https://pubsonline.informs.org/journal/opre

OPERATIONS RESEARCH
Vol. 73, No. 5, September–October 2025, pp. 2430–2457
ISSN 0030-364X (print), ISSN 1526-5463 (online)

Crosscutting Areas

Better Regularization for Sequential Decision Spaces: Fast Convergence Rates for Nash, Correlated, and Team Equilibria

Gabriele Farina,[a] Christian Kroer,[b,*] Tuomas Sandholm[c,d,e,f]

[a] MIT EECS, Cambridge, Massachusetts 02139; [b] Industrial Engineering and Operations Research Department, Columbia University, New York, New York 10027; [c] Computer Science Department, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213; [d] Strategy Robot, Inc., Pittsburgh, Pennsylvania 15232; [e] Optimized Markets, Inc., Pittsburgh, Pennsylvania 15232; [f] Strategic Machine, Inc., Pittsburgh, Pennsylvania 15232
*Corresponding author
Contact: gfarina@mit.edu, https://orcid.org/0000-0002-3976-0061 (GF); ck2945@columbia.edu, https://orcid.org/0000-0002-9009-8683 (CK); sandholm@cs.cmu.edu, https://orcid.org/0000-0001-8861-9366 (TS)

Received: September 24, 2021
Revised: October 27, 2022; December 5, 2023; May 13, 2024; November 19, 2024
Accepted: November 26, 2024
Published Online in Articles in Advance: February 28, 2025

Area of Review: Market Analytics and Revenue Management

https://doi.org/10.1287/opre.2021.0633

Abstract. We study the application of iterative first-order methods to the problem of computing equilibria of large-scale extensive-form games. First-order methods must typically be instantiated with a regularizer that serves as a distance-generating function (DGF) for the decision sets of the players. In this paper, we introduce a new weighted entropy-based distance-generating function. We show that this function is equivalent to a particular set of new weights for the dilated entropy distance–generating function on a treeplex while retaining the simpler structure of the regular entropy function for the unit cube. This function achieves significantly better strong-convexity properties than existing weight schemes for the dilated entropy while maintaining the same easily implemented closed-form proximal mapping as the prior state of the art. Extensive numerical simulations show that these superior theoretical properties translate into better numerical performance as well. We then generalize our new entropy distance function, as well as general dilated distance functions, to the scaled extension operator. The scaled extension operator is a way to recursively construct convex sets, which generalizes the decision polytope of extensive-form games as well as the convex polytopes corresponding to correlated and team equilibria. Correspondingly, we give the first efficiently computable distance-generating function for all those strategy polytopes. By instantiating first-order methods with our regularizers, we achieve several new results, such as the first method for computing ex ante correlated team equilibria with a guaranteed $1/T$ rate of convergence and efficient proximal updates. Similarly, we show that our regularizers can be used to speed up the computation of correlated solution concepts.

Funding: G. Farina was supported by the National Science Foundations [Grant CCF-2443068] and by T. Sandholm's grants listed below and a Facebook fellowship. C. Kroer was supported by the Office of Naval Research [Grants N00014-22-1-2530 and N00014-23-1-2374] and the National Science Foundation [Grants IIS-2147361 and IIS-2238960]. T. Sandholm was supported by the Vannevar Bush Faculty Fellowship, Office of Naval Research [Grant ONR N00014-23-1-2876], the National Science Foundation Division of Information and Intelligent Systems [Grants RI-1718457, RI-2312342, RI-1901403, and CCF-1733556], the Army Research Office [Grants W911NF2010081 and W911NF2210266], and the National Institutes of Health [Grant A240108S001]. This work was further supported by the National Science Foundation Division of Information and Intelligent Systems [Grant 1617590] and the Army Research Office [Grant W911NF-17-1-0082].

Supplemental Material: All supplemental materials, including the code, data, and files required to reproduce the results, are available at https://doi.org/10.1287/opre.2021.0633.

Keywords: extensive-form games • first-order methods • distance-generating functions • equilibrium computation • adversarial team games

1. Introduction

Large-scale *extensive-form game (EFG)* models have been used in several recent AI milestones, where equilibrium approximation was used as the approach for building AI agents (Bowling et al. 2015; Brown and Sandholm 2017, 2019b; Moravčík et al. 2017). A crucial component for constructing these agents is a fast method for computing approximate Nash equilibria in large and very large game models. For the two-player zero-sum setting, an EFG can be solved in polynomial time using a linear program (LP) whose size is linear in the size of the game tree (Romanovskii 1962, von Stengel 1996). However, this LP-based approach was not used in any of these AI milestones. Instead, fast iterative methods are preferred (Zinkevich et al. 2007, Hoda et al. 2010, Bowling et al. 2015, Brown and Sandholm

2019a, Farina et al. 2019a, Kroer et al. 2020, Farina et al. 2021b) as well as sampling-based variants (Lanctot et al. 2009; Gibson et al. 2012; Kroer et al. 2015; Brown and Sandholm 2017, 2019b; Schmid et al. 2019; Farina et al. 2020b). The reason for this is that constructing the LP, and running simplex or interior-point methods on it, is too expensive for these large-scale models.[1] In contrast, iterative methods only require oracle access to one or two gradient computations, or even estimates thereof, in order to perform an iteration.

From a theoretical standpoint, the fastest iterative methods for solving two-player zero-sum games are *first-order methods* (FOMs) such as the excessive gap technique (Nesterov 2005a) or mirror prox (MP) (Nemirovski 2004), which converge at a rate of $1/T$, where T is the number of iterations.[2] In order to apply these methods to EFGs, they must be instantiated with a *distance-generating function* (DGF), which yields an appropriate notion of how to measure distances between strategies in the game (Hoda et al. 2010). The choice of DGF affects the FOM in two ways: first, the strong-convexity properties of the DGF, as well as the induced polytope diameter, affect the convergence rate in terms of the number of iterations. Second, each iteration requires solving one or more *proximal steps* that use the DGF to penalize large steps away from the current iterate. It is thus crucial that the chosen DGF has good strong-convexity properties but also that it is efficiently computable. This has been studied for several other classes of decision sets in the past, for example, the simplex case, where both the negative entropy DGF and the Euclidean DGF are known to have good properties (Held et al. 1974, Beck and Teboulle 2003, Condat 2016) and the case of positive semidefinite matrices with a trace constraint, where the matrix entropy performs well (Ben-Tal and Nemirovski 2005).

In the formulation of solving a two-player zero-sum EFG via FOMs, the convex set of all strategies belonging to a player is referred to as the *sequence-form polytope* and alternatively as a *treeplex* (Hoda et al. 2010), which is a treelike structure of scaled simplexes. Essentially the only sequence-form polytope DGFs that are known are based on the *dilated DGF framework* introduced by Hoda et al. (2010) (apart from using the standard ℓ_2 distance, which is unsuitable because of projection requirements at each iteration). For example, the dilated entropy distance yields the best current rate of convergence for $1/T$ methods (Kroer et al. 2020). One drawback of the dilated entropy DGF, as well as other dilated DGFs, is that current analyses incur a dependence of the form $2^{\mathfrak{D}}$, where $\mathfrak{D}$ is the depth of the decision space (Hoda et al. 2010; Kroer et al. 2020, 2018). In some cases, this is reasonable because the decision space itself may have exponential size in the depth of the game tree. However, in other cases, the decision space may have substantial structure such that this exponential complexity in depth makes the bounds exponential in the size of the game tree.

In this paper, we introduce the first DGF for sequence-form polytopes whose strong convexity is not derived from its structure as a dilated distance function (again, the standard Euclidean distance also satisfies this, but it requires difficult projections). In particular, we show that a weighted version of the negative entropy for the nonnegative unit cube is a superior DGF for sequence-form polytopes. First, we show that this DGF can achieve strong-convexity modulus $1/M_Q$ (where M_Q is the maximum value of the ℓ_1 norm on the sequence-form polytope Q), with the largest weights at individual decision points being on the order of $M_Q \log n$ (where $\log n$ is the largest number of actions at any decision point). This improves upon the diameter of the dilated entropy DGF by removing the exponential dependence on depth and only retaining a dependence on the ℓ_1-norm of the sequence-form polytope. This also translates into an improvement to the theoretical convergence rate of FOMs by a factor $2^{\mathfrak{D}+2}$. Importantly, this improvement does not come at the cost of giving up the fast—in particular, linear-time—computation of the associated proximal steps required by, for example, mirror prox or EGT. Indeed, we show that the weights in our new DGF are chosen in a way that allows us to show that this new DGF corresponds to a particular dilated entropy DGF on the sequence-form polytope (while being different outside the sequence-form polytope). This enables us to use existing results on fast proximal-step computation for dilated entropy (Hoda et al. 2010). We call our new DGF the *dilatable global entropy* (DGE).

After introducing DGE for sequence-form polytopes, we switch our focus to studying DGFs for the more general *scaled extension* operator (Farina et al. 2019d). The scaled extension operator is a method for iteratively constructing a convex set as a sequence of convexity-preserving compositions of convex sets. This operator can be used to construct the sequence-form polytope, but more importantly, for our purposes, it can also be used to construct more general sets such as the polytope of correlation plans needed for computing optimal extensive-form correlated equilibria and ex ante coordinated team strategies in certain classes of games. First, we show how to extend the class of dilated DGFs to polytopes constructed via scaled extension, thereby generalizing the framework of Hoda et al. (2010) beyond sequence-form polytopes while also giving a simpler proof of strong convexity. This enables DGFs such as the dilated entropy or dilated Euclidean distance to be applied to a significantly broader class of polytopes. Then, we show that our DGE construction can also be extended to scaled extension. Taken together, we generalize the entire class of known "nice" DGFs for sequence-form polytopes to the set of polytopes that can be constructed

via scaled extension. Applying these results to the problems of computing optimal correlated solution concepts and ex ante coordinated team strategies yields the *first* method for iteratively solving these problems at a rate of $1/T$ while enjoying fast closed-form solutions at each iteration. In contrast, the only prior result of this form required using the standard Euclidean distance and thus had to perform expensive projections at every iteration of the algorithm (Farina et al. 2019c). Furthermore, in cases where those solution concepts are known to be computable in polynomial time, we recover polynomial time complexity (in the game tree size) for each proximal step induced by our regularizers. In all cases, our regularizers enable proximal steps with time complexity linear in the smallest description known today for the polytope of correlation strategies.

Extensive experiments validate the efficacy of our new DGFs. We find that these new DGFs lead to significantly smaller amounts of smoothing while still ensuring the correctness of the algorithms. Intuitively, this means that we can safely take much larger steps at each iteration.

1.1. Paper Outline

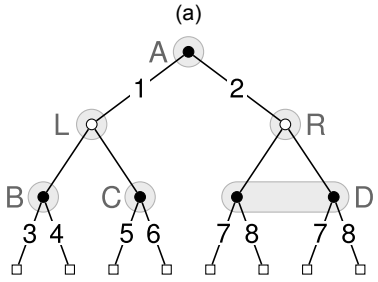
The paper is structured as follows. Section 2 gives a brief introduction to extensive-form games. We will formalize our problem directly in terms of the sequence-form polytope, which is given in Section 4. Section 3 presents a background on first-order methods, which includes the description of the DGFs needed for setting up these methods. That section can be skimmed for notation if the reader is already familiar with FOMs. Section 4 introduces the sequence-form polytope and dilated DGFs. Section 5 presents our new DGF for the sequence-form polytope, along with the convergence rate obtained when combined with an FOM. Section 6 develops DGFs for the scaled extension operator and shows how this leads to efficient FOMs with a $1/T$ convergence rate for correlated and team equilibria. Section 7 provides an extensive set of computational evaluations of our new DGFs for various games and types of equilibrium.

2. Preliminaries on Extensive-Form Games

An EFG is a game played on a tree. Every node in the tree belongs to some player, whose turn it is to act, and the set of branches at the node correspond to the set of actions available to the player. In general, a strategy for a player may consist of choosing a probability distribution over the actions at each node in the game. Additionally, there may be special nodes called *chance nodes*, which have a fixed distribution over actions associated with them. These nodes model stochastic outcomes, for example, the dealing out of cards in a card game or the valuation signals sent to buyers in a sequential auction. At leaf nodes, the game ends, and each leaf node is associated with a vector of payoffs, one payoff per player. The goal of each player in the game is to maximize the expected value of their leaf-node payoffs. Finally, an EFG can model imperfect information: an *information set* is a group of nodes belonging to a player such that the player cannot distinguish among those nodes and is therefore required to have the same probability distribution over actions at each node in the information set. An example of an information set would be in a poker game, where the information set represents all the cards seen by the player, as well as all bets (which are public). Each node in the information set would correspond to different possible hands held by the other player(s). Figure 1, which can be found in Section 4, shows a small example game.

A *solution concept* provides a definition of rationality. For a given EFG, the application of a solution concept yields a set of equilibria, where each equilibrium has one strategy per player. A strategy describes how players act at every one of their information sets. For example, a *Nash equilibrium* is a set of strategies such that each player cannot improve their expected utility

Figure 1. Example Extensive-Form Game and Associated Sequence-Form Constraints for One of the Players



(b)

Sequence-form constraints for Player 1:

$$\begin{cases} x_{\varnothing} = 1, \\ x_{A1} + x_{A2} = x_{\varnothing}, \\ x_{B3} + x_{B4} = x_{A1}, \\ x_{C5} + x_{C6} = x_{A1}, \\ x_{D7} + x_{D8} = x_{A2}. \end{cases}$$

Notes. (a) Example extensive-form games with two players. Black round nodes belong to player 1, white round nodes belong to player 2, and white square nodes are terminal states (payoffs are not shown). Nodes in the same information set are connected with a light gray background. The numbers on the edges uniquely identify actions for player 1. (b) The constraints that define the sequence-form polytope Q for player 1 (besides nonnegativity) in the game shown in (a).

by switching to another strategy when the strategies of other players are held fixed. We will introduce various solution concepts in the sections where we give algorithms for them.

3. Preliminaries on First-Order Methods

The types of FOMs that we will consider rely on access to a function d, which is used to construct a notion of distance between pairs of points in the decision space. The soundness of the algorithm requires such a function to satisfy a number of properties:

Definition 1 (Distance-Generating Function). A *DGF* d for a compact and convex set $\mathcal{X} \subseteq \mathbb{R}^n$ is a function $d : \mathcal{X} \to \mathbb{R}$ such that

- it is continuous on $\mathcal{X}$ and differentiable in the relative interior of $\mathcal{X}$;
- it is strongly convex in the relative interior of $\mathcal{X}$ with respect to some norm $\|\cdot\|$; that is, there exists a constant $\mu > 0$ such that

$$(\nabla d(\boldsymbol{x}) - \nabla d(\boldsymbol{x}'))^\top (\boldsymbol{x} - \boldsymbol{x}') \geq \mu \|\boldsymbol{x} - \boldsymbol{x}'\|^2 \quad \forall \boldsymbol{x}, \boldsymbol{x}' \in \text{relint } \mathcal{X}.$$

For twice-differentiable d, the strong-convexity condition is automatically verified as long as

$$\boldsymbol{m}^\top \nabla^2 d(\boldsymbol{x}) \boldsymbol{m} \geq \mu \|\boldsymbol{m}\|^2, \ \forall \boldsymbol{x} \in \text{relint } \mathcal{X}, \boldsymbol{m} \in \mathbb{R}^n; \quad (1)$$

- the minimizer of d (also called *prox-center*), $\arg\min_{x \in \mathcal{X}} d(\boldsymbol{x})$, belongs to relint $\mathcal{X}$.

Furthermore, we make the common assumption that

$$\min_{\boldsymbol{x} \in \mathcal{X}} d(\boldsymbol{x}) = 0. \quad (2)$$

This can always be assumed without loss of generality, as d can always be shifted by a constant amount without losing the other properties.

Given a convex set $\mathcal{X} \subseteq \mathbb{R}^n$ and a distance-generating function d for it, several important tools can be defined, which collectively form a *proximal setup* for $\mathcal{X}$:

- The *Bregman divergence* $D_d : \mathcal{X} \times \text{relint } \mathcal{X} \to \mathbb{R}_{\geq 0}$ associated with d yields a notion of distance between points defined as[3]

$$D_d(\boldsymbol{x} \| \boldsymbol{x}') := d(\boldsymbol{x}) - d(\boldsymbol{x}') - \nabla d(\boldsymbol{x}')^\top (\boldsymbol{x} - \boldsymbol{x}') \quad \forall \boldsymbol{x} \in \mathcal{X}, \boldsymbol{x}' \in \text{relint } \mathcal{X}.$$

- The *d-diameter of* $\mathcal{X}$ is

$$\Omega_{d,\mathcal{X}} := \max_{\boldsymbol{x} \in \mathcal{X}} d(\boldsymbol{x}) - \min_{\boldsymbol{x} \in \mathcal{X}} d(\boldsymbol{x}) = \max_{\boldsymbol{x} \in \mathcal{X}} d(\boldsymbol{x}),$$

where the last equality follows from (2). We remark that the prox-center $\boldsymbol{c} := \arg\min_{x \in \mathcal{X}} d(\boldsymbol{x})$ of d satisfies

$$\Omega_{d,\mathcal{X}} = \max_{\boldsymbol{x} \in \mathcal{X}} D_d(\boldsymbol{x} \| \boldsymbol{c}).$$

- Finally, we denote the largest possible value of the ℓ_1 norm on a $\mathcal{X}$ with the symbol

$$M_{\mathcal{X}} := \max_{\boldsymbol{x} \in \mathcal{X}} \|\boldsymbol{x}\|_1.$$

3.1. Nice Distance-Generating Functions

Whereas it is not a part of the assumptions on the DGF d, it is typically assumed that d allows one to efficiently compute the following two quantities, which come up at every iteration of most FOMS:

- the *gradient* $\nabla d(\boldsymbol{x})$ of d at any point $\boldsymbol{x} \in \text{relint } \mathcal{X}$;
- the gradient of the convex conjugate d^* of d at any point $\boldsymbol{g} \in \mathbb{R}^n$:

$$\nabla d^*(\boldsymbol{g}) = \arg\max_{\boldsymbol{x} \in \mathcal{X}} \{\boldsymbol{g}^\top \boldsymbol{x} - d(\boldsymbol{x})\}.$$

(We remark that the maximizer on the right-hand side exists, and it is unique because d is strongly convex by hypothesis.)

The gradient of the convex conjugate can be intuitively thought of as a linear maximization problem over $\mathcal{X}$ (i.e., the *support function* of $\mathcal{X}$, which is a nonsmooth convex optimization problem), *smoothed* by the regularizer d. For that reason, in this paper, we shall refer to $\nabla d^*(\boldsymbol{g})$ either symbolically or occasionally as the *smoothed support function*.

