Compositional Calculus of Regret Minimizers*

Gabriele Farina
Computer Science Department
Carnegie Mellon University
gfarina@cs.cmu.edu

Christian Kroer
IEOR Department
Columbia University
christian.kroer@columbia.edu

Tuomas Sandholm
Computer Science Department, CMU
Strategic Machine, Inc.
Strategy Robot, Inc.
Optimized Markets, Inc.
sandholm@cs.cmu.edu

Abstract

Regret minimization is a powerful tool for solving large-scale problems; it was recently used in breakthrough results for large-scale extensive-form game solving. This was achieved by composing simplex regret minimizers into an overall regret-minimization framework for extensive-form game strategy spaces. In this paper we study the general composability of regret minimizers. We derive a calculus for constructing regret minimizers for composite convex sets that are obtained from convexity-preserving operations on simpler convex sets. We show that local regret minimizers for the simpler sets can be combined with additional regret minimizers into an aggregate regret minimizer for the composite set. As one application, one can show that the CFR framework can be constructed easily from our framework. One can also show ways to include curtailing (constraining) operations into our framework. For one, they enable the construction of CFR generalization for extensive-form games with general convex strategy constraints that can cut across decision points.

1 Introduction

In this paper we introduce a general methodology for composing regret minimizers. We derive a set of rules for how regret minimizers can be constructed for composite convex sets via a *calculus* of regret minimization: given regret minimizers for convex sets $\mathcal{X}, \mathcal{Y}$ we show how to compose these regret minimizers for various convexity-preserving operations (e.g., intersection, convex hull, Cartesian product), in order to arrive at a regret minimizer for the resulting composite set.

Our approach treats the regret minimizers for individual convex sets as black boxes, and builds a regret minimizer for the resulting composite set by combining the outputs of the individual regret minimizers. This is important because it allows freedom in choosing the best regret minimizer for each individual set (from either a practical or theoretical perspective). For example, in practice the *regret matching* (Hart & Mas-Colell, 2000) and *regret matching*$^+$ (RM$^+$) (Tammelin et al., 2015) regret minimizers are known to perform better than theoretically-superior regret minimizers such as *Hedge* (Brown et al., 2017), while Hedge may give better theoretical results when trying to prove the convergence rate of a construction through our calculus.

*Parts of this short paper appeared at ICML 2019. Full paper available at http://www.cs.cmu.edu/~gfarina/2019/regret-circuits-icml19/regret-circuits.icml19.pdf

Abstract for: Smooth Games Optimization and Machine Learning Workshop (NeurIPS 2019), Vancouver, Canada.

One way to conceptually view our construction is as *regret circuits*: in order to construct a regret minimizer for some convex set $\mathcal{X}$ that consists of convexity-preserving operations on (say) two sets $\mathcal{X}_1, \mathcal{X}_2$, we construct a regret circuit consisting of regret minimizers for $\mathcal{X}_1$ and $\mathcal{X}_2$, along with a sequence of operations that aggregate the results of those circuits in order to form an overall circuit for $\mathcal{X}$. As an application, the correctness and convergence rate of the CFR algorithm can be proven easily through our calculus. The recent *Constrained CFR* algorithm (Davis et al., 2019) can also be constructed via our framework. Finally, this approach can be used to construct the first efficient regret minimizer for extensive-form correlated equilibrium in two-player general-sum games with no chance moves (Farina et al., 2019)

2 Regret Minimization

In online convex optimization (Zinkevich, 2003), a decision maker repeatedly interacts with an unknown environment by making a sequence of decisions $\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots$ from a convex and compact set $\mathcal{X} \subseteq \mathbb{R}^n$. After each decision $\boldsymbol{x}^t$, the decision maker faces a linear *loss function* $\ell^t(\boldsymbol{x}^t)$, which is unknown to the decision maker until after the decision is made. So, we are constructing a device that supports two operations: (i) it provides the next decision $\boldsymbol{x}^{t+1} \in \mathcal{X}$ and (ii) it receives/observes the linear loss function ℓ^t used to "evaluate" decision $\boldsymbol{x}^t$.

