$ where the supremum is over all measurable subsets S of K . For every bounded f , let $\|f\|_2^2 = \int f(x)^2 d\pi(x)$. For any fixed f , let $E_k f$ be the map that takes K to K $f(y) d\pi(y)$. Then if $\int f(x) d\pi(x) = 0$, $\|E_k f\|_2^2 \leq \frac{1}{k} \|f\|_2^2$. In summary, Lemma 5 provides a lower bound on conductance, while Theorem 6 ensures reduction of the norm whenever conductance is large enough. In the next section, these two are put together. We will show that reduction in the

norm guarantees that the distribution after one step of the random walk ($k = 1$ in Theorem 6) is close to the desired distribution μ_t . 3.2

Tracking the distributions

Let $\{\mu_i\}_{i=1}$ be the probability measures with bounded density, supported on K , corresponding to the distribution of a point during different steps of the evolution of the algorithm. For $i \in \mathbb{N}$, let $\|\cdot\|_i$ denote the L_2 norm with respect to the measure μ_i . We shall write $\|\cdot\|_i$ for brevity. Hence, $\int_K f^2 d\mu_i \leq 2 \int_K f d\mu_i$ for a measurable function $f: K \rightarrow \mathbb{R}$, $\|f\|_i^2 = \int_K f^2 d\mu_i$. Furthermore, $\int_K \mu_i(x) q_0(x) e^{\mu_i(x)} dx \leq \sup_{x \in K} \mu_i(x) \int_K q_0(x) e^{\mu_i(x)} dx \leq (1 + \epsilon) \int_K \mu_{i+1}(x) dx \leq 2 \int_K \mu_{i+1}(x) dx$

(4)

where we used the fact that $\mu_{i+1}(x) \leq 1$ and ϵ is an appropriate multiple of ϵ , e.g. $\epsilon = (e^2 - 1)\epsilon$ does the job. Analogously, $\int_K \mu_{i+1} d\mu_i \leq 1 + \epsilon$ over K . It then follows that the norms at time i and $i + 1$ are comparable: $\|f\|_i \leq (1 + \epsilon) \|f\|_{i+1}$ and $\|f\|_{i+1} \leq (1 + \epsilon) \|f\|_i$

(5)

The mixing results of Lemma 5 together with Theorem 6 imply Corollary 7. For any $i, 1 \leq i \leq n$, $\int_K \mu_{i+1} d\mu_i \leq 1 + \epsilon$ and $\int_K \mu_i d\mu_{i+1} \leq 1 + \epsilon$. $\int_K \mu_i d\mu_i \leq 1 + \epsilon$

Corollary 7 says that μ_{i+1} is closer to μ_i by a multiplicative constant. We now show that the distance of μ_{i+1} to μ_i is (additively) not much worse than its distance to μ_i . The multiplicative reduction in distance is shown to be dominating the additive increase, concluding the proof that μ_i is close to μ_i for all i (Theorem 9). Lemma 8. For any i , it holds that $\int_K \mu_{i+1} d\mu_i \leq 1 + \epsilon$ and $\int_K \mu_i d\mu_{i+1} \leq 1 + \epsilon$. Proof.

$\int_K \mu_{i+1} d\mu_i \leq 1 + \epsilon$ and $\int_K \mu_i d\mu_{i+1} \leq 1 + \epsilon$

$\int_K \mu_{i+1} d\mu_i \leq 1 + \epsilon$ and $\int_K \mu_i d\mu_{i+1} \leq 1 + \epsilon$

(6) (7)

We first establish a bound of $C\epsilon$ on (6). For any function $f: K \rightarrow \mathbb{R}$, let $f^+(x) = \max(0, f(x))$ and $f^-(x) = \min(0, f(x))$. By the triangle inequality, $\int_K \mu_{i+1} d\mu_i \leq \int_K \mu_{i+1}^+ d\mu_i + \int_K \mu_{i+1}^- d\mu_i$. Now, using (4) and (5), $\int_K \mu_{i+1}^+ d\mu_i \leq \int_K \mu_{i+1}^+ d\mu_{i+1} \leq 1 + \epsilon$ and $\int_K \mu_{i+1}^- d\mu_i \leq \int_K \mu_{i+1}^- d\mu_{i+1} \leq 1 + \epsilon$. Thus, $\int_K \mu_{i+1} d\mu_i \leq 1 + \epsilon$ and $\int_K \mu_i d\mu_{i+1} \leq 1 + \epsilon$.