Because the above two quantities arise so frequently in optimization methods, it is important that the chosen distance-generating function allows for efficient computation of them. In particular, in this paper, we are concerned with nice DGFs that enable linear-time (in the dimension n) exact computation of those two quantities.

Definition 2. A distance-generating function d is said to be nice if $d(\boldsymbol{x})$, $\nabla d(\boldsymbol{x})$ and $\nabla d^*(\boldsymbol{g})$ can be computed exactly in linear time in the dimension of the domain of d.

Hoda et al. (2010) also introduce a notion of a nice DGF. Their definition is similar to ours but only states that $\nabla d^*(\boldsymbol{g})$ should be "easily computable." In contrast, we attach a concrete meaning to that statement: we take it to mean linear time in the dimension of the domain.

Finally, we mention a closely related operation that comes up often in optimization methods: the *proximal operator* (or *prox operator*), defined as

$$\begin{aligned} \text{prox}_d(\boldsymbol{g} \| \tilde{\boldsymbol{x}}) &:= \arg\min_{\boldsymbol{x} \in \mathcal{X}} \{\boldsymbol{g}^\top \boldsymbol{x} + D_d(\boldsymbol{x} \| \tilde{\boldsymbol{x}})\} \\ &= \nabla d^*(-\boldsymbol{g} + \nabla d(\tilde{\boldsymbol{x}})) \in \mathcal{X} \end{aligned} \quad (3)$$

for any $\tilde{\boldsymbol{x}} \in \mathcal{X}$ and $\boldsymbol{g} \in \mathbb{R}^n$. In light of (3), the prox operator can be implemented efficiently provided that ∇d and ∇d^* can. So, prox operators can be computed exactly in linear time in the dimension n for nice DGFs.

3.2. Bilinear Saddle-Point Problems

We will be interested in solving *bilinear saddle-point problems* (BSPPs), whose general form is

$$\min_{\boldsymbol{x} \in \mathcal{X}} \max_{\boldsymbol{y} \in \mathcal{Y}} \boldsymbol{x}^\top \boldsymbol{A} \boldsymbol{y}, \quad (4)$$

where $\boldsymbol{A} \in \mathbb{R}^{n \times m}$ and $\mathcal{X}, \mathcal{Y}$ are convex and compact sets. We will now present the EGT and mirror prox algorithms

for solving BSPPs. These algorithms depend on two proximal setups: one for $\mathcal{X}$ and one for $\mathcal{Y}$, denoted d_x and d_y, respectively. Let $\|\cdot\|_x$ and $\|\cdot\|_y$ be the norms associated with the strong convexity of d_x and d_y in the given proximal setup. The convergence rate then depends on the following *operator norm* of the payoff matrix $\boldsymbol{A}$:

$$\|\boldsymbol{A}\| := \max\{\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{y} : \|\boldsymbol{x}\|_x \le 1, \|\boldsymbol{y}\|_y \le 1\}.$$

We will primarily be concerned with DGFs that are strongly convex with respect to either the ℓ_1 or ℓ_2 norms. The magnitude of $\|\boldsymbol{A}\|$ is the primary way in which the norm matters: if both d_x and d_y are strongly convex with respect to the ℓ_2 norm, then $\|\boldsymbol{A}\|$ can be on the order of $\sqrt{nm}$, whereas if both are with respect to the ℓ_1 norm, then $\|\boldsymbol{A}\|$ is simply equal to its largest entry.

3.3. The Excessive Gap Technique

The *excessive gap technique* (EGT) is a first-order method introduced by Nesterov (2005b), and one of the primary applications is to solve BSPPs such as Equation (4). EGT assumes access to a proximal setup for $\mathcal{X}$ and $\mathcal{Y}$, with one-strongly-convex DGFs d_x, d_y, and constructs smoothed approximations of the optimization problems faced by the x and y players. Based on this setup, we formally state the EGT of Nesterov (2005a) in Algorithm 1. EGT alternatingly takes steps focused on decreasing one or the other smoothing parameter. These steps are called ShrinkX and ShrinkY in Algorithm 1.

Algorithm 1 (Excessive Gap Technique Algorithm)

function Initialize
  $t \leftarrow 0$
  $\mu_x^0 \leftarrow \|\boldsymbol{A}\|, \quad \mu_y^0 \leftarrow \|\boldsymbol{A}\|$
  $\tilde{\boldsymbol{x}} \leftarrow \arg\min_{\hat{x}\in\mathcal{X}} d_x(\hat{x})$
  $\boldsymbol{y}^0 \leftarrow \nabla d_y^*(\boldsymbol{A}^\top \tilde{\boldsymbol{x}}/\mu_y^0)$
  $\boldsymbol{x}^0 \leftarrow \text{prox}_{d_x}(\boldsymbol{A}\boldsymbol{y}^0/\mu_x^0 \| \tilde{\boldsymbol{x}})$

function Iterate
  $t \leftarrow t+1, \quad \tau \leftarrow 2/(t+2)$
  if *t is even* **then** ShrinkX
  else ShrinkY

function ShrinkX
  $\bar{\boldsymbol{x}} \leftarrow -\nabla d_x^*(-\boldsymbol{A}\boldsymbol{y}^{t-1}/\mu_x^{t-1})$
  $\hat{\boldsymbol{x}} \leftarrow (1-\tau)\boldsymbol{x}^{t-1} + \tau\bar{\boldsymbol{x}}$
  $\bar{\boldsymbol{y}} \leftarrow \nabla d_y^*(\boldsymbol{A}^\top \hat{\boldsymbol{x}}/\mu_y^{t-1})$
  $\tilde{\boldsymbol{x}} \leftarrow \text{prox}_{d_x}\left(\frac{\tau}{(1-\tau)\mu_x^{t-1}}\boldsymbol{A}\bar{\boldsymbol{y}} \middle\| \bar{\boldsymbol{x}}\right)$
  $\boldsymbol{x}^t \leftarrow (1-\tau)\boldsymbol{x}^{t-1} + \tau\tilde{\boldsymbol{x}}$
  $\boldsymbol{y}^t \leftarrow (1-\tau)\boldsymbol{y}^{t-1} + \tau\bar{\boldsymbol{y}}$
  $\mu_x^t \leftarrow (1-\tau)\mu_x^{t-1}$

function ShrinkY
  $\bar{\boldsymbol{y}} \leftarrow \nabla d_y^*(\boldsymbol{A}^\top \boldsymbol{x}^{t-1}/\mu_y^{t-1})$
  $\hat{\boldsymbol{y}} \leftarrow (1-\tau)\boldsymbol{y}^{t-1} + \tau\bar{\boldsymbol{y}}$
  $\bar{\boldsymbol{x}} \leftarrow -\nabla d_x^*(-\boldsymbol{A}\hat{\boldsymbol{y}}/\mu_x^{t-1})$
  $\tilde{\boldsymbol{y}} \leftarrow \text{prox}_{d_y}\left(\frac{-\tau}{(1-\tau)\mu_y^{t-1}}\boldsymbol{A}^\top \bar{\boldsymbol{x}} \middle\| \bar{\boldsymbol{y}}\right)$
  $\boldsymbol{y}^t \leftarrow (1-\tau)\boldsymbol{y}^{t-1} + \tau\tilde{\boldsymbol{y}}$
  $\boldsymbol{x}^t \leftarrow (1-\tau)\boldsymbol{x}^{t-1} + \tau\bar{\boldsymbol{x}}$
  $\mu_y^t \leftarrow (1-\tau)\mu_y^{t-1}$

Algorithm 1 shows how initial points are selected and the alternating steps and step sizes are computed. Nesterov (2005a) proves that the EGT algorithm converges at a rate of $O(1/T)$:

Theorem 1 (Nesterov 2005a, Theorem 6.3). *At every iteration $t \ge 1$ of the EGT algorithm, the solution $(\boldsymbol{x}^t, \boldsymbol{y}^t)$ satisfies $\boldsymbol{x}^t \in \mathcal{X}$, $\boldsymbol{y}^t \in \mathcal{Y}$, and*

$$\max_{\boldsymbol{y}\in\mathcal{Y}} (\boldsymbol{x}^t)^\top \boldsymbol{A}\boldsymbol{y} - \min_{\boldsymbol{x}\in\mathcal{X}} \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{y}^t \le \frac{4\|\boldsymbol{A}\|\sqrt{\Omega_{d_x,\mathcal{X}}\Omega_{d_y,\mathcal{Y}}}}{t+1}.$$

3.4. Mirror Prox

Next, we consider the *mirror prox* algorithm (Nemirovski 2004). Rather than construct smoothed approximations, mirror prox directly uses the DGFs to take first-order steps. Hence, the MP algorithm is best understood as an algorithm that operates on the product space $\mathcal{X}\times\mathcal{Y}$ directly. As such, in most analyses of the MP algorithm, a single 1-strongly convex DGF for the product space $\mathcal{X}\times\mathcal{Y}$ is required. To better align with the setup used for EGT, we will define the DGF for the product space $\mathcal{X}\times\mathcal{Y}$ starting from proximal setups for both $\mathcal{X}$ and $\mathcal{Y}$, with 1-strongly convex DGFs d_x, d_y with respect to norms $\|\cdot\|_x$ and $\|\cdot\|_y$, respectively. With this setup, it is immediate to see that the function

$$d : \mathcal{X}\times\mathcal{Y} \ni (\boldsymbol{x},\boldsymbol{y}) \longmapsto d_x(\boldsymbol{x}) + d_y(\boldsymbol{y})$$

is a DGF for the product space $\mathcal{X}\times\mathcal{Y}$, which is strongly convex with modulus one with respect to the norm $\|(\boldsymbol{x},\boldsymbol{y})\| := \sqrt{\|\boldsymbol{x}\|_x^2 + \|\boldsymbol{y}\|_y^2}$. Furthermore, each proximal step taken with respect to d can be expressed as two independent proximal steps with respect to d_x and d_y: for any $\boldsymbol{g}_x, \boldsymbol{g}_y$,

$$\begin{aligned}
&\text{prox}_d\left(\begin{pmatrix}\boldsymbol{g}_x\\ \boldsymbol{g}_y\end{pmatrix} \middle\| \begin{pmatrix}\tilde{\boldsymbol{x}}\\ \tilde{\boldsymbol{y}}\end{pmatrix}\right)\\
&= \arg\min_{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}\times\mathcal{Y}} \left\{ \begin{pmatrix}\boldsymbol{g}_x\\ \boldsymbol{g}_y\end{pmatrix}^\top \begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix} + D_d\left(\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix} \middle\| \begin{pmatrix}\tilde{\boldsymbol{x}}\\ \tilde{\boldsymbol{y}}\end{pmatrix}\right)\right\}\\
&= \begin{pmatrix} \arg\min_{\boldsymbol{x}\in\mathcal{X}}\{\boldsymbol{g}_x^\top \boldsymbol{x} + d_x(\boldsymbol{x}) - \nabla d_x(\tilde{\boldsymbol{x}})^\top \boldsymbol{x}\} \\ \arg\min_{\boldsymbol{y}\in\mathcal{Y}}\{\boldsymbol{g}_y^\top \boldsymbol{y} + d_y(\boldsymbol{y}) - \nabla d_y(\tilde{\boldsymbol{y}})^\top \boldsymbol{y}\} \end{pmatrix}\\
&= \begin{pmatrix} \text{prod}_{d_x}(\boldsymbol{g}_x \| \tilde{\boldsymbol{x}}) \\ \text{prox}_{d_y}(\boldsymbol{g}_y \| \tilde{\boldsymbol{y}}) \end{pmatrix}.
\end{aligned}$$

Similarly, the d-diameter of the product space $\mathcal{X} \times \mathcal{Y}$ is equal to the sum of the diameters of $\mathcal{X}$ and $\mathcal{Y}$ in their respective proximal setups. Finally, we note that the function

$$F : \mathcal{X} \times \mathcal{Y} \ni (\boldsymbol{x}, \boldsymbol{y}) \longmapsto \begin{pmatrix} \boldsymbol{A}\boldsymbol{y} \\ -\boldsymbol{A}^\top \boldsymbol{x} \end{pmatrix},$$

critical in the analysis of MP (Ben-Tal and Nemirovski 2001), satisfies the following bound relative to the dual norm $\|\cdot\|_*$ of $\|\cdot\|$:

$$\begin{aligned} &\left\| F\begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} - F\begin{pmatrix} \boldsymbol{x}' \\ \boldsymbol{y}' \end{pmatrix} \right\|_* \\ &= \sqrt{\|\boldsymbol{A}(\boldsymbol{y}-\boldsymbol{y}')\|_{*\boldsymbol{x}}^2 + \|\boldsymbol{A}^\top(\boldsymbol{x}-\boldsymbol{x}')\|_{*\boldsymbol{y}}^2} \\ &\le \sqrt{\left[\max_{\|\tilde{\boldsymbol{x}}\|_{\boldsymbol{x}} \le 1} \tilde{\boldsymbol{x}}^\top \boldsymbol{A}(\boldsymbol{y}-\boldsymbol{y}')\right]^2 + \left[\max_{\|\tilde{\boldsymbol{y}}\|_{\boldsymbol{y}} \le 1} (\boldsymbol{x}-\boldsymbol{x}')^\top \boldsymbol{A}\tilde{\boldsymbol{y}}\right]^2} \\ &\le \sqrt{\|\boldsymbol{A}\|^2 \cdot \|\boldsymbol{y}-\boldsymbol{y}'\|_{\boldsymbol{y}}^2 + \|\boldsymbol{A}\|^2 \cdot \|\boldsymbol{x}-\boldsymbol{x}'\|_{\boldsymbol{x}}^2} = \|\boldsymbol{A}\| \cdot \left\| \begin{pmatrix} \boldsymbol{x}-\boldsymbol{x}' \\ \boldsymbol{y}-\boldsymbol{y}' \end{pmatrix} \right\|; \end{aligned}$$

that is, it is $\|\boldsymbol{A}\|$-Lipschitz with respect to the norm $\|\cdot\|$ on $\mathcal{X} \times \mathcal{Y}$.

Algorithm 2 shows the sequence of steps taken in every iteration of the MP algorithm. Compared with EGT, mirror prox has a somewhat simpler structure: it simply takes repeated extrapolated proximal steps. First, a proximal step in the descent direction is taken for both x and y. Then, the gradient at those new points is used to take a proximal step starting from the previous iterate (this is the extrapolation part: a step is taken starting from the previous iterate, but with the extrapolated gradient). Finally, the *average* strategy is output.

Algorithm 2 (Mirror Prox Algorithm)

function INITIALIZE()
  $t \leftarrow 0$
  $z_x^0 \leftarrow \arg\min_{\hat{\boldsymbol{x}} \in \mathcal{X}} d_x(\hat{\boldsymbol{x}})$
  $z_y^0 \leftarrow \arg\min_{\hat{\boldsymbol{y}} \in \mathcal{Y}} d_y(\hat{\boldsymbol{y}})$
function ITERATE()
  $t \leftarrow t + 1$
  $\boldsymbol{w}_x^t \leftarrow \text{prox}_{d_x}(\eta^t \boldsymbol{A} \boldsymbol{z}_y^{t-1} \| \boldsymbol{z}_x^t)$
  $\boldsymbol{w}_y^t \leftarrow \text{prox}_{d_y}(-\eta^t \boldsymbol{A}^\top \boldsymbol{z}_x^{t-1} \| \boldsymbol{z}_y^t)$
  $\boldsymbol{z}_x^{t+1} \leftarrow \text{prox}_{d_x}(\eta^t \boldsymbol{A} \boldsymbol{w}_y^t \| \boldsymbol{z}_x^t)$
  $\boldsymbol{z}_y^{t+1} \leftarrow \text{prox}_{d_y}(-\eta^t \boldsymbol{A}^\top \boldsymbol{w}_x^t \| \boldsymbol{z}_y^t)$
  $\boldsymbol{x}^t \leftarrow [\sum_{\tau=1}^t \eta^\tau]^{-1} \sum_{\tau=1}^t \eta^\tau \boldsymbol{w}_x^\tau$
  $\boldsymbol{y}^t \leftarrow [\sum_{\tau=1}^t \eta^\tau]^{-1} \sum_{\tau=1}^t \eta^\tau \boldsymbol{w}_y^\tau$

Note: $\{\eta^t\}$ is a sequence of step-size parameters. A well-known and theoretically sound choice for η^t is $\eta^t := \frac{1}{\|A\|}$ for all $t = 0, 1, \ldots$ (see also Theorem 2).

As we recall in the next theorem, like EGT, the mirror prox algorithm converges at rate $O(1/T)$.

Theorem 2 (Ben-Tal and Nemirovski 2001, Theorem 5.5.1). Suppose *the step size in Algorithm 2 is set as* $\eta_t = 1/\|A\|$. *Then, we have*

$$\max_{\boldsymbol{y} \in \mathcal{Y}} (\boldsymbol{x}^t)^\top \boldsymbol{A}\boldsymbol{y} - \min_{\boldsymbol{x} \in \mathcal{X}} \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{y}^t \le \frac{\|\boldsymbol{A}\|(\Omega_{d_x,\mathcal{X}} + \Omega_{d_y,\mathcal{Y}})}{2t}.$$

4. The Sequence-Form Polytope

Conceptually, a strategy for a player is an assignment of probability distributions over the actions available at each node. This representation of strategies, known as *behavioral* representation, has the drawback of making basic quantities such as the expected utility of a player a nonlinear function in the player's strategy. This is because, in this representation, the probability of following any terminal path of actions, a key quantity when defining expected utility, requires taking the *product* of several strategy variables (that is, the probabilities of the individual actions on the path).

A different representation of strategies, known as the *sequence-form* representation, is able to circumvent this limitation (Romanovskii 1962, Koller et al. 1996, von Stengel 1996). The idea of the sequence form is to associate strategy variables to *products* of probabilities of actions along each possible path from the root, and not to each individual action at each decision point. By operating this change of variables, each player's expected utility becomes a multilinear function of all the players' sequence-form strategies. On the flip side, this change of variables makes the set of valid strategies more intricate to define. Nonetheless, it is known that the set of valid strategies available to each player is a low-dimensional convex polytope, which we call the *sequence-form polytope* of that player.

For each player (aka, agent), we denote with $\mathcal{J}$ the set of their of information sets, which we will often call *decision points*. As mentioned earlier, information sets group together nodes that the player cannot distinguish given the information they have acquired thus far in the game. Because a player always knows what actions are available at a node, any two nodes at the same decision point j must have the same action set. We denote the set of actions available to the player when playing at (any node belonging to) j as A_j.