The *cumulative regret* measures the difference between the loss cumulated by the sequence of decisions $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^T$ and the loss that would have been cumulated by playing the best-in-hindsight time-independent decision $\hat{\boldsymbol{x}}$. Formally, the cumulative regret up to time T is

$$R^T_{(\mathcal{X},\mathcal{F})} := \sum_{t=1}^{T} \ell^t(\boldsymbol{x}^t) - \min_{\hat{\boldsymbol{x}} \in \mathcal{X}} \left\{ \sum_{t=1}^{T} \ell^t(\hat{\boldsymbol{x}}) \right\}. \tag{1}$$

The device is called a *regret minimizer* if it satisfies the desirable property of *Hannan consistency*: the average regret approaches zero, that is, $R^T_{(\mathcal{X},\mathcal{F})}$ grows *sublinearly* in T. Formally, in our notation, we have the following definition.

Definition 1 ($(\mathcal{X},\mathcal{F})$-regret minimizer)**.** *Let $\mathcal{X}$ be a convex and compact set and let $\mathcal{L}$ be the set of real linear functions on the domain $\mathcal{X}$. An $(\mathcal{X},\mathcal{L})$-*regret minimizer *is a function that selects the next decision $\boldsymbol{x}^{t+1} \in \mathcal{X}$ given the history of decisions $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^t$ and observed loss functions $\ell^1, \ldots, \ell^t \in \mathcal{L}$, so that the cumulative regret $R^T_{(\mathcal{X},\mathcal{L})} = o(T)$.*

Regret minimizers are useful for converging to convex-concave saddle-point problems. CFR is a very well-known regret minimizer for the strategy space of an extensive-form game.

3 Cartesian Product

In this section, we show how to combine an $(\mathcal{X},\mathcal{L})$- and a $(\mathcal{Y},\mathcal{L})$-regret minimizer to form an $(\mathcal{X}\times\mathcal{Y},\mathcal{L})$-regret minimizer. Any linear function $\ell : \mathcal{X}\times\mathcal{Y} \to \mathbb{R}$ can be written as $\ell(\boldsymbol{x},\boldsymbol{y}) = \ell_{\mathcal{X}}(\boldsymbol{x}) + \ell_{\mathcal{Y}}(\boldsymbol{y})$ where the linear functions $\ell_{\mathcal{X}} : \mathcal{X} \to \mathbb{R}$ and $\ell_{\mathcal{Y}} : \mathcal{Y} \to \mathbb{R}$ are defined as $\ell_{\mathcal{X}} : \boldsymbol{x} \mapsto \ell(\boldsymbol{x}, \boldsymbol{0})$ and $\ell_{\mathcal{Y}} : \boldsymbol{y} \mapsto \ell(\boldsymbol{0}, \boldsymbol{y})$. It is immediate to verify that

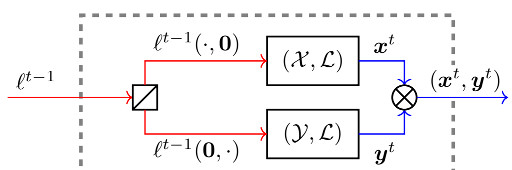

Figure 1: Regret circuit for the Cartesian product $\mathcal{X}\times\mathcal{Y}$.

$$R^T_{(\mathcal{X}\times\mathcal{Y},\mathcal{L})} = \left(\sum_{t=1}^{T} \ell^t_{\mathcal{X}}(\boldsymbol{x}^t) - \min_{\hat{\boldsymbol{x}} \in \mathcal{X}} \left\{ \sum_{t=1}^{T} \ell^t_{\mathcal{X}}(\hat{\boldsymbol{x}}) \right\} \right) + \left(\sum_{t=1}^{T} \ell^t_{\mathcal{Y}}(\boldsymbol{y}^t) - \min_{\hat{\boldsymbol{y}} \in \mathcal{Y}} \left\{ \sum_{t=1}^{T} \ell^t_{\mathcal{Y}}(\hat{\boldsymbol{y}}) \right\} \right) = R^T_{(\mathcal{X},\mathcal{L})} + R^T_{(\mathcal{Y},\mathcal{L})}.$$

In other words, it is possible to minimize regret on $\mathcal{X}\times\mathcal{Y}$ by simply minimizing it on $\mathcal{X}$ and $\mathcal{Y}$ independently and then combining the decisions, as in Figure 1.