Thus, (6) is bounded as $\int_K \mu_{i+1} d\mu_i \leq 1 + \epsilon$ and $\int_K \mu_i d\mu_{i+1} \leq 1 + \epsilon$. Next, a bound on (7) follows simply by the norm comparison inequality (5): $\int_K \mu_{i+1} d\mu_i \leq 1 + \epsilon$ and $\int_K \mu_i d\mu_{i+1} \leq 1 + \epsilon$.

The statement follows by rearranging the terms. $\int_K \mu_{i+1} d\mu_i \leq 1 + \epsilon$ and $\int_K \mu_i d\mu_{i+1} \leq 1 + \epsilon$, where $\epsilon = (e^2 - 1)\epsilon$, then for all i , $\int_K \mu_{i+1} d\mu_i \leq 1 + \epsilon$ and $\int_K \mu_i d\mu_{i+1} \leq 1 + \epsilon$.

(9)

$= 1$, and a direct computation shows that

$$dH(x, y) = \lim_{t \rightarrow 0} \frac{1}{t} \log \frac{P_t(x, y)}{P_t(x, x)} = \lim_{t \rightarrow 0} \frac{1}{t} \log \frac{P_t(x, y)}{P_t(x, x)}$$

Hence, using Theorem 10, the Hilbert metric and the Riemannian metric satisfy $dH(x, y) \leq 2(1 + 3\epsilon)dH(x, y)$.

The statement of the theorem is now an immediate consequence of the following result due to Lovász and Vempala [18]: If S_1 and S_2 are measurable subsets of K and μ a probability measure supported on K that possesses a density whose logarithm is concave, then $\mu((K \setminus S_1) \cap S_2) \leq \mu(S_1) \mu(S_2)$.

7

Proof of Lemma 5. Let S_1 be a measurable subset of K such that $\mu(S_1) \geq 1/2$ and $S_2 := K \setminus S_1$ be its complement. Let $S_1^* = S_1 \cap \{x : P_x(S_2) \geq 1/C\}$ and $S_2^* = S_2 \cap \{y : P_y(S_1) \geq 1/C\}$. By the reversibility of the chain, which is easily checked, $\int P_x(S_2) d\mu(x) = \int P_y(S_1) d\mu(y)$. S_1

S_2

If $x \in S_1^*$ and $y \in S_2^*$ then,

$$d_T(V(P_x, P_y)) \leq 1/C$$

"

min

$$K \cap \{x : P_x(S_2) \geq 1/C\}$$

)

$$\int P_y(S_1) d\mu(y) = \int P_x(S_2) d\mu(x) = \int P_x(S_2) d\mu(x)$$

then $d_T(V(P_x, P_y)) \leq 1/C$. Therefore $r(S_1^*, S_2^*) := \inf_{x \in S_1^*, y \in S_2^*} d_T(V(P_x, P_y)) \leq 1/C$. Therefore Theorem 3 implies that $r(S_1^*, S_2^*) \leq \min(\mu(S_1^*), \mu(S_2^*))$. $\mu((K \setminus S_1^*) \cap S_2^*) \leq \min(\mu(S_1^*), \mu(S_2^*)) \leq 2(1 + 3\epsilon)/C$. n Lemma 4 implies that if $\mu(x, y) \geq$

First suppose $\mu(S_1^*) \leq (1 - 1/C)\mu(S_1)$ and $\mu(S_2^*) \leq (1 - 1/C)\mu(S_2)$. Then, $\int C\mu(S_1^*) \leq \int C\mu(S_1) \leq \int P_x(S_2) d\mu(x) \leq \mu((K \setminus S_1^*) \cap S_2^*) \leq 2(1 + 3\epsilon)/C$ and we are done. Otherwise, without loss of generality, suppose $\mu(S_1^*) \geq (1 - 1/C)\mu(S_1)$ and $\mu(S_2^*) \leq (1 - 1/C)\mu(S_2)$.