Throughout the paper, we make the standard assumption that the game has *perfect recall*; that is, the information sets of each player never encode a loss of acquired information. A well-known consequence of perfect recall is that the decision points $\mathcal{J}$ can be ordered into a forest. More precisely, let $j' \prec j$ (read "j' is an ancestor of j") if $j \ne j'$, and there exist nodes $h \in j, h' \in j'$ such that the path from the root of the game tree to h passes through h'. Under the perfect recall assumption, the

partially ordered set $(\mathcal{J}, \prec)$ is a forest, that is, a disjoint union of trees. The relations $\preceq$, $\succ$, and $\succeq$ are derived from the relation $\prec$ as usual. In particular, $j \succeq j'$ denotes that j is in the subtree of $(\mathcal{J}, \prec)$ rooted in j'. We use the symbol $\mathfrak{D}$ to denote the maximum depth of any element j in the forest $(\mathcal{J}, \prec)$.

A pair ja consisting of a decision point j and action a is referred to as a *sequence*. The special element $\varnothing$ is called the *empty* sequence. We use the symbol $\Sigma := \{\varnothing\} \cup \{ja : j \in \mathcal{J}, a \in A_j\}$ to denote the set of all sequences, including the empty one. Given a decision point $j \in \mathcal{J}$, we denote its parent sequence, defined as the last sequence (decision point–action pair) encountered on the path from the root to j with the symbol p_j. If the player does not act before j, then we conventionally let p_j be set to the empty sequence $\varnothing$. If the agent takes a given action $a \in A_j$ at decision point j, then $\mathcal{C}_{ja} \subset \mathcal{J}$ denotes the set of next potential decision points that the agent may face (which may be empty if no more decisions can occur after taking action a at j). More generally, for any $\sigma \in \Sigma$,

$$\mathcal{C}_\sigma := \{j : p_j = \sigma\}.$$

Example 1. In the extensive-form game of Figure 1(a), player 1 has four decision points corresponding to the information sets $\mathcal{J} = \{\mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D}\}$. The partial ordering $\prec$ between the decision points is $\mathsf{A} \prec \mathsf{B}, \mathsf{A} \prec \mathsf{C}, \mathsf{A} \prec \mathsf{D}$. The player has nine sequences, including the empty sequence $\varnothing$. For decision point D, the parent sequence is $p_{\mathsf{D}} = \mathsf{A2}$; for B, it is $p_{\mathsf{B}} = \mathsf{A1}$; and for A, it is the empty sequence $p_{\mathsf{A}} = \varnothing$. The set of children decision points $\mathcal{C}_{\mathsf{A1}}$ of sequence $\mathsf{A1}$ is $\{\mathsf{B}, \mathsf{C}\}$; the set of children decision points $\mathcal{C}_\varnothing$ of the empty sequence is the singleton set $\{\mathsf{A}\}$. The set of children of sequence $\mathsf{B3}$ is empty.

As mentioned at the beginning of this section, the sequence-form representation represents a strategy as a vector $\boldsymbol{x} \in \mathbb{R}^{\Sigma}_{\geq 0}$ whose entries are indexed over Σ. The entry corresponding to sequence $ja \in \Sigma$ contains the *product* of the probabilities of all actions at all decision points on the path from the root down to action a at decision point j included. To be a valid sequence-form strategy, the vector $\boldsymbol{x} \in \mathbb{R}^{\Sigma}_{\geq 0}$ must satisfy the constraints

$$x_\varnothing = 1, \quad \text{and} \quad \sum_{a \in A_j} x_{ja} = x_{p_j} \quad \forall j \in \mathcal{J}. \tag{5}$$

We call the set of all valid sequence-form strategies (that is, all vectors $\boldsymbol{x} \in \mathbb{R}^{\Sigma}_{\geq 0}$ that satisfy the constraints in (5)) the *sequence-form polytope* of the player.

Example 2. The sequence-form constraints (5) for player 1 in the small game of Figure 1(a) are shown in Figure 1(b).

For a two-player zero-sum EFG with perfect recall, the problem of computing a Nash equilibrium can be cast as a BSPP in the form of Equation (4). In this formulation, $\mathcal{X}$ and $\mathcal{Y}$ are the sequence-form polytopes for player 1 and player 2, respectively. The *payoff matrix* $\boldsymbol{A}$ is such that for a pair of sequence-form strategies $\boldsymbol{x}, \boldsymbol{y}$, the objective $\boldsymbol{x}^\top \boldsymbol{A} \boldsymbol{y}$ is equal to the expected value achieved by the second player under those strategies. Thus, the second player wishes to maximize this objective, whereas the first player wishes to minimize it. Each cell in $\boldsymbol{A} \in \mathbb{R}^{\Sigma_1 \times \Sigma_2}$, where Σ_1 and Σ_2 denote the sets of sequences of player 1 and player 2, respectively, corresponds to a pair of sequences, one for each player. The matrix is often sparse: each nonzero entry corresponds to a pair of sequences such that they are the last sequences on the path to some leaf node (and thus, we have zeroes for all cells such that the corresponding sequences are never the last pair of sequences before the game ends). The value at a cell is the payoff to the second player at that leaf times the product of all chance probabilities on the path to the leaf.

4.1. Euclidean Projections onto Sequence-Form Polytopes

The standard Euclidean DGF applied to the overall polytope Q is not nice, as it requires performing a somewhat complicated recursive variant of the sorting trick for projecting onto a simplex (Condat 2016). Gilpin et al. (2012) show that standard Euclidean projection can be computed exactly by first constructing a strictly monotonic, piecewise-linear function that depends on the specific structure of the game tree. In particular, this piecewise-linear function is constructed recursively in a bottom-up fashion, assigning to each information set j a piecewise-linear function that is obtained by combining the functions for the information sets that follow; that is, $\cup_{a \in A_j} \mathcal{C}_{ja}$. Each piecewise-linear function is represented in memory explicitly via its ordered list of breakpoints and corresponding slopes. The number of breakpoints required by the function associated with each information set j is linear in the sum of the number of actions at all descendants $j' \succeq j$. Gilpin et al. (2012) do not state a runtime for obtaining these breakpoints. However, to combine the functions associated with the immediate descendants $\cup_{a \in A_j} \mathcal{C}_{ja}$ of j, the breakpoints of the functions to be combined can be merge sorted, and an additional linear number of operations in the total amount of breakpoints is required (Farina et al. 2022a). Hence, the recursive construction of the piecewise-linear functions associated with the information sets cumulatively requires $O(|\Sigma| \mathfrak{D} \log |\Sigma|)$ operations, where we recall that $\mathfrak{D}$ denotes the depth of the forest $(\mathcal{J}, \prec)$. Hence, the time needed to compute the Euclidean projection onto the sequence-form polytope is generally significantly worse than computing a proximal step using a dilated DGF and even requires a potentially prohibitively quadratic time in $|\Sigma|$ when $\mathfrak{D} = \Theta(|\Sigma|)$. We refer the reader to Farina et al. (2022a) for additional computational details and a slight generalization of the projection algorithm described above.

4.2. Preliminaries on Dilated Distance–Generating Functions

Dilated distance–generating functions are a general framework for constructing nice DGFs (in the sense of Section 3.1) for (the relative interior of the) sequence-form polytopes (Hoda et al. 2010). Specifically, a dilated DGF for a sequence-form polytope is constructed by taking a weighted sum over *local* distance-generating functions $d_\varnothing$ and d_j ($j \in \mathcal{J}$) and is of the form

$$d : Q \ni \boldsymbol{x} \longmapsto \alpha_\varnothing d_\varnothing(x_\varnothing) + \sum_{j \in \mathcal{J}} \alpha_j d_j^\square(x_{p_j}, (x_{ja})_{a \in A_j}), \tag{6}$$

where

$$d_j^\square(x_{p_j}, \boldsymbol{x}) := \begin{cases} 0 & \text{if } x_{p_j} = 0 \\ x_{p_j} d_j\left(\dfrac{(x_{ja})_{a \in A_j}}{x_{p_j}}\right) & \text{otherwise.} \end{cases} \tag{7}$$

Each local function $d_j : \Delta^{A_j} \to \mathbb{R}$ is assumed to be continuously differentiable and strongly convex modulus one on the relative interior of the probability simplex Δ^{A_j}. By dividing $(x_{ja})_{a \in A_j}$ by x_{p_j} in (7), we renormalize $(x_{ja})_{a \in A_j}$ to the simplex, measure the DGF there, and then scale that value back by x_{p_j}. Finally, the weight α_j is a flexible-weight term that can be chosen to ensure good properties. Hoda et al. (2010) showed that if each local DGF d_j is strongly convex, then the dilated DGF d is also strongly convex (although they do not give an explicit modulus), and they show that the associated smoothed support function can easily be computed, provided that the smoothed support function for each d_j can easily be computed.

The gradient of a dilated DGF and its convex conjugate can be computed exactly in closed form by combining the gradients of each d_j and their convex conjugates, as shown in Algorithm 1. Finally, we remark that the computation of the *prox-center* of the DGF is a special case of the computation of the gradient of the convex conjugate because, by definition,

$$\nabla d^*(\mathbf{0}) = \arg\max_{x \in \mathcal{X}} \{-d(\boldsymbol{x})\} = \arg\min_{x \in \mathcal{X}} d(x).$$

Algorithm 3 (Gradient and Smoothed Support Function Implementation for General Dilated DGFs)

function GRADIENT($\boldsymbol{x} \in \text{relint } Q$)
  $\boldsymbol{g} \leftarrow \mathbf{0} \in \mathbb{R}^\Sigma$
  for $j \in \mathcal{J}$ *in bottom-up order* **do**
    $(g_{ja})_{a \in A_j} \leftarrow (g_{ja})_{a \in A_j} + \alpha_j \nabla d_j\left(\frac{(x_{ja})_{a \in A_j}}{x_{p_j}}\right)$
    $g_{p_j} \leftarrow g_{p_j} + \alpha_j d_j\left(\frac{(x_{ja})_{a \in A_j}}{x_{p_j}}\right)$
    $g_{p_j} \leftarrow g_{p_j} - \alpha_j \nabla d_j\left(\frac{(x_{ja})_{a \in A_j}}{x_{p_j}}\right)^\top \left(\frac{(x_{ja})_{a \in A_j}}{x_{p_j}}\right)$
  $g_\varnothing \leftarrow g_\varnothing + \alpha_\varnothing \nabla d_\varnothing(x_\varnothing)$
  return $\boldsymbol{g}$

function CONJUGATEGRADIENT($\boldsymbol{g} \in \mathbb{R}^\Sigma$)
  $\boldsymbol{z} \leftarrow \mathbf{0} \in \mathbb{R}^\Sigma$
  $z_\varnothing \leftarrow 1$
  for $j \in \mathcal{J}$ *in bottom-up order* **do**
    $(z_{ja})_{a \in A_j} \leftarrow \nabla d_j^*((g_{ja})_{a \in A_j})$
    $g_{p_j} \leftarrow g_{p_j} - d_j((z_{ja})_{a \in A_j}) + \sum_{a \in A_j} g_{ja} z_{ja}$
  for $j \in \mathcal{J}$ *in top-down order* **do**
    for $a \in A_j$ **do**
      $z_{ja} \leftarrow z_{p_j} \cdot z_{ja}$
  return $\boldsymbol{z}$

The local DGFs must be chosen so that they are compatible with the relative interior of the simplex. For a given simplex Δ^k, these are usually chosen either as the *entropy DGF* $d(\boldsymbol{y}) = \log k + \sum_i y_i \log y_i$ (we let $y_i \log y_i = 0$ whenever $y_i = 0$) or *Euclidean DGF* $d(\boldsymbol{y}) = \frac{1}{2} \sum_i (y_i - 1/k)^2$. These are both 1-strongly convex on relint Δ^k (for entropy with respect to the ℓ_1 norm and for Euclidean with respect to the ℓ_2 norm), and their associated smoothed support functions can be computed in $O(k)$ time (see, e.g., Ben-Tal and Nemirovski 2001 and Condat 2016).

One of the most important properties of dilated DGFs is that they lead to a nice DGF as long as each local convex conjugate gradient ∇d_j^* can be computed in time linear in $|A_j|$. In particular, this makes the dilated entropy and dilated Euclidean DGFs nice DGFs. This is in contrast to the standard Euclidean distance, as discussed in Section 4.1.

For dilated DGFs, the strongest general result on the strong-convexity modulus comes from Farina et al. (2019b), where the authors show that if each local DGF d_j is strongly convex modulus one with respect to the ℓ_2 norm and the weights of the decision points are set recursively according to $\alpha_j = 2 + 2\max_{a \in A_j} \sum_{j' \in \mathcal{C}_{ja}} \alpha_{j'}$ (so, in particular, decision points without descendants have a weight of two), then d is strongly convex modulus one with respect to the ℓ_2 norm on Q.

4.3. Preliminaries on the Dilated Entropy Distance–Generating Function

The dilated *entropy* DGF is the instantiation of the general dilated DGF framework of Section 4.2 with the particular choice of using the (negative) entropy function at each decision node. In particular, for any choice of weights $\alpha_\varnothing, \alpha_j > 0$ is the regularizer of the form[4]

$$Q \ni \boldsymbol{x} \longmapsto \alpha_\varnothing x_\varnothing \log x_\varnothing + \sum_{j \in \mathcal{J}} \alpha_j \left(x_{p_j} \log |A_j| + \sum_{a \in A_j} x_{ja} \log \left(\frac{x_{ja}}{x_{p_j}} \right) \right). \tag{8}$$

We will now briefly review existing results specific to the dilated entropy DGF, for which stronger results are known than for the general class of dilated DGFs.

First, as a direct consequence of the more general discussion in Section 4.2 and Algorithm 1, the dilated entropy DGF is a nice DGF (in the precise sense of Section 3.1) no matter the choice of weights α. In particular, in the case of the negative entropy functions $d_j(\boldsymbol{x}) = \log|A_j| + \sum_{a\in A_j} x_{ja}\log x_{ja}$, one has

$$(\nabla d_j(\boldsymbol{x}))_a = 1 + \log x_a, \quad (\nabla d_j^*(\boldsymbol{g}))_a = \frac{e^{g_a}}{\sum_{a'\in A_j} e^{g_{a'}}}$$
$$\forall a \in A_j, \boldsymbol{x} \in \operatorname{relint}\Delta^{A_j}, \boldsymbol{g} \in \mathbb{R}^{A_j}.$$

By plugging the above expression in the template of Algorithm 1, we obtain linear-time exact algorithms to compute $\nabla\varphi$ and $\nabla\varphi^*$. Furthermore, the prox-center of the DGF is the uniform strategy (that is, the strategy that assigns to each action a at any decision point j probability $1/|A_j| > 0$), and it therefore belongs to relint $Q = Q \cap \mathbb{R}^{\Sigma}_{>0}$.

Kroer et al. (2020) show that the dilated entropy DGF is strongly convex modulus $1/M_Q$ with respect to the ℓ_1 norm, when the weights α are chosen as in the following definition.

Definition 3 (Kroer et al. Dilated Entropy DGF, ψ). Define the DGF weights β_j recursively as

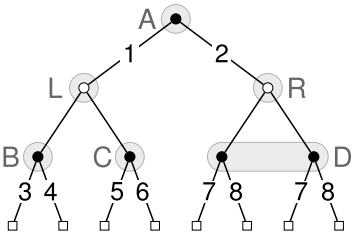
$$\beta_\varnothing := 2 + 2\sum_{j\in\mathcal{C}_\varnothing}\beta_j, \qquad \beta_j := 2 + 2\max_{a\in A_j}\sum_{j'\in\mathcal{C}_{ja}}\beta_{j'}, \qquad \forall j \in \mathcal{J}.$$

The resulting instantiation of the dilated entropy DGF is called the *Kroer et al. dilated entropy DGF* and denoted ψ:

$$\psi : Q \ni \boldsymbol{x} \mapsto \beta_\varnothing x_\varnothing \log x_\varnothing + \sum_{j\in\mathcal{J}}\beta_j\left(x_{p_j}\log|A_j| + \sum_{a\in A_j} x_{ja}\log\left(\frac{x_{ja}}{x_{p_j}}\right)\right).$$

Example 3. Continuing the example started in Figure 1 (reproduced at the top of the next column), in the case of the sequence-form polytope of player 1, the weights β in the construction of the Kroer et al.'s dilated entropy DGF ψ are computed as follows: $\beta_D = \beta_C = \beta_B = 2$; $\beta_A = 2 + 2\max\{\beta_B + \beta_C, \beta_D\} = 10$; $\beta_\varnothing = 2 + 2\beta_A = 22$. Correspondingly, the DGF ψ is the function

$$\begin{aligned}\psi(\boldsymbol{x}) = {} & 22\,x_\varnothing\log x_\varnothing + 10\,(x_\varnothing\log(2) + x_{A1}\log(x_{A1}/x_\varnothing) \\ & + x_{A2}\log(x_{A2}/x_\varnothing)) + 2\,(x_{A1}\log(2) \\ & + x_{B3}\log(x_{B3}/x_{A1}) + x_{B4}\log(x_{B4}/x_{A1})) \\ & + 2\,(x_{A1}\log(2) + x_{C5}\log(x_{C5}/x_{A1}) \\ & + x_{C6}\log(x_{C6}/x_{A1})) + 2\,(x_{A2}\log(2) \\ & + x_{D7}\log(x_{D7}/x_{A2}) + x_{D8}\log(x_{D8}/x_{A2})).\end{aligned}$$



Remark 1. On the surface, the strong-convexity modulus of $\frac{1}{M_Q}$ with respect to the ℓ_1 norm might appear less appealing than the modulus 1 obtained by using the ℓ_2 norm. However, recall that the norm that is used to measure strong convexity affects the value of the operator norm of A, which is significantly smaller under the $\ell_1 - \ell_\infty$ operator norm (where it is equal to $\max_{ij}|A_{ij}|$) than the $\ell_2 - \ell_2$ operator norm for strong convexity with respect to the ℓ_2 norm.

One drawback of both the general and entropy-specific dilated DGFs developed in the past is that they have an exponential dependence on the depth of the sequence-form polytope. In particular, note that the factor of two in the recursive definition of the weights means that the factor β_j for some root decision point is growing at least on the order of $2^{\mathfrak{D}_Q}$, where $\mathfrak{D}_Q$ is the depth of the tree-form sequential decision process. For many sequence-form polytopes, this might be acceptable: if the tree-form sequential decision process is reasonably balanced, then the number of decision points is also exponential in depth. However, for other sequence-form polytopes, this would be unacceptable: the most extreme case would be a single line of decision points, where the number of decision points is linear in $\mathfrak{D}_Q$, but the β_j at the root is exponentially large. This exponential dependence on depth also enters the convergence rate of the FOMs because it effectively acts as a scalar on the polytope diameter Ω induced by $\mathfrak{D}_Q$. The present paper was motivated by the need to soundly resolve that drawback, thus introducing the first nice DGF (in the sense of Section 3.1) with guaranteed polynomially small diameter for any decision point.