4 Affine Transformation and Minkowski Sum

Let $H : E \to F$ be an affine map between two Euclidean spaces E and F, and let $\mathcal{X} \subseteq E$ be a convex and compact set. We now show how an $(\mathcal{X},\mathcal{L})$-regret minimizer can be employed to construct a

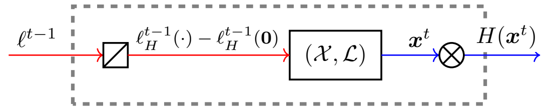

Figure 2: Regret circuit for the image $H(\mathcal{X})$ of $\mathcal{X}$ under the affine transformation H.

$(H(\mathcal{X}), \mathcal{L})$-regret minimizer. Since every $\boldsymbol{y} \in H(\mathcal{X})$ can be written as $\boldsymbol{y} = H(\boldsymbol{x})$ for some $\boldsymbol{x} \in \mathcal{X}$, the cumulative regret for an $(H(\mathcal{X}), \mathcal{L})$-regret minimizer can be expressed as

$$R^T_{(H(\mathcal{X}),\mathcal{L})} = \sum_{t=1}^{T} (\ell^t \circ H)(\boldsymbol{x}^t) - \min_{\hat{\boldsymbol{x}} \in \mathcal{X}} \left\{ \sum_{t=1}^{T} (\ell^t \circ H)(\hat{\boldsymbol{x}}) \right\}.$$

Since ℓ^t and H are affine, their composition $\ell^t_H := \ell^t \circ H$ is also affine. Hence, $R^T_{(H(\mathcal{X}),\mathcal{L})}$ is the same regret as an $(\mathcal{X}, \mathcal{L})$-regret minimizer that observes the linear function $\ell^t_H(\cdot) - \ell^t_H(\mathbf{0})$ instead of ℓ^t. The construction is summarized by the circuit in Figure 2. As an application, we can use the above construction to form a regret minimizer for the Minkowski sum $\mathcal{X} + \mathcal{Y} := \{\boldsymbol{x} + \boldsymbol{y} : \boldsymbol{x} \in \mathcal{X}, \boldsymbol{y} \in \mathcal{Y}\}$ of two sets.

5 Convex Hull

In this section, we show how to combine an $(\mathcal{X}, \mathcal{L})$- and a $(\mathcal{Y}, \mathcal{L})$-regret minimizer to form a $(\mathrm{co}\{\mathcal{X}, \mathcal{Y}\}, \mathcal{L})$-regret minimizer, where co denotes the convex hull operation, $\mathrm{co}\{\mathcal{X}, \mathcal{Y}\} = \{\lambda_1 \boldsymbol{x} + \lambda_2 \boldsymbol{y} : \boldsymbol{x} \in \mathcal{X}, \boldsymbol{y} \in \mathcal{Y}, (\lambda_1, \lambda_2) \in \Delta^2\}$, and $\Delta^2 := \{(\lambda_1, \lambda_2) \in \mathbb{R}^2_+ : \lambda_1 + \lambda_2 = 1\}$ is the two-dimensional *simplex*. We can think of a $(\mathrm{co}\{\mathcal{X}, \mathcal{Y}\}, \mathcal{L})$-regret minimizer as picking a triple $(\boldsymbol{\lambda}^t, \boldsymbol{x}^t, \boldsymbol{y}^t) \in \Delta^2 \times \mathcal{X} \times \mathcal{Y}$ at each time point t. One can show that in order to make "good decisions" in the convex hull $\mathrm{co}\{\mathcal{X}, \mathcal{Y}\}$, we can let two independent $(\mathcal{X}, \mathcal{L})$- and $(\mathcal{Y}, \mathcal{L})$-regret minimizers pick good decisions in $\mathcal{X}$ and $\mathcal{Y}$ respectively, and then use a third regret minimizer $(\Delta^2, \mathcal{L})$ that decides how to "mix" the two outputs. This way, we break the task of picking the next recommended triple $(\boldsymbol{\lambda}^t, \boldsymbol{x}^t, \boldsymbol{y}^t)$ into three different subproblems, two of which can be run independently. Figure 3 shows the regret circuit for the convex hull. The loss function ℓ^t_λ is defined as