$$\int C\mu(S_1^*) \leq \int C\mu(S_1) \leq \int P_x(S_2) d\mu(x) \leq \mu((K \setminus S_1^*) \cap S_2^*) \leq 2(1 + 3\epsilon)/C$$

(10)

Then

Proof of Lemma 1. We have that $\sum_{t=1}^n \log \frac{P_t(X_t)}{P_t(X_t)} = \log \frac{P_n(X_n)}{P_n(X_n)} = \log \frac{P_n(X_n)}{P_n(X_n)}$. Rearranging, canceling the telescoping terms, and using the fact that $Z_0 = 1$

$$\sum_{t=1}^n \log \frac{P_t(X_t)}{P_t(X_t)} = \log \frac{P_n(X_n)}{P_n(X_n)}$$

$$\sum_{t=1}^n \log \frac{P_t(X_t)}{P_t(X_t)} = \log \frac{P_n(X_n)}{P_n(X_n)}$$

$$\sum_{t=1}^n \log \frac{P_t(X_t)}{P_t(X_t)} = \log \frac{P_n(X_n)}{P_n(X_n)}$$

$$\sum_{t=1}^n \log \frac{P_t(X_t)}{P_t(X_t)} = \log \frac{P_n(X_n)}{P_n(X_n)}$$

Let U be a random variable with a probability distribution p_U . Then $\sum_{t=1}^n \log \frac{P_t(X_t)}{P_t(X_t)} = \log \frac{P_n(X_n)}{P_n(X_n)}$

and Computation, 108(2):212–261, 1994. [16] L. Lovász and M. Simonovits. Random walks in a convex body and an improved volume algorithm. *Random Structures and Algorithms*, 4(4):359–412, 1993. [17] L. Lovász and S. Vempala. Simulated annealing in convex bodies and an $O(n^4)$ volume algorithm. *J. Comput. Syst. Sci.*, 72(2):392–417, 2006. [18] L. Lovász and S. Vempala. The geometry of logconcave functions and sampling algorithms. *Random Struct. Algorithms*, 30(3):307–358, 2007. [19] H. Narayanan. Randomized interior point methods for sampling and optimization. *CoRR*, abs/0911.3950, 2009. [20] A.S. Nemirovskii. Interior point polynomial time methods in convex programming, 2004. [21] Y. E. Nesterov and A. S. Nemirovskii. *Interior Point Polynomial Algorithms in Convex Programming*. SIAM, Philadelphia, 1994. [22] Y.E. Nesterov and MJ Todd. On the Riemannian geometry defined by self-concordant barriers and interior-point methods. *Foundations of Computational Mathematics*, 2(4):333–361, 2008. [23] A. Rakhlin. Lecture notes on online learning, 2008. [http://stat.wharton.upenn.edu/~rakhlin/papers/online learning.pdf](http://stat.wharton.upenn.edu/~rakhlin/papers/online%20learning.pdf). [24] D. Shah and J. Shin. Dynamics in congestion games. In *Proceedings of ACM Sigmetrics*, 2010. [25] S. Shalev-Shwartz and Y. Singer. Convex repeated games and fenchel duality. In *NIPS*. 2007. [26] S. Vempala. Geometric random walks: A survey. In *Combinatorial and computational geometry*. *Math. Sci. Res. Inst. Publ.*, 52:577–616, 2005. [27] V. Vovk. Aggregating strategies. In *Proceedings of the Third Annual Workshop on Computational Learning Theory*, pages 372–383. Morgan Kaufmann, 1990. [28] V. Vovk. Competitive on-line statistics. *International Statistical Review*, 69:213–248, 2001. [29] K. Yamanishi. Minimax relative loss analysis for sequential prediction algorithms using parametric hypotheses. In *COLT’98*, pages 32–43, New York, NY, USA, 1998. ACM. [30] M. Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *ICML*, 2003.