5. The Dilatable Global Entropy Distance–Generating Function

We now develop our new DGF for sequence-form polytopes. The DGF is based on a scaled variant of the standard entropy DGF (which we will also refer to as the "global entropy" to distinguish it from the dilated entropy), which is strongly convex modulus one on the hypercube $[0,1]^\Sigma$.

Definition 4 (Dilatable Global Entropy). The dilatable global entropy distance–generating function $\tilde{\varphi}$ is the

function $\tilde{\varphi} : Q \to \mathbb{R}_{\geq 0}$ defined as

$$\tilde{\varphi} : Q \ni \boldsymbol{x} \longmapsto w_{\varnothing} x_{\varnothing} \log(x_{\varnothing}) + \sum_{j \in \mathcal{J}} \sum_{a \in A_j} w_{ja} x_{ja} \log x_{ja} + \sum_{j \in \mathcal{J}} \gamma_j x_{p_j} \log |A_j|,$$

where each $\gamma_j \geq 1$ ($j \in \mathcal{J}$) is defined recursively as

$$\gamma_{\varnothing} = 1 + \sum_{j \in \mathcal{C}_{\varnothing}} \gamma_j, \quad \gamma_j := 1 + \max_{a \in A_j} \left\{ \sum_{j' \in \mathcal{C}_{ja}} \gamma_{j'} \right\} \quad \forall j \in \mathcal{J}, \tag{9}$$

and each $w_{\sigma} \geq 1$ ($\sigma \in \Sigma$) is defined recursively as

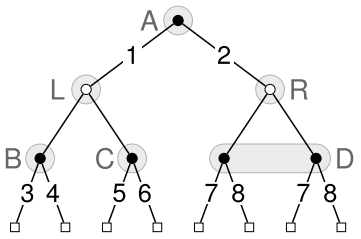
$$w_{\varnothing} := \gamma_{\varnothing} - \sum_{j \in \mathcal{C}_{\varnothing}} \gamma_j,$$

$$w_{ja} := \gamma_j - \sum_{j' \in \mathcal{C}_{ja}} \gamma_{j'} = 1 + \max_{a' \in A_j} \left\{ \sum_{j' \in \mathcal{C}_{ja'}} \gamma_{j'} \right\} - \sum_{j' \in \mathcal{C}_{ja}} \gamma_{j'} \quad \forall ja \in \Sigma.$$

The weights γ_j defined in (9) are very similar to the ones given for the dilated DGFs in the previous section except that the whole expression is smaller by a factor of two. Avoiding this factor of two is crucial because it allows us to avoid the exponential dependence on depth. Here, it is easy to see that γ_j is upper bounded by the number of decision points in the subtree rooted at j, so γ_j is at most polynomial in the size of the sequential decision problem. In fact, it is not hard to show that if j is the sole root decision point, then γ_j is equal to $\max_{x \in Q} \|\boldsymbol{x}\|_1$.

Example 4. Continuing the example started in Figure 1 (reproduced below), in the case of the sequence-form polytope of player 1, the weights γ in the construction of the dilatable global entropy DGF $\tilde{\varphi}$ are computed as follows: $\gamma_{\mathsf{D}} = \gamma_{\mathsf{C}} = \gamma_{\mathsf{B}} = 1$; $\gamma_{\mathsf{A}} = 1 + \max\{\gamma_{\mathsf{B}} + \gamma_{\mathsf{C}}, \gamma_{\mathsf{D}}\} = 3$; $\gamma_{\varnothing} = 1 + \gamma_{\mathsf{A}} = 4$. Correspondingly, all weights w_{ja} are equal to one except for $w_{\mathsf{A2}} = \gamma_{\mathsf{A}} - \gamma_{\mathsf{D}} = 2$. Correspondingly, the DGF ψ is the function

$$\begin{aligned}\tilde{\varphi}(\boldsymbol{x}) = {} & x_{\varnothing} \log(x_{\varnothing}) + x_{\mathsf{A1}} \log(x_{\mathsf{A1}}) + 2 x_{\mathsf{A2}} \log(x_{\mathsf{A2}}) \\ & + x_{\mathsf{B3}} \log(x_{\mathsf{B3}}) + x_{\mathsf{B4}} \log(x_{\mathsf{B4}}) + x_{\mathsf{C5}} \log(x_{\mathsf{C5}}) \\ & + x_{\mathsf{C6}} \log(x_{\mathsf{C6}}) + x_{\mathsf{D7}} \log(x_{\mathsf{D7}}) + x_{\mathsf{D8}} \log(x_{\mathsf{D8}}) \\ & + 3 x_{\varnothing} \log(2) + x_{\mathsf{A1}} \log(2) + x_{\mathsf{A1}} \log(2) + x_{\mathsf{A2}} \log(2).\end{aligned}$$



5.1. Dilatability

The adjective *dilatable* comes from the key property that the dilatable global entropy is equal to a specific dilated entropy regularizer φ *on the sequence-form strategy space* Q. More precisely, consider the dilated entropy DGF defined as

$$\varphi : Q \ni \boldsymbol{x} \longmapsto \gamma_{\varnothing} x_{\varnothing} \log x_{\varnothing} + \sum_{j \in \mathcal{J}} \gamma_j \left(x_{p_j} \log |A_j| + \sum_{a \in A_j} x_{ja} \log \left(\frac{x_{ja}}{x_{p_j}} \right) \right).$$

Then, we have the following.

Theorem 3. *The dilatable global entropy DGF and the dilated entropy DGF coincide on the polytope of sequence-form strategies Q; that is, $\tilde{\varphi}(\boldsymbol{x}) = \varphi(\boldsymbol{x})$ for all $\boldsymbol{x} \in Q$.*

Proof. We start by expanding the definition of $\varphi(\boldsymbol{x})$ for $\boldsymbol{x} \in Q$:

$$\begin{aligned}\varphi(\boldsymbol{x}) &:= \gamma_{\varnothing} x_{\varnothing} \log x_{\varnothing} + \sum_{j \in \mathcal{J}} \sum_{a \in A_j} \gamma_j x_{ja} \log \left(\frac{x_{ja}}{x_{p_j}} \right) + \sum_{j \in \mathcal{J}} \gamma_j x_{p_j} \log |A_j| \\ &= \sum_{j \in \mathcal{J}} \sum_{a \in A_j} \gamma_j x_{ja} \log x_{ja} - \sum_{j \in \mathcal{J}} \sum_{a \in A_j} \gamma_j x_{ja} \log x_{p_j} \\ &\quad + \sum_{j \in \mathcal{J}} \gamma_j x_{p_j} \log |A_j|, \end{aligned} \tag{10}$$

where, in the second equality, we have used the fact that $x_{\varnothing} = 1$, as well as the properties of logarithms. Given the assumption $\boldsymbol{x} \in Q$, it holds that $\sum_{a \in A_j} x_{ja} = x_{p_j}$ for all $j \in \mathcal{J}$, and so, we can simplify the middle summation in (10) and obtain

$$\varphi(\boldsymbol{x}) = \sum_{j \in \mathcal{J}} \sum_{a \in A_j} \gamma_j x_{ja} \log x_{ja} - \sum_{j \in \mathcal{J}} \gamma_j x_{p_j} \log x_{p_j} + \sum_{j \in \mathcal{J}} \gamma_j x_{p_j} \log |A_j|.$$

The middle summation multiplies the weights γ_j by a quantity that depends on the parent sequence p_j of j. It can, therefore, be rewritten equivalently by summing over nonterminal sequences and multiplying by the weight of the children decision points, as follows.

$$\begin{aligned}\varphi(\boldsymbol{x}) &= \sum_{j \in \mathcal{J}} \sum_{a \in A_j} \gamma_j x_{ja} \log x_{ja} - \sum_{j \in \mathcal{J}} \sum_{a \in A_j} \sum_{j' \in \mathcal{C}_{ja}} \gamma_{j'} x_{ja} \log x_{ja} \\ &\quad + \sum_{j \in \mathcal{J}} \gamma_j x_{p_j} \log |A_j| \\ &= \sum_{j \in \mathcal{J}} \sum_{a \in A_j} \gamma_j x_{ja} \log x_{ja} - \sum_{j \in \mathcal{J}} \sum_{a \in A_j} \left(\sum_{j' \in \mathcal{C}_{ja}} \gamma_{j'} \right) x_{ja} \log x_{ja} \\ &\quad + \sum_{j \in \mathcal{J}} \gamma_j x_{p_j} \log |A_j| \\ &= \sum_{j \in \mathcal{J}} \sum_{a \in A_j} w_{ja} x_{ja} \log x_{ja} + \sum_{j \in \mathcal{J}} \gamma_j x_{p_j} \log |A_j| = \tilde{\varphi}(\boldsymbol{x}),\end{aligned}$$

as we wanted to show. □

The equality established in Theorem 3 does not hold outside the sequence-form polytope—a set with no interior. Because of that, it is not surprising that, in general, $\nabla\tilde{\varphi}(\boldsymbol{x}) \neq \nabla\varphi(\boldsymbol{x})$, even at points $\boldsymbol{x}$ in the relative interior of the sequence-form polytope. We will return to the difference between the gradients of $\tilde{\varphi}$ and φ in Section 5.5, where we show that the mismatch between $\nabla\tilde{\varphi}(\boldsymbol{x})$ and $\nabla\varphi(\boldsymbol{x})$ is orthogonal to the sequence-form polytope when $\boldsymbol{x} \in Q$.

5.2. Niceness

We now show that our dilatable global entropy regularizer is nice in the sense of Section 3.1; that is, its gradient and the gradient of its convex conjugate can be computed exactly in linear time in $|\Sigma|$.

- The gradient of $\tilde{\varphi}$ can be trivially computed in closed form and linear time in $|\Sigma|$ starting from Definition 4 as

$$(\nabla\tilde{\varphi}(\boldsymbol{x}))_\sigma = (1 + \log x_\sigma) w_\sigma + \sum_{j \in \mathcal{C}_\sigma} \gamma_j \log|A_j| \qquad \forall \sigma \in \Sigma, \boldsymbol{x} \in \text{relint } Q.$$

- Using the dilatability property, we have that the gradient of the convex conjugate satisfies

$$\nabla\tilde{\varphi}^*(\boldsymbol{g}) = \arg\max_{\boldsymbol{x} \in Q}\{\boldsymbol{g}^\top \boldsymbol{x} - \tilde{\varphi}(\boldsymbol{x})\} = \arg\max_{\boldsymbol{x} \in Q}\{\boldsymbol{g}^\top \boldsymbol{x} - \varphi(\boldsymbol{x})\} = \nabla\varphi^*(\boldsymbol{g}), \tag{11}$$

where we used the dilatability property (Theorem 3) in the second equality. Therefore, because φ is a dilated DGF and its smoothed support function can be computed in linear time, the smoothed support function of $\tilde{\varphi}$ can be computed in linear time in $|\Sigma|$.

An immediate consequence of the niceness established in the previous bullet point is that proximal operators of $\tilde{\varphi}$ can be computed efficiently, as

$$\text{prox}_{\tilde{\varphi}}(\boldsymbol{g} \| \boldsymbol{c}) = \nabla\tilde{\varphi}^*(-\boldsymbol{g} + \nabla\tilde{\varphi}(\boldsymbol{c})) = \nabla\varphi^*(-\boldsymbol{g} + \nabla\tilde{\varphi}(\boldsymbol{c})),$$

where the last equality is a special case of (11).

5.3. Strong Convexity

On the other hand, we now show that $\tilde{\varphi}$ has the advantage of a better strong convex modulus compared with the existing dilated entropy DGFs.

Theorem 4. *The dilatable global entropy function $\tilde{\varphi} : Q \to \mathbb{R}_{\geq 0}$ is a DGF for the sequence-form polytope Q, 1-strongly convex on* relint Q *with respect to the ℓ_2 norm.*

Proof. The function $\tilde{\varphi}$ is twice differentiable on $(0,1)^\Sigma \supseteq$ relint Q. Using (1), we conclude that $\tilde{\varphi}$ is 1-strongly convex because the Hessian is

$$\nabla^2\tilde{\varphi}(\boldsymbol{x}) = \text{diag}\left(\left\{\frac{w_{ja}}{x_{ja}}\right\}_{ja \in \Sigma}\right) \geq I,$$

where we used the inequalities $0 \leq x_{ja} \leq 1$ and $w_{ja} \geq 1$. Next, we verify that the minimum of $\tilde{\varphi}$ is zero. Because of dilatability, $\arg\min_{\boldsymbol{x} \in Q} \tilde{\varphi}(\boldsymbol{x}) = \arg\min_{\boldsymbol{x} \in Q} \varphi(\boldsymbol{x}) = \nabla\varphi^*(\boldsymbol{0}) = \boldsymbol{m}$ is the uniform strategy, that is, the strategy that assigns probability $1/|A_j|$ to each action at each decision point $j \in \mathcal{J}$ (this corresponds, in the sequence-form representation, to $x_{ja} = x_{p_j}/|A_j|$ for all $ja \in \Sigma$ and $x_\varnothing = 1$). Substituting into the definition of $\tilde{\varphi}$, we obtain

$$\tilde{\varphi}(\boldsymbol{m}) = \sum_{j \in \mathcal{J}} \gamma_j \left(x_{p_j} \log|A_j| + \sum_{a \in A_j} \frac{x_{p_j}}{|A_j|} \log\frac{1}{|A_j|} \right) = \sum_{j \in \mathcal{J}} \gamma_j (x_{p_j} \log|A_j| - x_{p_j} \log|A_j|) = 0. \quad \square$$

Theorem 5. *The dilatable global entropy function $\tilde{\varphi}$ is strongly convex modulus $1/M_Q$ with respect to the ℓ_1 norm on* relint Q.

Proof. Using the second-order definition of strong convexity, we wish to show that the inequality $\boldsymbol{m}^\top \nabla^2\tilde{\varphi}(\boldsymbol{x})\boldsymbol{m} \geq \frac{1}{M_Q}\|\boldsymbol{m}\|_1^2$ holds for any $\boldsymbol{m} \in \mathbb{R}^\Sigma$. Expanding the Hessian matrix and using the fact that $w_{ja} \geq 1$ for all $ja \in \Sigma$ gives

$$\boldsymbol{m}^\top \nabla^2\tilde{\varphi}(\boldsymbol{x})\boldsymbol{m} = \boldsymbol{m}^\top \text{diag}\left(\left\{\frac{w_{ja}}{x_{ja}}\right\}_{ja \in \Sigma}\right)\boldsymbol{m} \geq \sum_{ja \in \Sigma} \frac{m_{ja}^2}{x_{ja}}. \tag{12}$$

On the other hand, by expanding the definition of $\|\boldsymbol{m}\|_1^2$ and applying the Cauchy-Schwarz inequality, we have

$$\|\boldsymbol{m}\|_1^2 = \left(\sum_{ja \in \Sigma} |m_{ja}|\right)^2 = \left(\sum_{ja \in \Sigma} \frac{|m_{ja}|}{\sqrt{x_{ja}}}\sqrt{x_{ja}}\right)^2 \leq \left(\sum_{ja \in \Sigma} \frac{m_{ja}^2}{x_{ja}}\right)\left(\sum_{ja \in \Sigma} x_{ja}\right) \leq \left(\sum_{ja \in \Sigma} \frac{m_{ja}^2}{x_{ja}}\right) M_Q.$$

Substituting (12) into the last inequality yields a proof of the desired strong-convexity modulus $1/M_Q$. $\square$

5.4. Diameter

The properties above immediately imply that our dilatable global entropy DGF satisfies all the requirements for a prox setup on the polytope of sequence-form strategies Q. Here, we complete the analysis by giving bounds on the diameter induced by $\tilde{\varphi}$.

Theorem 6. *The $\tilde{\varphi}$-diameter $\Omega_{\tilde{\varphi},Q}$ of Q is at most $M_Q^2 \max_{j' \in \mathcal{J}} \log|A_{j'}|$.*

Proof. By the definition of the polytope diameter and the fact that we chose our DGFs such that

$\min_{x\in\mathcal{A}} \tilde{\varphi}(x) = 0$, we have

$$\begin{aligned}\Omega_{\tilde{\varphi},Q} &= \max_{x\in Q} \tilde{\varphi}(x) \le \max_{x\in Q} \sum_{j\in\mathcal{J}} \gamma_j x_{p_j} \log|A_j| \\ &\le \max_{x\in Q} \max_{j'\in\mathcal{J}} \log|A_{j'}| \sum_{j\in\mathcal{J}} \gamma_j x_{p_j} \\ &\le M_Q \max_{x\in Q} \max_{j'\in\mathcal{J}} \log|A_{j'}| \sum_{j\in\mathcal{J}} x_{p_j} \\ &\le M_Q^2 \max_{j'\in\mathcal{J}} \log|A_{j'}|,\end{aligned}$$

where the first inequality is by noting that $\log x_{ja} \le 0$ because $x_{ja} \le 1$ for all $ja \in \Sigma$; the third inequality is by noting that γ_j is largest at root decision points, where it is at most M_Q; and the fourth inequality upper bounds $\sum_{j\in\mathcal{J}} x_{p_j}$ by M_Q. □

Kroer et al. (2020) show that the dilated entropy DGF with weights β leads to a polytope diameter

$$2^{2\mathfrak{D}_Q+2} M_Q^2 \max_{j'\in\mathcal{J}} \log|A_{j'}|.$$

Our DGF improves that polytope diameter by a factor of $2^{2\mathfrak{D}_Q+2}$. Thus, we are the first to achieve a polytope diameter with no exponential dependence on the depth $\mathfrak{D}_Q$ of the sequence-form polytope.

Example 5. Let Q be the sequence-form polytope corresponding to an optimal stopping problem: an agent has a sequence of k decision points, and at each decision point, the agent can either choose to stop the game, in which case a payoff is reached, or choose to continue the game, in which case they continue to the next decision point. For such a problem, Q has depth k, which is also the size of Q as measured by the ℓ_1 norm; that is, $M_Q = k$. In this case, the diameter associated to $\tilde{\varphi}$ is $k^2 \log 2$, whereas the dilated entropy DGF of Kroer et al. (2020) gives diameter $2^{k+2}k^2 \log 2$.