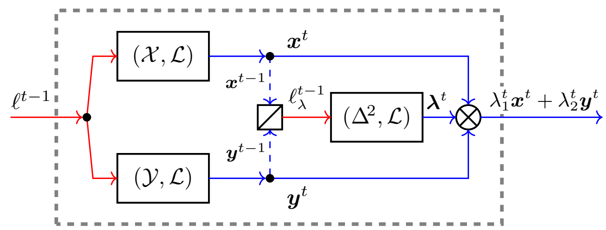

Figure 3: Regret circuit for the convex hull $\mathrm{co}\{\mathcal{X}, \mathcal{Y}\}$.

$$\ell^t_\lambda : \Delta^2 \ni (\lambda_1, \lambda_2) \mapsto \lambda_1 \ell^t(\boldsymbol{x}^t) + \lambda_2 \ell^t(\boldsymbol{y}^t).$$

One can show that the overall cumulative regret of the construction is

$$R^T_{(\mathrm{co}\{\mathcal{X},\mathcal{Y}\},\mathcal{L})} \le R^T_{(\Delta^2,\mathcal{L})} + \max\{R^T_{(\mathcal{X},\mathcal{L})}, R^T_{(\mathcal{Y},\mathcal{L})}\}.$$

The construction shown in Figure 3 can be extended to handle the convex hull $\mathrm{co}\{\mathcal{X}_1, \dots, \mathcal{X}_n\}$ of n sets as follows. First, the input loss function ℓ^{t-1} is fed into all the $(\mathcal{X}_i, \mathcal{L})$-regret minimizers $(i = 1, \dots, n)$. Then, the loss function ℓ^t_λ, defined as

$$\ell^t_\lambda : \Delta^n \ni (\lambda_1, \dots, \lambda_n) \mapsto \lambda_1 \ell(\boldsymbol{x}^t_1) + \cdots + \lambda_n \ell(\boldsymbol{x}^t_n),$$

is input into a $(\Delta^n, \mathcal{L})$-regret minimizer, where Δ^n is the n-dimensional simplex. Finally, at each time instant t, the n decisions $\boldsymbol{x}^t_1, \dots, \boldsymbol{x}^t_n$ output by the $(\mathcal{X}_i, \mathcal{L})$-regret minimizers are combined with the decision $\boldsymbol{\lambda}^t$ output by the $(\Delta^n, \mathcal{L})$-regret minimizer to form $\lambda^t_1 \boldsymbol{x}^t_1 + \cdots + \lambda^t_n \boldsymbol{x}^t_n$.

V-polytopes Our construction can be directly applied to construct an $(\mathcal{X}, \mathcal{L})$-regret minimizer for a V-polytope $\mathcal{X} = \mathrm{co}\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ where $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ are n points in a Euclidean space $\mathbb{E}$. Of course, any $(\{\boldsymbol{v}_i\}, \mathcal{L})$-regret minimizer outputs the constant decision $\boldsymbol{v}_i$. Hence, our construction (Figure 3) reduces to a single $(\Delta^n, \mathcal{L})$-regret minimizer that observes the (linear) loss function

$$\ell^t_\lambda : \Delta^n \ni (\lambda_1, \dots, \lambda_n) \mapsto \lambda_1 \ell^t(\boldsymbol{v}_1) + \cdots + \lambda_n \ell^t(\boldsymbol{v}_n).$$

The observation that a regret minimizer over a simplex can be used to minimize regret over a V-polytope already appeared in Zinkevich (2003), Schuurmans & Zinkevich (2016), and Farina et al. (2017, Theorem 3).

6 Application: Derivation of CFR

The strategy space of a single player in an extensive-form game is a treeplex, which can be viewed recursively as a series of convex hull and Cartesian product operations. This perspective is also used when constructing distance functions for first-order methods for EFGs (Hoda et al., 2010; Kroer et al., 2015, 2018).

Hence, one can apply our convex hull and Cartesian product constructions inductively to obtain a regret minimizer for the strategy space of a perfect-recall extensive-form game. The resulting algorithm is exactly *Counterfactual regret minimization (CFR)* (Zinkevich et al., 2007).

7 Intersection with a Closed Convex Set

In this section we consider constructing an $(\mathcal{X} \cap \mathcal{Y}, \mathcal{L})$-regret minimizer from an $(\mathcal{X}, \mathcal{L})$-regret minimizer, where $\mathcal{Y}$ is a closed convex set such that $\mathcal{X} \cap \mathcal{Y} \neq \emptyset$. As it turns out, this is always possible, and can be done by letting the $(\mathcal{X}, \mathcal{L})$-regret minimizer give decisions in $\mathcal{X}$, and then *projecting* them onto the intersection $\mathcal{X} \cap \mathcal{Y}$.

We will use a *Bregman divergence* $D(\boldsymbol{y}\|\boldsymbol{x}) := d(\boldsymbol{y}) - d(\boldsymbol{x}) - \langle \nabla d(\boldsymbol{x}), \boldsymbol{y} - \boldsymbol{x}\rangle$ as our notion of distance between the points x and y, where the *distance generating function* (DGF) d is μ-strongly convex and β-smooth (that is, d is differentiable and its gradient is Lipschitz continuous with Lipschitz constant β). Our construction makes no further assumptions on d, so the most appropriate DGF can be used for the application at hand. When $d(\boldsymbol{x}) = \|\boldsymbol{x}\|_2^2$ we obtain $D(\boldsymbol{y}\|\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{x}\|_2^2$, so we recover the usual Euclidean distance between $\boldsymbol{x}$ and $\boldsymbol{y}$. In accordance with our generalized notion of distance, we define the projection of a point $\boldsymbol{x} \in \mathcal{X}$ onto $\mathcal{X} \cap \mathcal{Y}$ as $\pi_{\mathcal{X}\cap\mathcal{Y}}(\boldsymbol{x}) = \mathrm{argmin}_{\boldsymbol{y}\in\mathcal{X}\cap\mathcal{Y}} D(\boldsymbol{y}\|\boldsymbol{x})$. For ease of notation, we will denote the projection of $\boldsymbol{x}$ onto $\mathcal{X} \cap \mathcal{Y}$ as $[\boldsymbol{x}]$; since $\mathcal{X} \cap \mathcal{Y}$ is closed and convex, and since $D(\cdot\|\boldsymbol{x})$ is strongly convex, such projection exists and is unique. The cumulative regret of the $(\mathcal{X} \cap \mathcal{Y}, \mathcal{L})$-minimizer is

$$R^T_{(\mathcal{X}\cap\mathcal{Y},\mathcal{L})} = \sum_{t=1}^T \ell^t([\boldsymbol{x}^t]) - \min_{\hat{\boldsymbol{x}}\in\mathcal{X}\cap\mathcal{Y}} \left\{\sum_{t=1}^T \ell^t(\hat{\boldsymbol{x}})\right\} = \sum_{t=1}^T \ell^t([\boldsymbol{x}^t]-\boldsymbol{x}^t) - \min_{\hat{\boldsymbol{x}}\in\mathcal{X}\cap\mathcal{Y}} \left\{\sum_{t=1}^T \ell^t(\hat{\boldsymbol{x}}-\boldsymbol{x}^t)\right\}, \quad (2)$$