Example 5 shows the most extreme type of decision problem in terms of our improvement over existing results. Another example of a polytope that partially embeds this structure is the decision spaces of a no-limit poker game. In those games, a player may raise by any amount between the minimum raise size and the size of that player's stack. Because of this, there exists a very deep decision path where players take turns raising by the minimum amount. However, the average depth is much smaller than this.

Summing up our results on the dilatable global entropy, we have shown that it enjoys the same fast smoothed-support-function computation as the dilated entropy DGF while having a better way to achieve strong-convexity modulus $1/M_Q$. In particular, the existing dilated entropy setup requires the weight parameters β to grow exponentially in the depth of the sequence-form polytope, whereas we have only a linear growth in those weights. More concretely, this means that the largest weights $\max_{j\in\mathcal{J}} \beta_j$ in the dilated entropy DGF are larger than the largest weights $\max_{j\in\mathcal{J}} \gamma_j$ in the dilatable global entropy DGF by a factor of more than $2^{\mathfrak{D}_Q}$. This, in turn, allowed us to achieve a better polytope diameter by a factor of $2^{\mathfrak{D}_Q+2}$ while retaining the same strong-convexity modulus.

5.5. Gradient Properties and Further Relationships Between $\tilde{\varphi}$ and φ

The previous subsections establish that the dilated global entropy DGF φ is dilatable (i.e., it coincides with a dilated entropy DGF, φ, on all of their domain Q) and nice (i.e., gradients, smoothed support functions, and proximal operators can be computed efficiently). Furthermore, we have seen that φ has very appealing metric properties, including a *diagonal* Hessian matrix that renders establishing strong convexity rather trivial, especially as compared with the analyses of dilated entropy DGFs in the prior literature (Kroer et al. 2020).

In this subsection, we go one step further: by studying the relationship between the gradients of $\tilde{\varphi}$ and φ, we will establish that the two enjoy the same strong-convexity parameter. In particular, we will use the following property, which we prove in Online Appendix A.

Lemma 1. *At any point $x \in$ relint Q, the gradients $\nabla\tilde{\varphi}(x)$ and $\nabla\varphi(x)$ differ only along orthogonal directions to the affine hull* aff Q *of Q; that is,*

$$(\nabla\tilde{\varphi}(x) - \nabla\varphi(x))^\top (y - z) = 0 \qquad \forall y, z \in Q.$$

Lemma 1 immediately implies that for any two points $x, x' \in$ relint Q:

$$(\nabla\varphi(x) - \nabla\varphi(x'))^\top (x - x') = (\nabla\tilde{\varphi}(x) - \nabla\tilde{\varphi}(x'))^\top (x - x');$$

that is, the strong-convexity properties of $\tilde{\varphi}$ established in Section 5.3 apply to φ as well. In particular, this shows that the specific dilated entropy DGE φ is 1-strongly convex on relint Q with respect to the ℓ_2 norm and $1/M_Q$-strongly convex with respect to the ℓ_1 norm. Additionally, the analysis of the diameter of $\tilde{\varphi}$ transfers directly to φ by dilatability (Section 5.1). Combined, these results enable us to introduce a 1-strongly convex *dilated entropy* DGF (and not just any nice DGF) φ that, for the first time, has worst-case polynomially bounded diameter.

We remark that it would be hard to establish such a result without first introducing our dilatable global regularizer, as analyzing the product $(\nabla\varphi(x) - \nabla\varphi(x'))^\top (x - x')$ for a dilated entropy DGF φ is notoriously involved. Instead, the good metric properties of $\tilde{\varphi}$—especially the diagonal Hessian matrix—make the analysis of the strong-convexity and diameter properties trivial.

In fact, every strong-convexity proof for an entropy DGF that we are aware of uses the Hessian-based condition (Ben-Tal and Nemirovski 2001, Kroer et al.

2020). If we had used this definition directly on the dilated DGF φ, then the lower bound from Kroer et al. (2020) implies that we would get an exponential dependence on the depth. This is because the Hessian condition is a sufficient but not necessary condition for strong convexity: it asks for the condition to hold along directions orthogonal to the affine hull of the sequence-form strategy space, which is not necessary. Arguably, the key insight for the dilatable global entropy is that it allows us to use the Hessian-based sufficient condition for proving strong convexity in a way that avoids paying the cost of maintaining strong convexity along irrelevant orthogonal directions. The fact that it also greatly simplifies the Hessian-based analysis is an added bonus.

Another consequence of Lemma 1 is that the proximal steps induced by $\tilde{\varphi}$ and φ coincide. Indeed, for all centers $\boldsymbol{y} \in \operatorname{relint} Q$ and gradients $\boldsymbol{g}$, one has

$$\begin{aligned}
&\arg\min_{\boldsymbol{y}\in Q}\{\boldsymbol{g}^\top \boldsymbol{y} + D_{\tilde{\varphi}}(\boldsymbol{y}\,\|\,\boldsymbol{x})\}\\
&\quad= \arg\min_{\boldsymbol{y}\in Q}\{\boldsymbol{g}^\top \boldsymbol{y} + \tilde{\varphi}(\boldsymbol{y}) - \nabla\tilde{\varphi}(\boldsymbol{x})^\top(\boldsymbol{y}-\boldsymbol{x})\}\\
&\quad= \arg\min_{\boldsymbol{y}\in Q}\{\boldsymbol{g}^\top \boldsymbol{y} + \varphi(\boldsymbol{y}) - \nabla\tilde{\varphi}(\boldsymbol{x})^\top(\boldsymbol{y}-\boldsymbol{x})\}\\
&\qquad\qquad\qquad\qquad\qquad\qquad\text{(dilatability, Theorem 3)}\\
&\quad= \arg\min_{\boldsymbol{y}\in Q}\{\boldsymbol{g}^\top \boldsymbol{y} + \varphi(\boldsymbol{y}) - \nabla\varphi(\boldsymbol{x})^\top(\boldsymbol{y}-\boldsymbol{x})\}\quad\text{(Lemma 1)}\\
&\quad= \arg\min_{\boldsymbol{y}\in Q}\{\boldsymbol{g}^\top \boldsymbol{y} + D_{\varphi}(\boldsymbol{y}\,\|\,\boldsymbol{x})\}.
\end{aligned}$$

Hence, mirror descent, online mirror descent, follow-the-regularizer-leader, mirror prox, and EGT (among others) all produce the same iterates when instantiated with φ or $\tilde{\varphi}$. It is conceivable that some methods—for example, methods that require taking norms of gradients or methods that rely on norms induced by the Hessian matrix of the regularizer (e.g., Abernethy and Rakhlin 2009)—would, in general, not be equivalent when set up using φ or $\tilde{\varphi}$.

6. Scaled Extension and Correlated Decision Spaces

In this section, we extend and generalize both the framework of dilated DGFs and the dilatable global entropy DGF to more complex combinatorial domains than sequence-form polytopes. In particular, we show that dilated DGFs and the dilatable global entropy apply to sets that can be constructed through composition of *scaled extension*, a convexity-preserving operation that was recently proposed as a general way of constructing sequential decision spaces in the presence of correlation between the strategies of two or more players (Farina et al. 2019d).

Our generalization begets the first nice regularizers (in a sense similar to Section 3.1; see Definition 6) for correlated strategy spaces, which, in turn, enables us to construct the first FOMs that guarantee convergence to optimal correlated equilibria and optimal ex ante team-coordinated equilibria at a rate of $1/T$ in certain classes of games where these equilibria can be found in polynomial time.

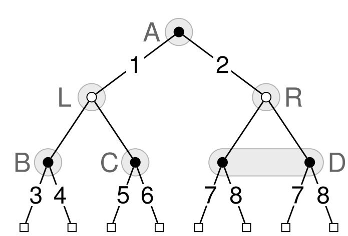
We start by recalling the definition of scaled extension and give an example showing how sequence-form polytopes arise from composition of such an operation.

Definition 5 (Scaled Extension (Farina et al. 2019d)). Let $\mathcal{U}$ and $\mathcal{V}$ be nonempty convex sets, and let $h : \mathcal{U} \to \mathbb{R}_{\geq 0}$ be a nonnegative affine real function. The *scaled extension* of $\mathcal{U}$ with $\mathcal{V}$ via h is defined as the convex set

$$\mathcal{U} \overset{h}{\triangleleft} \mathcal{V} := \{(\boldsymbol{u}, h(\boldsymbol{u})\,\boldsymbol{v}) : \boldsymbol{u} \in \mathcal{U}, \boldsymbol{v} \in \mathcal{V}\}.$$

Example 6. This example illustrates how scaled extension operations can be chained to capture the sequence-form strategy space

$$\begin{cases}
x_\varnothing = 1,\\
x_{\mathsf{A1}} + x_{\mathsf{A2}} = x_\varnothing,\\
x_{\mathsf{B3}} + x_{\mathsf{B4}} = x_{\mathsf{A1}},\\
x_{\mathsf{C5}} + x_{\mathsf{C6}} = x_{\mathsf{A1}},\\
x_{\mathsf{D7}} + x_{\mathsf{D8}} = x_{\mathsf{A2}}.
\end{cases}$$



of the small game in Figure 1(a), reproduced above.

(i) First, the empty sequence is set to value $x_\varnothing = 1$.

(ii) (Decision point A) Next, the value $x_\varnothing$ is partitioned into the two nonnegative values

$$x_{\mathsf{A1}} + x_{\mathsf{A2}} = x_\varnothing.$$

(iii) (Decision point B) Next, the value x_{A1} is partitioned into two nonnegative values

$$x_{\mathsf{B3}} + x_{\mathsf{B4}} = x_{\mathsf{A1}}.$$

(iv) (Decision point C) Next, the value x_{A1} is partitioned into two nonnegative values

$$x_{\mathsf{C5}} + x_{\mathsf{C6}} = x_{\mathsf{A1}}.$$

(v) (Decision point D) Next, the value x_{A2} is partitioned into two nonnegative values

$$x_{\mathsf{D7}} + x_{\mathsf{D8}} = x_{\mathsf{A2}}.$$

The incremental choices in the above recipe can be directly translated—in the same order—into set operations by using scaled extensions, as follows:

(i) First, the set of all feasible values of sequence $x_\varnothing$ is the singleton $\mathcal{X}_0 := \{1\}$.

(ii) Then, the set of all feasible values of $(x_\varnothing, x_{\mathsf{A1}}, x_{\mathsf{A2}})$ is the set

$$\mathcal{X}_1 := \mathcal{X}_0 \times \Delta^2 = \mathcal{X}_0 \overset{h_1}{\triangleleft} \Delta^2,$$

where h_1 is the linear function $h_1 : \mathcal{X}_0 \ni x_\varnothing \longmapsto x_\varnothing$ (the identity function).

(iii) In order to characterize the set of all feasible values of $(x_\varnothing, \dots, x_{\mathsf{B4}})$, we start from $\mathcal{X}_1$ and *extend* any element $(x_\varnothing, x_{\mathsf{A1}}, x_{\mathsf{A2}}) \in \mathcal{X}_1$ with the two sequences x_{B3} and x_{B4}, drawn from the set

$$\{(x_{\mathsf{B3}}, x_{\mathsf{B4}}) \in \mathbb{R}^2_{\geq 0} : x_{\mathsf{B3}} + x_{\mathsf{B4}} = x_{\mathsf{A1}}\} = x_{\mathsf{A1}}\Delta^2.$$

We can express this extension using scaled extension $\mathcal{X}_2 := \mathcal{X}_1 \overset{h_2}{\triangleleft} \Delta^2$, where $h_2 : \mathcal{X}_1 \ni (x_\varnothing, x_{\mathsf{A1}}, x_{\mathsf{A2}}) \longmapsto x_{\mathsf{A1}}$.

(iv) Similarly, we can extend every element in $\mathcal{X}_2$ to include $(x_{\mathsf{C5}}, x_{\mathsf{C6}}) \in x_{\mathsf{A1}}\Delta^2$: in this case, $\mathcal{X}_3 := \mathcal{X}_2 \overset{h_3}{\triangleleft} \Delta^2$, where $h_3 : \mathcal{X}_2 \ni (x_\varnothing, x_{\mathsf{A1}}, x_{\mathsf{A2}}, x_{\mathsf{B3}}, x_{\mathsf{B4}}) \longmapsto x_{\mathsf{A1}}$.

(v) The set of all feasible $(x_\varnothing, \dots, x_{\mathsf{C8}})$ is $\mathcal{X}_4 := \mathcal{X}_3 \overset{h_4}{\triangleleft} \Delta^2$, where $h_4 : \mathcal{X}_3 \ni (x_\varnothing, \dots, x_{\mathsf{C6}}) \longmapsto x[2]$.

In the Online Appendix, we prove the following property, which will be useful for the construction of our DGF for sets obtained through scaled extension.

Lemma 2. *Let $\mathcal{U} \subset \mathbb{R}^m, \mathcal{V} \subset \mathbb{R}^n$ be bounded sets and $h : \mathcal{U} \longrightarrow \mathbb{R}$ be an affine function $h : \boldsymbol{u} \longmapsto \boldsymbol{a}^\top \boldsymbol{u} + b$, nonnegative on $\mathcal{U}$ and strictly positive on* relint $\mathcal{U}$*. Then,*

$$\operatorname{relint}(\mathcal{U} \overset{h}{\triangleleft} \mathcal{V}) = (\operatorname{relint} \mathcal{U}) \overset{h}{\triangleleft} (\operatorname{relint} \mathcal{V}).$$

To keep the treatment concrete, we are especially interested in the following setup, which captures both the set of sequence-form strategies Q of a generic player as well as the set of correlation plans Ξ used in correlated solution concepts, recalled in Section 6.2, which will be of interest in the remainder of the paper.

Setup 1. Let $\mathcal{X}$ be a set expressible in the form

$$\mathcal{X} = \mathcal{X}_1 \overset{h_1}{\triangleleft} \mathcal{X}_2 \overset{h_2}{\triangleleft} \cdots \overset{h_{n-1}}{\triangleleft} \mathcal{X}_n, \tag{13}$$

where

- Each $\mathcal{X}_k \subseteq \mathbb{R}^{s_k}$ is a compact and convex set such that $\max_{\boldsymbol{v} \in \mathcal{X}_k} \|\boldsymbol{v}\|_2 \leq 1$. (In the case of $\mathcal{X} = Q$ and $\mathcal{X} = \Xi$, each $\mathcal{X}_k$ is a probability simplex $\mathcal{X}_k = \Delta^{s_k}$.)
- Each h_k is a linear function nonnegative on $\mathcal{U}_k := \mathcal{X}_1 \overset{h_1}{\triangleleft} \cdots \overset{h_{k-1}}{\triangleleft} \mathcal{X}_k$, strictly positive in the relative interior of $\mathcal{U}_k$. Furthermore, we assume that each h_k can be written in the form

$$h_k : \mathcal{U}_k \ni \boldsymbol{u} := (\boldsymbol{u}_1, \dots, \boldsymbol{u}_k) \longmapsto \boldsymbol{a}_k^\top \boldsymbol{u} = \boldsymbol{a}_{k,1}^\top \boldsymbol{u}_1 + \cdots + \boldsymbol{a}_{k,k}^\top \boldsymbol{u}_k, \tag{14}$$

for an appropriate vector $\boldsymbol{a}_k := (\boldsymbol{a}_{k,1}, \dots, \boldsymbol{a}_{k,k}) \in [0,1]^{s_1} \times \cdots \times [0,1]^{s_k}$ and that $h_k(\boldsymbol{u}) \leq 1$ for all $\boldsymbol{u} \in \mathcal{U}_k$;
- For each $k = 1, \dots, n$, a 1-strongly convex DGF d_k for $\mathcal{X}_k$ has been chosen, where strong convexity is measured with respect to the Euclidean norm.

6.1. Niceness in the Context of Scaled Extensions

The structure imposed by scaled extension often allows answering questions on a generic set of the form $\mathcal{Z} := \mathcal{U} \overset{h}{\triangleleft} \mathcal{V} \subseteq \mathbb{R}^{m+n}$, where $h : \boldsymbol{u} \longmapsto \boldsymbol{a}^\top \boldsymbol{u} + b$, by looking at a similar question on $\mathcal{U}$ and $\mathcal{V}$. For example, checking whether a point $(\boldsymbol{u}, \boldsymbol{w}) \in \mathbb{R}^{m+n}$ belongs to $\mathcal{Z}$ can be decomposed into checking whether $\boldsymbol{u} \in \mathcal{U}$ and $\boldsymbol{w}/(\boldsymbol{a}^\top \boldsymbol{u} + b) \in \mathcal{V}$. Similarly, it is easy to see that a generic linear optimization problem on $\mathcal{Z}$,

$$\boldsymbol{z}^* \in \arg\max_{\boldsymbol{z} \in \mathcal{Z}} \boldsymbol{p}^\top \boldsymbol{u} + \boldsymbol{q}^\top \boldsymbol{w},$$

can be decomposed into the two linear optimization problems

1. $\boldsymbol{v}^* \in \arg\max_{\boldsymbol{v} \in \mathcal{V}} \boldsymbol{q}^\top \boldsymbol{v}$,
2. $\boldsymbol{u}^* \in \arg\max_{\boldsymbol{u} \in \mathcal{U}} (\boldsymbol{p} + (\boldsymbol{q}^\top \boldsymbol{v}^*)\boldsymbol{a})^\top \boldsymbol{u}$

so that an optimal solution to the original problem is then $\boldsymbol{z}^* = (\boldsymbol{u}^*, h(\boldsymbol{u}^*)\boldsymbol{v}^*) \in \mathcal{Z}$. In both of these cases, we are able to decompose an operation on $\mathcal{Z}$ into two smaller, similar operations on $\mathcal{U}$ and $\mathcal{V}$, *plus an additional overhead that is linear in the number of nonzeros in* $\boldsymbol{a}$. Indeed, to check whether $(\boldsymbol{u}, \boldsymbol{w}) \in \mathcal{Z}$, we needed to divide $\boldsymbol{w}$ by $h(\boldsymbol{u})$, and computing the latter quantity can be carried out in $O(\|\boldsymbol{a}\|_0)$ time. Similarly, to perform linear optimization on $\mathcal{Z}$, we had to multiply $\boldsymbol{v}^*$ by $h(\boldsymbol{u}^*)$ and to adjust the objective function for the optimization problem on $\mathcal{U}$ from $\boldsymbol{p}$ to $\boldsymbol{p} + \boldsymbol{a}$; again, both operations require $O(\|\boldsymbol{a}\|_0)$ time.

In light of the preceding discussion, we generalize the definition of niceness given in Section 3.1 to sets complying with setup 1, accounting for this overhead.