where the second equality holds by linearity of ℓ^t. The first-order optimality condition for the projection problem is $\langle \nabla d(\boldsymbol{x}^t) - \nabla d([\boldsymbol{x}^t]), \hat{\boldsymbol{x}} - [\boldsymbol{x}^t]\rangle \le 0 \quad \forall\, \hat{\boldsymbol{x}} \in \mathcal{X} \cap \mathcal{Y}$. Consequently, provided $\alpha^t \ge 0$ for all t,

$$\min_{\hat{\boldsymbol{x}}\in\mathcal{X}\cap\mathcal{Y}} \left\{\sum_{t=1}^T \ell^t(\hat{\boldsymbol{x}} - \boldsymbol{x}^t)\right\} \ge \min_{\hat{\boldsymbol{x}}\in\mathcal{X}} \left\{\sum_{t=1}^T \ell^t(\hat{\boldsymbol{x}} - \boldsymbol{x}^t) + \sum_{t=1}^T \alpha^t \langle \nabla d(\boldsymbol{x}^t) - \nabla d([\boldsymbol{x}^t]), \hat{\boldsymbol{x}} - [\boldsymbol{x}^t]\rangle\right\} \quad (3)$$

(note the change in the domain of the minimum between the left- and right-hand side). The role of the α^t coefficients is to penalize choices of $\boldsymbol{x}^t$ that are in $\mathcal{X} \setminus \mathcal{Y}$. In particular, if

$$\frac{1}{\mu} \sum_{t=1}^T \ell_t([\boldsymbol{x}^t] - \boldsymbol{x}^t) \le \sum_{t=1}^T \alpha^t \|[\boldsymbol{x}^t] - \boldsymbol{x}^t\|^2, \quad (4)$$

then, by μ-strong convexity of d, we have

$$\sum_{t=1}^T \ell_t([\boldsymbol{x}^t]-\boldsymbol{x}^t) \le \sum_{t=1}^T \alpha^t \langle \nabla d(\boldsymbol{x}^t)-\nabla d([\boldsymbol{x}^t]), \boldsymbol{x}^t-[\boldsymbol{x}^t]\rangle. \quad (5)$$

Substituting (5) and (3) into Equation (2) we get

$$R^T_{(\mathcal{X}\cap\mathcal{Y},\mathcal{L})} \le \left(\sum_{t=1}^T \ell^t(\boldsymbol{x}^t) + \alpha^t\langle \nabla d(\boldsymbol{x}^t) - \nabla d([\boldsymbol{x}^t]), \boldsymbol{x}^t\rangle\right) - \min_{\hat{\boldsymbol{x}}\in\mathcal{X}} \left\{\sum_{t=1}^T \ell^t(\hat{\boldsymbol{x}}) + \alpha^t \langle \nabla d(\boldsymbol{x}^t) - \nabla d([\boldsymbol{x}^t]), \hat{\boldsymbol{x}}\rangle\right\},$$

which is the regret observed by an $(\mathcal{X}, \mathcal{L})$-regret minimizer that at each time t observes the linear loss function

$$\tilde{\ell}^t : \boldsymbol{x} \mapsto \ell^t(\boldsymbol{x}) + \alpha^t \langle \nabla d(\boldsymbol{x}^t) - \nabla d([\boldsymbol{x}^t]), \boldsymbol{x}\rangle. \quad (6)$$

Hence, as long as condition (4) holds, the regret circuit of Figure 4 is guaranteed to be Hannan consistent. On the other hand, condition (4) can be trivially satisfied by the deterministic choice

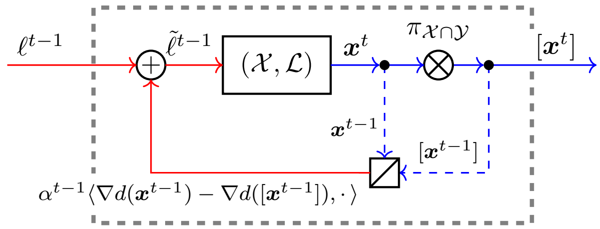

Figure 4: Regret circuit representing the construction of an $(\mathcal{X} \cap \mathcal{Y}, \mathcal{L})$-regret minimizer using a $(\mathcal{X}, \mathcal{L})$-regret minimizer.