Definition 6. Let $\mathcal{X} = \mathcal{X}_1 \overset{h_1}{\triangleleft} \mathcal{X}_2 \overset{h_2}{\triangleleft} \cdots \overset{h_{n-1}}{\triangleleft} \mathcal{X}_n$ be a set obtained through repeated scaled extension, where the sets $\mathcal{X}_k \subseteq \mathbb{R}^{s_k}$ and nonnegative linear functions $h_k(\boldsymbol{u}_k) = \boldsymbol{a}_k^\top \boldsymbol{u}_k$ satisfy the conditions of setup 1. A distance-generating function d for $\mathcal{X}$ is said to be nice if $d(\boldsymbol{z})$, $\nabla d(\boldsymbol{z})$, and $\nabla d^*(\boldsymbol{g})$ can be computed exactly in time $O(\sum_{k=1}^n s_k + \sum_{k=1}^{n-1} \|\boldsymbol{a}_k\|_0)$.

We remark that in a sequence-form strategy space Q, the scaling functions involved in the representation of Q via scaled extension are such that each scaling function $h : \boldsymbol{u} \longmapsto \boldsymbol{a}^\top \boldsymbol{u} + b$ has exactly one nonzero in $\boldsymbol{a}$ (corresponding to the parent sequence of the decision point; see Example 6). So, in the particular case of sequence-form strategy spaces, the more general definition of niceness given in Definition 6 reduces to the notion of Section 3.1.

6.2. Preliminaries on Correlation and Triangle Freeness

As reviewed earlier in the paper, Nash equilibria in two-player zero-sum EFGs can be expressed as BSPPs. It turns out that several other solution concepts can also be formulated as BSPPs via more intricate convex-polytope constructions. In this section, we briefly describe two

important solution concepts that can be expressed as BSPPs: several variants of optimal correlated equilibria and ex ante team-coordinated equilibria.

In correlated equilibria, the rationality assumption of Nash equilibrium is relaxed in order to allow for coordination between the players. It is assumed that a *mediator* will recommend actions to be taken. In *correlated equilibria*, each player sees the recommended action before deciding whether to take it. In *coarse correlated equilibria*, the players must commit to acting according to the recommended strategy before the recommendation is revealed. In all these solution concepts, the recommended strategy is sampled by a mediator from some correlated distribution that is known to the players.

We mention three types of correlated equilibria in EFGs, which differ in how the mediator's recommendations are given and when players have to choose whether to abide by them:

- Extensive-form correlated equilibrium (EFCE): The mediator incrementally recommends individual moves to the players. Every time a player faces a decision point, the mediator privately reveals a recommended move for that decision point to that player. If a player chooses to disregard a recommendation, then the mediator immediately stops issuing recommendations to that player forever (von Stengel and Forges 2008).
- Extensive-form coarse correlated equilibrium (EFCCE): The mediator incrementally recommends individual moves to the players, but at each decision point, the player must decide whether to follow the recommendation *before* seeing the recommendation (Farina and Sandholm 2020).
- Normal-form coarse correlated equilibrium (NFCCE): each player will be recommended a strategy from the normal-form representation of the EFG, but they must decide whether to commit to playing the recommended strategy before seeing the recommendation (Moulin and Vial 1978).

Farina and Sandholm (2020) show that EFCE is a subset of EFCCE and that EFCCE is a subset of NFCCE. They also show that for triangle-free decision problems and triangle-free EFGs,[5] the set Ξ of all possible correlated plans between two players can be obtained via composition of a polynomial number (in the size of the original game tree) of scaled extension operators. In a recent paper, Zhang et al. (2022) extend the result to general multiplayer games and give parameterized complexity guarantees on the number of scaled extension operations needed. A correlated solution concept with social welfare at least some arbitrary value τ (if it exists) can be expressed as a BSPP

$$\arg\min_{\boldsymbol{x}\in\Xi} \max_{\boldsymbol{y}\in\mathcal{Y}} \boldsymbol{x}^\top \mathcal{A}\boldsymbol{y}, \tag{15}$$

where $\mathcal{A}$ is a matrix that depends on τ, the minimization over $\boldsymbol{x}$ represents the choice of a correlated plan for the players, and the maximization over $\mathcal{Y}$, intuitively, represents different ways of rejecting the mediator's recommendation. By carefully choosing $\mathcal{Y}$, we can enforce different types of correlated equilibrium behavior. Now, as long as we have a nice DGF for scaled extensions and a nice DGF for the polytope $\mathcal{Y}$, we can apply fast FOMs to the computation of the corresponding type of correlated equilibrium. Appropriate characterizations for $\mathcal{Y}$ exist in the case of EFCE (Farina et al. 2019c), EFCCE (Farina et al. 2020a), and NFCCE (Farina et al. 2020a). In each case, $\mathcal{Y}$ is itself a polytope that can be constructed via scaled extension.

Reviewing the details of the scaled extension–based construction of the polytope of correlation plans Ξ is beyond the scope of this paper. Here, we take the decomposition as a given, and we are concerned with the task of constructing a suitable proximal setup for sets that, like $\mathcal{Y}$ and Ξ, can be expressed through composition of scaled extension operations.

In addition to correlated equilibrium problems discussed above, one can also capture *adversarial team games* with the scaled extension DGFs that we will construct. In the adversarial team game that we consider, players on a team (meaning that they share the same payoffs) are trying to correlate their strategies so as to maximize utility against an opposing team whose utility is exactly the opposite of theirs (i.e., it is a zero-sum game between the two teams). This solution concept is called *team-maxmin equilibrium with coordination (TMECor)* (Celli and Gatti 2018). The set of correlated plans for the players on the team can again be expressed with Ξ (Farina et al. 2021a). An alternative and more concise representation—again based on scaled extension—was recently proposed by Zhang et al. (2023). It follows that DGFs for both polytopes can be chosen as the DGE from our paper.

By applying this DGF to a method such as EGT or mirror prox, we get a $1/T$ iterative method for converging to each of these solution concepts, with only a linear cost per iteration. In contrast, prior iterative approaches at the time of submission of this article converged at a rate of $1/\sqrt{T}$ and required significantly more expensive projections at every iteration (e.g., Farina et al. 2019c, d).

6.3. Dilated Distance-Generating Functions for Scaled Extension

In this section, we show that the construction of dilated DGFs can be generalized to sets obtained through scaled extension. Let $\mathcal{Z} := \mathcal{U} \overset{h}{\triangleleft} \mathcal{V}$ be a set constructed by scaled extension of $\mathcal{U}$ with $\mathcal{V}$ using a linear function h, and assume that nice (in the sense of Section 3.1) DGFs d_u, d_v for $\mathcal{U}$ and $\mathcal{V}$, respectively, have been chosen. In Proposition 1, we show that d_u and d_v can be combined to give a composite DGF for $\mathcal{Z}$ with desirable properties.

Proposition 1. *Let*

- $\mathcal{Z} := \mathcal{U} \overset{h}{\triangleleft} \mathcal{V}$, *where* $\mathcal{U} \subseteq \mathbb{R}^m$, $\mathcal{V} \subseteq \mathbb{R}^n$ *are compact convex sets, and* h *is a linear function* $\boldsymbol{u} \mapsto \boldsymbol{a}^\top \boldsymbol{u}$ *such that* $h(\boldsymbol{u}) \geq 0$ *for all* $\boldsymbol{u} \in \mathcal{U}$, *and* $h(\boldsymbol{u}) > 0$ *for all* $\boldsymbol{u} \in \operatorname{relint} \mathcal{U}$;

- $d_u : \mathcal{U} \to \mathbb{R}$ *and* $d_v : \mathcal{V} \to \mathbb{R}$ *be DGFs for* $\mathcal{U}$ *and* $\mathcal{V}$, *strongly convex with respect to norms* $\|\cdot\|_u$ *and* $\|\cdot\|_v$, *respectively;*
- $\alpha_v > 0$ *be a positive scalar.*

Then, the function

$$d_z : \mathcal{Z} \ni (u, w) \longmapsto d_u(u) + \alpha_v d_v^{\square}(u, w),$$
$$d_v^{\square}(u, w) := \begin{cases} 0 & \text{if } h(u) = 0 \\ h(u) d_v\left(\frac{w}{h(u)}\right) & \text{if } h(u) > 0 \end{cases} \tag{16}$$

is a DGF for $\mathcal{Z}$. *Furthermore, the functions* d_z, ∇d_z, *and* ∇d_z^* *can be evaluated on* $\mathcal{Z}$, *each using a single call to an oracle for* $d_u, \nabla d_u$, *or* ∇d_u^*, *as well as a constant number of calls to an oracle for* $d_v, \nabla d_v$, *and/or* ∇d_v^*, *plus an additional overhead linear in* $\|a\|_0$ *and the dimension* n *of* $\mathcal{V} \subseteq \mathbb{R}^n$.

Proof. We verify that all conditions stated in Definition 1 are satisfied.

- *Continuity*: In order to show that d_z is continuous, it suffices to show that $d_v^{\square}$ is. Let $(u', w') \in \mathcal{Z}$. If $h(u') > 0$, then by continuity of h, there exists a neighborhood of (u', w') where $d_v^{\square}(u, w)$ coincides with the function $h(u) d_v(w/h(u))$, which is clearly continuous because d_v is continuous by hypothesis and h is a linear function. Hence, we are only left with checking continuity in the case where $h(u') = 0$. First, note that $d_v^{\square}$ is a nonnegative function given the hypotheses and that $M := \max_{v \in \mathcal{V}} d_v(v)$ is a finite constant by Weierstrass's theorem because d_v is a continuous function defined over a compact domain. For any $\epsilon > 0$, consider the neighborhood $N_{r(\epsilon)}$ of radius $r(\epsilon) := \epsilon/(2\|a\|_2 M)$ centered in (u', w'). Then, for any point (u, w) in the neighborhood, we have that either $h(u) = 0$, yielding $d_v^{\square}(u, w) = 0$, or $h(u) > 0$, yielding

$$d_v^{\square}(u, w) = h(u) d_v\left(\frac{w}{h(u)}\right) = a^\top (u - u') d_v\left(\frac{w}{h(u)}\right)$$
$$\le \|a\|_2 \|u - u'\|_2 M \le \frac{\epsilon}{2} < \epsilon.$$

So, in either case, we have that $0 \le d^{\square}(u, w) < \epsilon$ for all $(u, w) \in N_{r(\epsilon)}$; that is, $d_v^{\square}$ is continuous at (u', w'), as we wanted to show.

- *Differentiability*: Because $h(u) > 0$ for all $u \in \operatorname{relint} \mathcal{U}$ by hypothesis, Lemma 2 guarantees that $\operatorname{relint} \mathcal{Z} = (\operatorname{relint} \mathcal{U}) \triangleleft^h (\operatorname{relint} \mathcal{V})$. Furthermore, using the definition of $d_v^{\square}$, we have that

$$d_z(u, w) = d_u(u) + \alpha_v h(u) d_v\left(\frac{w}{h(u)}\right) \qquad \forall (u, w) \in \operatorname{relint} \mathcal{Z}.$$

Hence, direct computation shows that at any $(u, w) \in \operatorname{relint} \mathcal{Z}$,

$$\nabla d_z(u, w) = \begin{pmatrix} \nabla d_u(u) + \alpha_v \left(d_v\left(\frac{w}{h(u)}\right) - \nabla d_v\left(\frac{w}{h(u)}\right)^\top \frac{w}{h(u)} \right) a \\ \alpha_v \nabla d_v\left(\frac{w}{h(u)}\right) \end{pmatrix}. \tag{17}$$

(Note that (17) only needs the evaluation of ∇d_u and ∇d_v in the relative interior of $\mathcal{U}$ and $\mathcal{V}$, respectively, because $\operatorname{relint} \mathcal{Z} = (\operatorname{relint} \mathcal{U}) \triangleleft^h (\operatorname{relint} \mathcal{V})$ by Lemma 2.)

- *Strong convexity*: We now verify that d_z is strongly convex on $\operatorname{relint} \mathcal{Z}$. We start from the following result, which follows easily from algebraic manipulation (see the Online Appendix for details).

Lemma 3. The *function* d_z *as defined in Proposition 1 satisfies*

$$(\nabla d_z(u, w) - \nabla d_z(u', w'))^\top \begin{pmatrix} u - u' \\ w - w' \end{pmatrix}$$
$$\ge (\nabla d_u(u) - \nabla d_u(u'))^\top (u - u')$$
$$+ \alpha_v h\left(\frac{u + u'}{2}\right) \left\| \frac{w}{h(u)} - \frac{w'}{h(u)} \right\|_2^2 \tag{18}$$

for all $(u, w), (u', w') \in \operatorname{relint} \mathcal{Z} = (\operatorname{relint} \mathcal{U}) \triangleleft^h (\operatorname{relint} \mathcal{V})$.

Because d_u is strongly convex and because $\|\cdot\|_u$, $\|\cdot\|_v$ are norms for finite-dimensional spaces, there exist constants $\mu_u, \mu_v > 0$ such that

$$(\nabla d_u(u) - \nabla d_u(u'))^\top (u - u') \ge \mu_u \|u - u'\|_2^2,$$

and

$$\left\| \frac{w}{h(u)} - \frac{w'}{h(u)} \right\|_v^2 \ge \mu_v \left\| \frac{w}{h(u)} - \frac{w'}{h(u)} \right\|_2^2,$$

and thus, we can write

$$(\nabla d_z(u, w) - \nabla d_z(u', w'))^\top \begin{pmatrix} u - u' \\ w - w' \end{pmatrix}$$
$$\ge \mu_u \|u - u'\|_2^2 + \mu_v \alpha_v h\left(\frac{u + u'}{2}\right) \left\| \frac{w}{h(u)} - \frac{w'}{h(u)} \right\|_2^2. \tag{19}$$

Now,

$$\|w - w'\|_2 = \left\| \frac{h(u) + h(u')}{2} \left(\frac{w}{h(u)} - \frac{w'}{h(u')} \right) + (h(u) - h(u')) \cdot \frac{1}{2} \left(\frac{w}{h(u)} + \frac{w'}{h(u')} \right) \right\|_2$$
$$\le \frac{h(u) + h(u')}{2} \left\| \frac{w}{h(u)} - \frac{w'}{h(u')} \right\|_2 + |h(u) - h(u')| \cdot \left\| \frac{1}{2} \left(\frac{w}{h(u)} + \frac{w'}{h(u')} \right) \right\|_2$$
$$= h\left(\frac{u + u'}{2}\right) \left\| \frac{w}{h(u)} - \frac{w'}{h(u')} \right\|_2 + |h(u) - h(u')| \cdot \left\| \frac{1}{2} \left(\frac{w}{h(u)} + \frac{w'}{h(u')} \right) \right\|_2.$$

Because $w/h(u), w'/h(u') \in \operatorname{relint} \mathcal{V} \subseteq \mathcal{V}$ and $\mathcal{V}$ is a convex set, then the argument of the last norm on the right

is a point in $\mathcal{V}$. Furthermore, because $\mathcal{U}$ and $\mathcal{V}$ are compact, there exist constants $\Omega_u, \Omega_v > 0$ such that $\max_{u\in\mathcal{V}} \|u\|_2 \le \Omega_u$ and $\max_{v\in\mathcal{V}} \|v\|_2 \le \Omega_v$. Hence, we can write

$$\|w - w'\|_2 \le h\left(\frac{u+u'}{2}\right)\left\|\frac{w}{h(u)} - \frac{w'}{h(u')}\right\|_2 + (\|a\|_2\,\Omega_v)\|u - u'\|_2,$$

which, in turn, implies that

$$\begin{aligned}&\left\|\begin{pmatrix}u\\w\end{pmatrix} - \begin{pmatrix}u'\\w'\end{pmatrix}\right\|_2\\ &\le \|u - u'\|_2 + \|w - w'\|_2\\ &\le (1 + \|a\|_2\,\Omega_v)\|u - u'\|_2 + h\left(\frac{u+u'}{2}\right)\left\|\frac{w}{h(u)} - \frac{w'}{h(u')}\right\|_2.\end{aligned}$$

Squaring and using the Cauchy-Schwarz inequality yields

$$\begin{aligned}&\left\|\begin{pmatrix}u\\w\end{pmatrix} - \begin{pmatrix}u'\\w'\end{pmatrix}\right\|_2^2\\ &\le \left((1 + \|a\|_2\,\Omega_v)\|u - u'\|_2 + h\left(\frac{u+u'}{2}\right)\left\|\frac{w}{h(u)} - \frac{w'}{h(u')}\right\|_2\right)^2\\ &\le \left(\frac{(1 + \|a\|_2\Omega_u)^2}{\mu_u} + \frac{h\left(\frac{u+u'}{2}\right)}{\mu_v\alpha_v}\right)\left(\mu_u\|u - u'\|_2^2 + \mu_v\alpha_v h\left(\frac{u+u'}{2}\right)\left\|\frac{w}{h(u)} - \frac{w'}{h(u')}\right\|_2^2\right)\\ &\le \left(\frac{(1 + \|a\|_2\Omega_u)^2}{\mu_u} + \frac{\|a\|_2\Omega_u}{\mu_v\alpha_v}\right)\left(\mu_u\|u - u'\|_2^2 + \mu_v\alpha_v h\left(\frac{u+u'}{2}\right)\left\|\frac{w}{h(u)} - \frac{w'}{h(u')}\right\|_2^2\right);\end{aligned}$$

that is,

$$\mu_u\|u - u'\|_2^2 + \mu_v\alpha_v h\left(\frac{u+u'}{2}\right)\left\|\frac{w}{h(u)} - \frac{w'}{h(u')}\right\|_2^2 \ge \frac{1}{\frac{(1+\|a\|_2\Omega_u)^2}{\mu_u} + \frac{\|a\|_2\Omega_u}{\mu_v\alpha_v}}\left\|\begin{pmatrix}u\\w\end{pmatrix} - \begin{pmatrix}u'\\w'\end{pmatrix}\right\|_2^2.$$

Finally, plugging the above inequality into (19) yields

$$(\nabla d_z(u,w) - \nabla d_z(u',w'))^\top\begin{pmatrix}u - u'\\w - w'\end{pmatrix} \ge \frac{1}{\frac{(1+\|a\|_2\Omega_u)^2}{\mu_u} + \frac{\|a\|_2\Omega_u}{\mu_v\alpha_v}}\left\|\begin{pmatrix}u\\w\end{pmatrix} - \begin{pmatrix}u'\\w'\end{pmatrix}\right\|_2^2.$$

Hence, d_z is strongly convex (with respect to the Euclidean norm) with strong-convexity modulus $1/\left(\frac{(1+\|a\|_2\Omega_u)^2}{\mu_u} + \frac{\|a\|_2\Omega_u}{\mu_v\alpha_v}\right) > 0$.

- *Minimum of d_z*: Because, by hypothesis $\min_{u\in\mathcal{U}} d_u(u) = \min_{v\in\mathcal{V}} d_v(v) = 0$, $\alpha_v > 0$, and $h(u) \ge 0$ for all $u \in \mathcal{U}$, from the Definition (16), it follows that $d_z(u,w) \ge 0$ for all $(u,w) \in \mathcal{Z}$. So, in order to conclude that $\min_{(u,w)\in\mathcal{Z}} d_z(u,w) = 0$, it is enough to show that the value 0 can be attained by d_z. That can be checked directly by considering the point $(u^*, h(u^*)v^*) \in \mathcal{Z}$, where $u^* := \arg\min_{u\in\mathcal{U}} d_u(u) = \nabla d_u^*(\mathbf{0})$ and $v^* := \arg\min_{v\in\mathcal{V}} d_v(v) = \nabla d_v^*(\mathbf{0})$. Furthermore, by definition of the distance-generating function (Definition 1), the prox-centers of d_u and d_v lie in relint $\mathcal{U}$ and relint $\mathcal{V}$, respectively. Hence, using Lemma 2, we conclude that the prox-center $(u^*, h(u^*)v^*)$ constructed above belongs to relint $\mathcal{Z}$, as needed by the definition of DGF.

So, d_z is a DGF. We now focus on the second part of the statement.

- *Complexity of operations*: The key observation is that the gradient of the convex conjugate at any point $g = (g_u, g_v) \in \mathbb{R}^m \times \mathbb{R}^n$ satisfies

$$\begin{aligned}&\max_{\substack{u\in\mathcal{U}\\v\in\mathcal{V}}}\{g_u^\top u + h(u)g_v^\top v - d_u(u) - \alpha_v h(u) d_v(v)\}\\ &= \max_{\substack{u\in\mathcal{U}\\v\in\mathcal{V}}}\left\{g_u^\top u - d_u(u) + \alpha_v h(u)\left[\left(\frac{g_v}{\alpha_v}\right)^\top v - d_v(v)\right]\right\}\\ &= \max_{u\in\mathcal{U}}\left\{g_u^\top u - d_u(u) + \alpha_v h(u)\max_{v\in\mathcal{V}}\left\{\left(\frac{g_v}{\alpha_v}\right)^\top v - d_v(v)\right\}\right\}\\ &= \max_{u\in\mathcal{U}}\left\{\left(g_u + \alpha_v d_v^*\left(\frac{g_v}{\alpha_v}\right)a\right)^\top u - d_u(u)\right\},\end{aligned}$$

where the second equality follows because $\alpha_v > 0$ and $h(u) \ge 0$ by hypothesis, and the third equality follows from the definition of $h : u \longmapsto a^\top u$. Hence,

$$\nabla d_z^*(g_u, g_v) = \begin{pmatrix}\nabla d_u^*\left(g_u + \alpha_v d_v^*\left(\frac{g_v}{\alpha_v}\right)a\right)\\ h\left(\nabla d_u^*\left(g_u + \alpha_v d_v^*\left(\frac{g_v}{\alpha_v}\right)a\right)\right)\nabla d_v^*\left(\frac{g_v}{\alpha_v}\right)\end{pmatrix} \in \mathcal{U} \overset{h}{\triangleleft} \mathcal{V} = \mathcal{Z}. \tag{20}$$

This shows that ∇d_z^* can be computed starting from ∇d_u^* and ∇d_v^* as follows:

- First, we construct the vector g_v/α_v, which costs a number of arithmetic operations equal to the dimension n of $\mathcal{V} \subseteq \mathbb{R}^n$.
- Then, we compute $\nabla d_v^*(g_v/\alpha_v)$.
- From there, we compute

$$\begin{aligned}d_v^*\left(\frac{g_v}{\alpha_v}\right) &:= \max_{v\in\mathcal{V}}\left\{\left(\frac{g_v}{\alpha_v}\right)^\top v - d_v(v)\right\}\\ &= \left(\frac{g_v}{\alpha_v}\right)^\top \nabla d_v^*\left(\frac{g_v}{\alpha_v}\right) - d_v\left(\nabla d_v^*\left(\frac{g_v}{\alpha_v}\right)\right),\end{aligned}$$

which incurs an operation cost equal to n (for computing the dot product), plus a contribution equal to the complexity of evaluating d_v.

- Then, we compute

$$\nabla d_u^*\left(\boldsymbol{g}_u + \alpha_v\, d_v^*\left(\frac{\boldsymbol{g}_v}{\alpha_v}\right)\boldsymbol{a}\right)$$

for a cost of $n + \|\boldsymbol{a}\|_0$ (to evaluate the argument), plus a single call to the oracle for ∇d_u^*.

- Last, we evaluate $h\left(\nabla d_u^*\left(\boldsymbol{g}_u + \alpha_v\, d_v^*\left(\frac{\boldsymbol{g}_v}{\alpha_v}\right)\boldsymbol{a}\right)\right)$ (for a cost of $\|\boldsymbol{a}\|_0$) and use that to multiply $\nabla d_v^*\left(\frac{\boldsymbol{g}_v}{\alpha_v}\right)$ computed previously.

The gradient ∇d_z, whose formula was given in (17), can be computed similarly: we only need to take inner products, which takes time linear in the dimension of $\mathcal{V}$; compute the value of $h(\boldsymbol{u})$, which is linear in $\|\boldsymbol{a}\|_0$; and compute corresponding values of d_u and d_v (via a single oracle call to each). □

By repeatedly composing the construction of Proposition 1 using DGFs d_k for each $\mathcal{X}_k$ in setup 1, we obtain the DGF

$$d : \mathcal{X}_1 \overset{h_1}{\triangleleft} \mathcal{X}_2 \overset{h_2}{\triangleleft} \cdots \overset{h_{n-1}}{\triangleleft} \mathcal{X}_n \ni \boldsymbol{x} := (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \longmapsto \alpha_1\, d_1(\boldsymbol{x}_1) + \sum_{k=2}^{n} \alpha_k d_k^{\square}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_k), \tag{21}$$

where, for all $k = 2, \dots, n$,

$$d_k^{\square}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_k) := \begin{cases} 0 & \text{if } h_{k-1}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{k-1}) = 0 \\ h_{k-1}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{k-1}) & \\ \quad d_k\left(\dfrac{\boldsymbol{x}_k}{h_{k-1}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{k-1})}\right) & \text{if } h_{k-1}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{k-1}) > 0. \end{cases}$$

If all DGFs d_k ($k = 1, \dots, n$) are nice, then Proposition 1 guarantees that d_z is also nice in the sense of Definition 6. This is directly shown by induction on the number n of scaled extensions being applied. When $n = 1$, there is nothing to prove because $d_z = d_1$. Assume now that the composite DGF d_z' constructed for set $\mathcal{X}_1 \overset{h_1}{\triangleleft} \mathcal{X}_2 \overset{h_2}{\triangleleft} \cdots \overset{h_{n-1}}{\triangleleft} \mathcal{X}_n$ is nice (in the sense of Definition 6). From the proposition, the DGF d_z for $\mathcal{X}_1 \overset{h_1}{\triangleleft} \mathcal{X}_2 \overset{h_2}{\triangleleft} \cdots \overset{h_n}{\triangleleft} \mathcal{X}_{n+1}$ constructed by an additional application of Proposition 1 is such that $d_z, \nabla d_z$, and ∇d_z^* each require a single oracle evaluation of $d_z', \nabla d_z', \nabla(d_z')^*$—whose cost is linear in $\sum_{k=1}^{n} s_k + \sum_{k=1}^{n-1} \|\boldsymbol{a}_k\|_0$ by definition of niceness for d_z'—plus an overhead linear in the costs of evaluating $d_{n+1}, \nabla d_{n+1}$, and ∇d_{n+1}^*, as well as $\|\boldsymbol{a}_n\|_0$ and the dimension s_{n+1} of $\mathcal{X}_{n+1} \subseteq \mathbb{R}^{s_{n+1}}$. Because d_{n+1} is a nice DGF (in the sense of Section 3.1), then the overhead is linear in s_{k+1} and $\|\boldsymbol{a}_n\|_0$, establishing that each evaluation of $d_z, \nabla d_z$, or ∇d_z^* has a cost linear in $\sum_{k=1}^{n+1} s_k + \sum_{k=1}^{n} \|\boldsymbol{a}_k\|_0$; that is, d_k is a nice DGF in the sense of Definition 6, as we wanted to show.

In the case of the sequence-form polytope $\mathcal{X} = Q$, where $\mathcal{X}_k = \Delta^{n_k}$ and each h_k maps a sequence-form vector to the parent sequence of the k-th decision point, this yields exactly a dilated DGF in the precise sense of (6). So, the construction in Proposition 1 subsumes that of Section 4.2, which only applied to sequence-form strategy spaces. Additionally, in the case of the polytope of correlation plans $\mathcal{X} = \Xi$, to our knowledge, our construction yields the first nice DGF.

However, before the DGF (21) can be put to use in an optimization method such as EGT or mirror prox, the weights α_k need to be specified so that the DGF is 1-strongly convex with respect to some norm. The next proposition provides a general way of doing so, with respect to the Euclidean norm, under conditions met for both $\mathcal{X} = Q$ and $\mathcal{X} = \Xi$. From now on, we will denote the i-th coordinate of a vector $\boldsymbol{v}$ with the symbol $v[i]$.

Proposition 2. *Consider the composite DGF defined in* (21) *for the set* $\mathcal{X}$ (13) *under setup* 1. *Then, for all* $\boldsymbol{x} := (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$ *and* $\boldsymbol{x}' := (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$ *such that* $\boldsymbol{x}, \boldsymbol{x}' \in \operatorname{relint} \mathcal{X} = (\operatorname{relint} \mathcal{X}_1) \overset{h_1}{\triangleleft} \cdots \overset{h_{n-1}}{\triangleleft} (\operatorname{relint} \mathcal{X}_n)$, *the composite DGF d satisfies*

$$\langle \nabla d(\boldsymbol{x}) - \nabla d(\boldsymbol{x}'), \boldsymbol{x} - \boldsymbol{x}' \rangle \geq \sum_{k=1}^{n} \sum_{i=1}^{s_k} \left(\frac{\alpha_k}{2} - \sum_{p=k}^{n-1} \alpha_{p+1} \|\boldsymbol{a}_p\|_0\, a_{p,k}[i] \right) (x_k[i] - x_k'[i])^2.$$

Proof. By induction on the sequence of scaled extension operations,

- *Base case*: The base case corresponds to the case where $n = 1$ and $\mathcal{X} = \mathcal{X}_1 \subseteq \mathbb{R}^{s_1}$; that is, no scaled extension is performed. In that case, $d : \mathcal{X} \ni \boldsymbol{x} \longmapsto \alpha_1 d_1(\boldsymbol{x})$, where d_1 is a 1-strongly convex DGF for $\mathcal{X}_1$ with respect to the Euclidean norm by hypothesis. Hence,

$$\begin{aligned} &\langle \nabla d(\boldsymbol{x}) - \nabla d(\boldsymbol{x}'), \boldsymbol{x} - \boldsymbol{x}' \rangle \\ &\quad \geq \alpha_1 \|\boldsymbol{x} - \boldsymbol{x}'\|_2^2 = \sum_{i=1}^{s_1} \alpha_1 (x[i] - x'[i])^2 \\ &\quad \geq \sum_{i=1}^{s_1} \frac{\alpha_1}{2} (x[i] - x'[i])^2, \end{aligned}$$

which satisfies the statement.

- *Inductive step*: Let $\mathcal{U} := \mathcal{X}_1 \overset{h_1}{\triangleleft} \dots \overset{h_{n-2}}{\triangleleft} \mathcal{X}_{n-1}$, and let $\tilde{d}$ be the dilated DGF constructed for $\mathcal{U}$. Assume by induction that the statement of this proposition holds for $\tilde{d}$. We will show that the statement continues to hold after one further application of the construction, that is, for the dilated DGF

$$d : (\mathcal{U} \triangleleft^{h_{n-1}} \mathcal{X}_n) \ni (\boldsymbol{u}, \boldsymbol{w}) \longmapsto \tilde{d}(\boldsymbol{u}) + \alpha_n d_n^{\square}(\boldsymbol{u}, \boldsymbol{w}),$$

where

$$d_n^{\square}(\boldsymbol{u}, \boldsymbol{w}) := \begin{cases} 0 & \text{if } h_{n-1}(\boldsymbol{u}, \boldsymbol{w}) = 0 \\ h_{n-1}(\boldsymbol{u}, \boldsymbol{w}) d_n\left(\dfrac{\boldsymbol{w}}{h(\boldsymbol{u})}\right) & \text{if } h_{n-1}(\boldsymbol{u}, \boldsymbol{w}) > 0. \end{cases}$$

To lighten notation, from now on in this proof, we will let h be a shorthand for h_{n-1}. From Lemma 3, it

follows that for any $(u,w),(u',w') \in \text{relint}(\mathcal{U} \overset{h}{\triangleleft} \mathcal{X}_n) = (\text{relint}\,\mathcal{U}) \overset{h}{\triangleleft} (\text{relint}\,\mathcal{X}_n)$,

$$\begin{aligned}&\left\langle \nabla d(u,w) - \nabla d(u',w'), \begin{pmatrix} u-u' \\ w-w' \end{pmatrix} \right\rangle \\ &\quad \geq \langle \nabla \tilde{d}(u) - \nabla \tilde{d}(u'), u-u' \rangle \\ &\qquad + \alpha_n h\left(\frac{u+u'}{2}\right) \left\| \frac{w}{h(u)} - \frac{w'}{h(u')} \right\|_2^2. \end{aligned} \tag{22}$$

Now,

$$\begin{aligned}\|w-w'\|_2 &= \left\| h(u)\frac{w}{h(u)} - h(u')\frac{w'}{h(u')} \right\|_2 \\ &= \left\| h\left(\frac{u+u'}{2}\right)\left(\frac{w}{h(u)} - \frac{w'}{h(u')}\right) \right. \\ &\qquad \left. + (h(u)-h(u'))\left[\frac{1}{2}\left(\frac{w}{h(u)} + \frac{w'}{h(u')}\right)\right] \right\|_2 \\ &\leq h\left(\frac{u+u'}{2}\right) \left\| \frac{w}{h(u)} - \frac{w'}{h(u')} \right\|_2 + (h(u)-h(u')),\end{aligned}$$

where, in the last step, we used the triangle inequality and the hypothesis that $\max_{v \in \mathcal{X}_n} \|v\| \leq 1$ together with the fact that $\frac{1}{2}\left(\frac{w}{h(u)} + \frac{w'}{h(u')}\right) \in \text{relint}\,\mathcal{X}_n$ by convexity. By taking squares, rearranging, and using the fact that $h(u) \in [0,1]$ for all $u \in \mathcal{U}$, we have

$$\begin{aligned}&h\left(\frac{u+u'}{2}\right) \left\| \frac{w}{h(u)} - \frac{w'}{h(u')} \right\|_2^2 \\ &\quad \geq h\left(\frac{u+u'}{2}\right)^2 \left\| \frac{w}{h(u)} - \frac{w'}{h(u')} \right\|^2 \\ &\quad \geq \frac{1}{2}\|w-w'\|^2 - (h(u)-h(u'))^2 \\ &\quad = \frac{1}{2}\|w-w'\|^2 - (a_{n-1}^\top (u-u'))^2. \end{aligned} \tag{23}$$

Plugging the inductive hypothesis and (23) into (22), defining $u = (x_1,\dots,x_{n-1}), u' = (x_1',\dots,x_{n-1}') \in \text{relint}\,\mathcal{U}$, $w := x_n, x := (u,x_n)$, and $x' := (u',x_n')$ yields

$$\begin{aligned}&\langle \nabla d(x) - \nabla d(x'), x - x' \rangle \\ &\geq \left[\sum_{k=1}^{n-1} \sum_{i=1}^{s_k} \left(\frac{\alpha_k}{2} - \sum_{p=k}^{n-2} \alpha_{p+1} \|a_p\|_0 a_{p,k}[i] \right) (x_k[i] - x_k'[i])^2 \right] \\ &\quad + \frac{\alpha_n}{2}\|x_n - x_n'\|_2^2 - \alpha_n (a_{n-1}^\top (u-u'))^2 \\ &\geq \left[\sum_{k=1}^{n-1} \sum_{i=1}^{s_k} \left(\frac{\alpha_k}{2} - \sum_{p=k}^{n-2} \alpha_{p+1} \|a_p\|_0 a_{p,k}[i] \right) (x_k[i] - x_k'[i])^2 \right] \\ &\quad + \frac{\alpha_n}{2}\|x_n - x_n'\|_2^2 - \alpha_n \|a_{n-1}\|_0 \left(\sum_{q=1}^{n-1} \sum_{i=1}^{s_q} a_{n-1,q}[i](x_q[i] - x_q'[i])^2 \right) \\ &= \sum_{k=1}^{n} \sum_{i=1}^{s_k} \left(\frac{\alpha_k}{2} - \sum_{p=k}^{n-1} \alpha_{p+1} \|a_p\|_0 a_{p,k}[i] \right) (x_k[i] - x_k'[i])^2,\end{aligned}$$

where the second inequality follows from applying the Cauchy-Schwarz inequality and the fact that $\|a_p\|_0 \geq \|a_p\|_1$ because $a_p \in [0,1]^{s_1+\cdots+s_{p-1}}$ by the hypotheses in setup 1. □

Proposition 2 implies that when the α_k's are such that

$$\frac{\alpha_k}{2} - \sum_{p=k}^{n-1} \alpha_{p+1} \|a_p\|_0 a_{p_k}[i] \geq 1$$
$$\text{for all } k = 1,\dots,n, i = 1,\dots,s_k,$$

then d is 1-strongly convex with respect to the Euclidean norm. So, we have the following.

Corollary 1. *Consider the composite DGF defined in (21) for the set $\mathcal{X}$ (13) under setup 1, where the coefficients α_k are defined recursively as*

$$\alpha_n = 2, \quad \text{and} \quad \alpha_k = 2 + 2 \left\| \sum_{p=k}^{n-1} \alpha_{p+1} \right\| a_p \|_0 \, a_{p,k} \|_\infty \quad \forall k = n-1,\dots,1.$$

Then, d is 1-strongly convex with respect to the Euclidean distance.

Because sets $\mathcal{X}$ obtained via scaled extension of *simplex* domains are prevalent in game theory, we show a stronger result for the case where each d_k is set to the negative entropy function.