$$\alpha^t = \begin{cases} 0 & \text{if } \boldsymbol{x}^t \in \mathcal{X} \cap \mathcal{Y} \\ \max\left\{0, \dfrac{\ell^t([\boldsymbol{x}^t] - \boldsymbol{x}^t)}{\mu \|[\boldsymbol{x}^t] - \boldsymbol{x}^t\|^2}\right\} & \text{otherwise.} \end{cases}$$

The fact that α^t can be arbitrarily large (when $\boldsymbol{x}^t$ and $[\boldsymbol{x}^t]$ are very close) is not an issue. Indeed, α^t is only used in $\tilde{\ell}^t$ (Equation 6) and is always multiplied by a term whose magnitude grows proportionally with the distance between $\boldsymbol{x}^t$ and $[\boldsymbol{x}^t]$. In fact, the norm of the functional $\tilde{\ell}^t$ is bounded:

$$\|\tilde{\ell}^t\| \le \|\ell^t\| + \left|\frac{\ell^t([\boldsymbol{x}^t] - \boldsymbol{x}^t)}{\mu \|[\boldsymbol{x}^t] - \boldsymbol{x}^t\|^2}\right| \cdot \|\nabla d(\boldsymbol{x}^t) - \nabla d([\boldsymbol{x}^t])\| \le \left(1 + \frac{\beta}{\mu}\right) \|\ell^t\|,$$

In other words, our construction dilates the loss functions by at most a factor $1 + \beta/\mu$. For instance, when $d(\boldsymbol{x}) = \|\boldsymbol{x}\|^2$, this dilation factor is equal to 2.

References

Brown, N., Kroer, C., and Sandholm, T. Dynamic thresholding and pruning for regret minimization. In *AAAI Conference on Artificial Intelligence (AAAI)*, 2017.

Davis, T., Waugh, K., and Bowling, M. Solving large extensive-form games with strategy constraints. In *AAAI Conference on Artificial Intelligence (AAAI)*, 2019.

Farina, G., Kroer, C., and Sandholm, T. Regret minimization in behaviorally-constrained zero-sum games. In *International Conference on Machine Learning (ICML)*, 2017.

Farina, G., Ling, C. K., Fang, F., and Sandholm, T. Efficient regret minimization algorithm for extensive-form correlated equilibrium. 2019.

Hart, S. and Mas-Colell, A. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68:1127–1150, 2000.

Hoda, S., Gilpin, A., Peña, J., and Sandholm, T. Smoothing techniques for computing Nash equilibria of sequential games. *Mathematics of Operations Research*, 35(2), 2010.

Kroer, C., Waugh, K., Kılınç-Karzan, F., and Sandholm, T. Faster first-order methods for extensive-form game solving. In *Proceedings of the ACM Conference on Economics and Computation (EC)*, 2015.

Kroer, C., Waugh, K., Kılınç-Karzan, F., and Sandholm, T. Faster algorithms for extensive-form game solving via improved smoothing functions. *Mathematical Programming*, pp. 1–33, 2018.

Schuurmans, D. and Zinkevich, M. A. Deep learning games. In *Advances in Neural Information Processing Systems*, pp. 1678–1686, 2016.

Tammelin, O., Burch, N., Johanson, M., and Bowling, M. Solving heads-up limit Texas hold'em. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI)*, 2015.

Zinkevich, M. Online convex programming and generalized infinitesimal gradient ascent. In *International Conference on Machine Learning (ICML)*, pp. 928–936, Washington, DC, USA, 2003.

Zinkevich, M., Bowling, M., Johanson, M., and Piccione, C. Regret minimization in games with incomplete information. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NIPS)*, 2007.