Definition 7. Consider a set $\mathcal{X}$ obtained via scaled extension of *simplex* domains $\mathcal{X}_k = \Delta^{s_k}$ in accordance with setup 1, and consider the dilated DGF for $\mathcal{X}$ obtained as described in (21), when d_k is set to the negative entropy function

$$d_k : \Delta^{s_k} \ni x_k \longmapsto \log(s_k) + \sum_{i=1}^{s_k} x_k[i] \log x_k[i] \geq 0 \tag{24}$$

for all k. Then, the resulting function $\psi : (\mathcal{X}_1 \overset{h_1}{\triangleleft} \cdots \overset{h_{n-1}}{\triangleleft} \mathcal{X}_n) \to \mathbb{R}_{\geq 0}$ is

$$\begin{aligned}\psi : (x_1,\dots,x_n) \longmapsto{}& \alpha_1 \left(\log(s_1) + \sum_{i=1}^{s_1} x_1[i] \log x_1[i] \right) \\ &+ \sum_{k=2}^{n} \alpha_k \left(h_{k-1}(x_1,\dots,x_{k-1}) \log(s_k) \right. \\ &\left. + \sum_{i=1}^{s_k} x_k[i] \log \frac{x_k[i]}{h_{k-1}(x_1,\dots,x_{k-1})} \right)\end{aligned} \tag{25}$$

for the particular choice of weights α_k, defined in Corollary 1, and called the *dilated entropy DGF*.

Note that Definition 7 extends the name *dilated entropy DGF*, already used for the DGF in Definition 3

in the case of the sequence-form strategy $\mathcal{X} = Q$ to any scaled extension of simplexes. The overload is sound in the sense that when $\mathcal{X} = Q$, Definition 7 recovers exactly the function in Definition 3, together with the state-of-the-art coefficients defined by Kroer et al. (2017). A consequence of this observation is that the coefficients defined by Corollary 1 grow exponentially fast in the dimension of $\mathcal{X}$, showing that the composite DGF constructed by means of Proposition 1 suffers from the same problem as the Kroer et al. dilated entropy DGF discussed in Section 4.2. We show that the strong-convexity result of Kroer et al. (2020) for treeplexes generalizes as well; the proof can be found in Online Appendix C.

Proposition 3. *Let $\mathcal{X}$ be obtained via scaled extension of simplex domains $\mathcal{X}_k = \Delta^{s_k}$. Then, the dilated entropy DGF* (Definition 7) *for $\mathcal{X}$ is 1-strongly convex with respect to the Euclidean norm and $(1/M_{\mathcal{X}})$-strongly convex with respect to the ℓ_1 norm.*

In the next section, we show how the ideas already used in Section 5 generalize to scaled extension–based sets, yielding the first nice DGF with polynomially small range on $\mathcal{X}$.

6.4. Dilatable Global Entropy for Scaled Extension

We instantiate the generic framework of dilated DGFs, as defined in the previous section, to the chains of scaled extensions with *simplex domains*. More specifically, we consider the same setting as setup 1 under the further assumption that each convex and compact set $\mathcal{X}_k$ in the decomposition of $\mathcal{X}$ (13) is a probability simplex; that is, for all $k = 1, \dots, n$, $\mathcal{X}_k = \Delta^{s_k}$, where $s_k \in \mathbb{N}_{\geq 1}$. This setup encompasses both sequence-form strategy spaces and the polytope of correlation plans.

The *dilated entropy DGF* ψ for such a set is the dilated DGF (21) obtained by recursively applying the general construction of Proposition 1 with the (negative) entropy function at each Δ^{s_k}. Specifically, ψ can be written as[6]

$$\begin{aligned}\psi : \mathcal{X} \ni (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \mapsto \alpha_1 \left(\log s_1 + \sum_{i=1}^{s_1} x_1[i] \log x_1[i] \right) \\ + \sum_{k=2}^{n} \alpha_k \left(h_{k-1}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{k-1}) \log s_k \right. \\ \left. + \sum_{i=1}^{s_k} x_k[i] \log \frac{x_k[i]}{h_{k-1}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{k-1})} \right).\end{aligned}$$

By using the same manipulations of the logarithms that we used in Theorem 3, in Online Appendix B, we show that ψ coincides on $\mathcal{X}$ with the function

$$\begin{aligned}\tilde{\varphi} : \mathcal{X} \ni (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \mapsto \sum_{k=1}^{n} \left(\alpha_k \sum_{i=1}^{s_k} x_k[i] \log x_k[i] \right) \\ - \sum_{k=2}^{n} \alpha_k h_{k-1}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{k-1}) \log h_{k-1}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{k-1}) \\ + \alpha_1 \log s_1 + \sum_{k=2}^{n} \alpha_k h_{k-1}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{k-1}) \log s_k. \qquad (26)\end{aligned}$$

For this reason, similarly to what we did for extensive-form strategy spaces, we coin $\tilde{\varphi}_k$ the *dilatable global entropy DGF*. It is immediate to see by induction that $\nabla\psi$ can be computed exactly in linear time. Furthermore, because ψ is a nice DGF by virtue of Proposition 1, and $\psi = \tilde{\varphi}$ on $\mathcal{X}$, we immediately obtain that $\tilde{\varphi}$ is a nice DGF.

We conclude this section by showing that there exists polynomially small (in the dimension of $\mathcal{X}$) DGF weights α_k such that $\tilde{\varphi}$ defined in (26) is 1-strongly convex with respect to the Euclidean norm. For any $\boldsymbol{m} := (\boldsymbol{m}_1, \dots, \boldsymbol{m}_n) \in \mathbb{R}^{s_1} \times \cdots \times \mathbb{R}^{s_n}$ and $\boldsymbol{x} := (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \in \mathcal{X}$, the Hessian matrix of $\tilde{\varphi}$ satisfies

$$\begin{aligned}&\boldsymbol{m}^\top \nabla^2 \tilde{\varphi}(\boldsymbol{x}) \boldsymbol{m} \\ &= \sum_{k=1}^{n} \left(\alpha_k \sum_{i=1}^{s_k} \frac{m_k[i]^2}{x_k[i]} \right) - \sum_{k=2}^{n} \left(\alpha_k \frac{\left(\sum_{p=1}^{k-1} \boldsymbol{a}_{k-1,p}^\top \boldsymbol{m}_p\right)^2}{h_{k-1}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{k-1})} \right) \\ &= \sum_{k=1}^{n} \left(\alpha_k \sum_{i=1}^{s_k} \frac{m_k[i]^2}{x_k[i]} \right) - \sum_{k=2}^{n} \left(\alpha_k \frac{\left(\sum_{p=1}^{k-1} \boldsymbol{a}_{k-1,p}^\top \boldsymbol{m}_p\right)^2}{\sum_{p=1}^{k-1} \boldsymbol{a}_{k-1,p}^\top \boldsymbol{x}_p} \right), \qquad (27)\end{aligned}$$

where in the second equality, we expanded the definition of h_{k-1} according to (14). Now, expanding the following product, we find that

$$\begin{aligned}&\|\boldsymbol{a}_{k-1}\|_0 \left(\sum_{p=1}^{k-1} \boldsymbol{a}_{k-1,p}^\top \boldsymbol{x}_p \right) \left(\sum_{p=1}^{k-1} \sum_{i=1}^{s_p} \frac{a_{k-1,p}[i] m_p[i]^2}{x_p[i]} \right) \\ &\geq \|\boldsymbol{a}_{k-1}\|_0 \left(\sum_{p=1}^{k-1} \sum_{i=1}^{s_p} a_{k-1,p}[i] x_p[i] \frac{a_{k-1,p}[i] m_p[i]^2}{x_p[i]} \right) \\ &= \|\boldsymbol{a}_{k-1}\|_0 \left(\sum_{p=1}^{k-1} \sum_{i=1}^{s_p} a_{k-1,p}[i]^2 m_p[i]^2 \right) \\ &\geq \left(\sum_{p=1}^{k-1} \sum_{i=1}^{s_p} a_{k-1,p}[i] m_p[i] \right)^2 \\ &= \left(\sum_{p=1}^{k-1} \boldsymbol{a}_{k-1,p}^\top \boldsymbol{m}_p \right)^2.\end{aligned}$$

Hence, plugging the above inequality into (27), we have

$$\begin{aligned}\boldsymbol{m}^\top \nabla^2 \tilde{\varphi}(\boldsymbol{x})\boldsymbol{m} &\geq \sum_{k=1}^{n}\left(\alpha_k \sum_{i=1}^{s_k}\frac{m_k[i]^2}{x_k[i]}\right) \\ &\quad - \sum_{k=2}^{n}\left(\alpha_k \|\boldsymbol{a}_{k-1}\|_0 \sum_{p=1}^{k-1}\sum_{i=1}^{s_p}\frac{a_{k-1,p}[i]\, m_p[i]^2}{x_p[i]}\right) \\ &= \sum_{k=1}^{n}\sum_{i=1}^{s_k}\left(\alpha_k - \sum_{p=k}^{n-1}\alpha_{p+1}\|\boldsymbol{a}_p\|_0 a_{p,k}[i]\right)\frac{m_k[i]^2}{x_k[i]} \qquad (28) \\ &\geq \sum_{k=1}^{n}\sum_{i=1}^{s_k}\left(\alpha_k - \sum_{p=k}^{n-1}\alpha_{p+1}\|\boldsymbol{a}_p\|_0 a_{p,k}[i]\right) m_k[i]^2, \qquad (29)\end{aligned}$$

where the last inequality follows from the fact that all entries of $\boldsymbol{x} \in \mathcal{X} = \Delta^{s_1} \overset{h_1}{\triangleleft} \cdots \overset{h_{n-1}}{\triangleleft} \Delta^{s_n}$ belong to $[0,1]$, given the assumption that $h_k(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_k) \in [0,1]$ for all $k = 1, \ldots, n-1$.

In particular, (29) implies that when the coefficients α_k are chosen so that

$$\alpha_k - \sum_{p=k}^{n-1}\alpha_{p+1}\|\boldsymbol{a}_p\|_0 a_{p,k}[i] \geq 1 \qquad \forall k = 1,\ldots,n, \quad i = 1,\ldots,s_k, \qquad (30)$$

then $\tilde{\varphi}$ is 1-strongly convex with respect to the Euclidean norm, and for the ℓ_1 norm, we have

$$\begin{aligned}\|\boldsymbol{m}\|_1^2 = \left(\sum_{k=1}^{n}\sum_{i=1}^{s_k} m_k[i]\right)^2 &= \left(\sum_{k=1}^{n}\sum_{i=1}^{s_k}\frac{m_k[i]}{\sqrt{x_k[i]}}\sqrt{x_k[i]}\right)^2 \\ &\leq \left(\sum_{k=1}^{n}\sum_{i=1}^{s_k}\frac{m_k[i]^2}{x_k[i]}\right)\left(\sum_{k=1}^{n}\sum_{i=1}^{s_k} x_k[i]\right) \\ &\leq M_{\mathcal{X}}\boldsymbol{m}^\top\nabla^2\tilde{\varphi}(\boldsymbol{x})\boldsymbol{m},\end{aligned}$$

which follows by Cauchy-Schwarz, (30), and (28). In other words, we have the following:

Theorem 7. *Consider the dilatable global entropy DGF defined in* (26) *for the set* $\mathcal{X}$ (13) *under setup* 1 *and the further assumption that each* $\mathcal{X}_k = \Delta^{s_k}$*, where the coefficients* α_k *are defined recursively as*

$$\alpha_n = 1, \quad \text{and} \quad \alpha_k = 1 + \left\|\sum_{p=k}^{n-1}\alpha_{p+1}\|\boldsymbol{a}_p\|_0\, \boldsymbol{a}_{p,k}\right\|_\infty \qquad \forall k = n-1,\ldots,1.$$

Then, d *is* 1*-strongly convex with respect to the Euclidean distance and* $(1/M_{\mathcal{X}})$*-strongly convex with respect to the* ℓ_1 *norm.*

Theorem 7 gives an analogue of Theorem 4 and Theorem 5 for the case of scaled extensions. Next, one would similarly want an analogue of Theorem 6 for scaled extension: a bound on the polytope diameter as measured by $\tilde{\varphi}$. By dilatability, ψ and $\tilde{\varphi}$ coincide on $\mathcal{X}$, so their diameters are the same. Furthermore, the minimum of ψ is zero, and so, the diameter of $\tilde{\varphi}$ is equal to $\max_{\boldsymbol{x}\in\mathcal{X}}\psi(\boldsymbol{x})$.

By using the fact that $a\log\left(\frac{a}{b}\right) = b\cdot\left(\frac{a}{b}\log\left(\frac{a}{b}\right)\right) \leq 0$, for all $0 < a \leq b$, we can then upper bound the expression in (25) as

$$\psi(\boldsymbol{x}) \leq \alpha_1\log(s_1) + \sum_{k=2}^{n}\alpha_k h_{k-1}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_{k-1})\log(s_k). \qquad (31)$$

Thus, this implies a bound on the diameter, similar to that of the sequence-form case (cf. Theorem 6), stated next.

Proposition 4. *For any set* $\mathcal{X}$ *satisfying setup* 1*, the diameter of the global entropy DGF* $\tilde{\varphi}$ *satisfies the inequality*

$$\Omega_{\tilde{\varphi},\mathcal{X}} \leq \max_{\boldsymbol{x}\in\mathcal{X}}\left\{\alpha_1\log(s_1) + \sum_{k=2}^{n}\alpha_k h_{k-1}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_{k-1})\log(s_k)\right\}.$$

In the sequence-form case, we were able to simplify this bound further by noting that in this case, $h_{k-1}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k-1})$ is replaced by x_{p_j}, which allows us to express the bound in terms of the maximum ℓ_1 norm. Unfortunately, such a simplification is not possible here because the scaling functions h_k are abstract. For a specific application of scaled extensions, one would need to use problem-specific structure in order to get a more explicit closed-form bound based on Equation (31). At any rate, we note that by using the trivial upper bound $h_{k-1}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_{k-1}) \leq 1$ implied by setup 1, we nonetheless obtain, as a corollary of the previous proposition, the following loose, problem-independent guarantee on the diameter.

Corollary 2. *Under the same hypotheses as Proposition 4,* $\Omega_{\tilde{\varphi},\mathcal{X}} \leq \alpha_1\log(s_1) + \sum_{k=2}^{n}\alpha_k\log(s_k)$.

7. Experiments

In this section, we study the numerical performance of our DGFs. First, we study the performance of the dilatable global entropy for computing Nash equilibria in zero-sum EFGs, and second, we study the performance for computing correlated equilibria and team equilibria.

Our experiments will be shown on nine different games, which span a variety of poker games, other recreational games, as well as a pursuit-evasion game played on a graph. All games are standard benchmarks in the computational game theory literature, and a full description of the games is given in the Online Appendix. In Table 1(a), we summarize some key dimensions of the game instances we use: the number of decision points $|\mathcal{J}_1|, |\mathcal{J}_2|$ for player 1 and 2, respectively; the number of sequences $|\Sigma_1|, |\Sigma_2|$; and the number of terminal nodes (leaves).

Table 1. Key Properties of the Game Instances Used in the Experimental Evaluation

	Decision points		Sequences		Leaves	Weights β		Weights γ	
Game instance	$\lvert\mathcal{J}_1\rvert$	$\lvert\mathcal{J}_2\rvert$	$\lvert\Sigma_1\rvert$	$\lvert\Sigma_2\rvert$	$\lvert Z\rvert$	Average	Max	Average	Max
Kuhn poker	6	6	13	13	30	8.857	38	2.286	7
Leduc poker (3 ranks)	144	144	337	337	1,116	11.766	686	2.117	43
Leduc poker (13 ranks)	2,574	2,574	6,007	6,007	98,956	12.057	12,326	2.131	703
Goofspiel	3,652	3,652	7,505	7,505	13,824	5.316	5,010	1.500	341
Battleship (3 turns)	18,152	62,875	73,130	253,940	552,132	3.294	2,894	1.242	99
Battleship (4 turns)	316,520	734,203	968,234	2,267,924	3,487,428	3.753	27,470	1.331	483
Liar's dice	6,144	6,144	18,427	18,427	147,420	14.885	32,762	2.054	865
Pursuit-evasion (4 turns)	34	348	52	2,029	15,898	8.286	62	1.943	5
Pursuit-evasion (6 turns)	58	11,830	78	68,951	118,514	19.424	254	2.508	7
	(a) — Game instances and sizes					(b)		(c)	

Notes. Column (a): various measures of the size of each of the games that we test algorithms on. Columns (b) and (c): the magnitude of the dilated entropy DGF and DGE weights.

Our experiments will show performance on three algorithms. First, we will plot the performance for both the EGT and mirror prox algorithms, with step sizes and smoothing chosen according to the theoretical values dictated by Theorems 1 and 2 with one minor variation: for each DGF, we do not scale by $M_{\mathcal{X}}$, which is required in order to achieve strong convexity with respect to the ℓ_1 norm. This is done because scaling by this last factor leads to extremely slow performance for all of the DGFs. Second, we will also show results on a tweaked variant of EGT called EGT/AS, which implements several heuristics that typically lead to better performance in practice, as seen in Hoda et al. (2010) and Kroer et al. (2020, 2018). These heuristics are:

1. *μ balancing*: At each iteration, we take a step on the player i whose smoothing parameter μ_i is larger.
2. *Aggressive step sizing*: The original step size of EGT at iteration t is $\tau = 2/(3+t)$, which is typically too conservative in practice. Instead, EGT/AS maintains some current value τ, initially set at $\tau = 0.5$. EGT/AS then repeatedly attempts to take steps with the current τ and, after every step, checks whether the invariant condition of EGT still holds. If not, then we undo the step, decrease τ, and repeat the process.
3. *Initial μ fitting*: The initial EGT values for μ_x, μ_y are much too conservative. Instead, at the beginning of the algorithm, we perform a search over initial values for $\mu_x = \mu_y$. The search starts at the candidate value $\mu = 10^{-6}$ and stops as soon as the choice of $\mu_x = \mu_y = \mu$ yields an excessive gap value above 0.1. If the current choice does not, μ is incremented by 20%, and the fitting continues.

For all parameters above, we use the same values as in Kroer et al. (2018), even though those values were tuned for the dilated entropy DGF, rather than dilatable global entropy.

In the presentation of the numerical performance, we will generally plot the number of iterations of the FOM on the x-axis, rather than plot wall-clock time. Because we hold the algorithmic setup fixed in each plot, apart from the DGF, this gives a fair representation of performance because they all use the same set of operations (in particular, the same number of gradient computations, which is typically the most expensive operation). For EGT/AS, we will instead plot the number of gradient computations on the x-axis because the number of gradient computations can vary for each DGF, depending on the amount of backtracking incurred.

7.1. Nash Equilibrium Computation

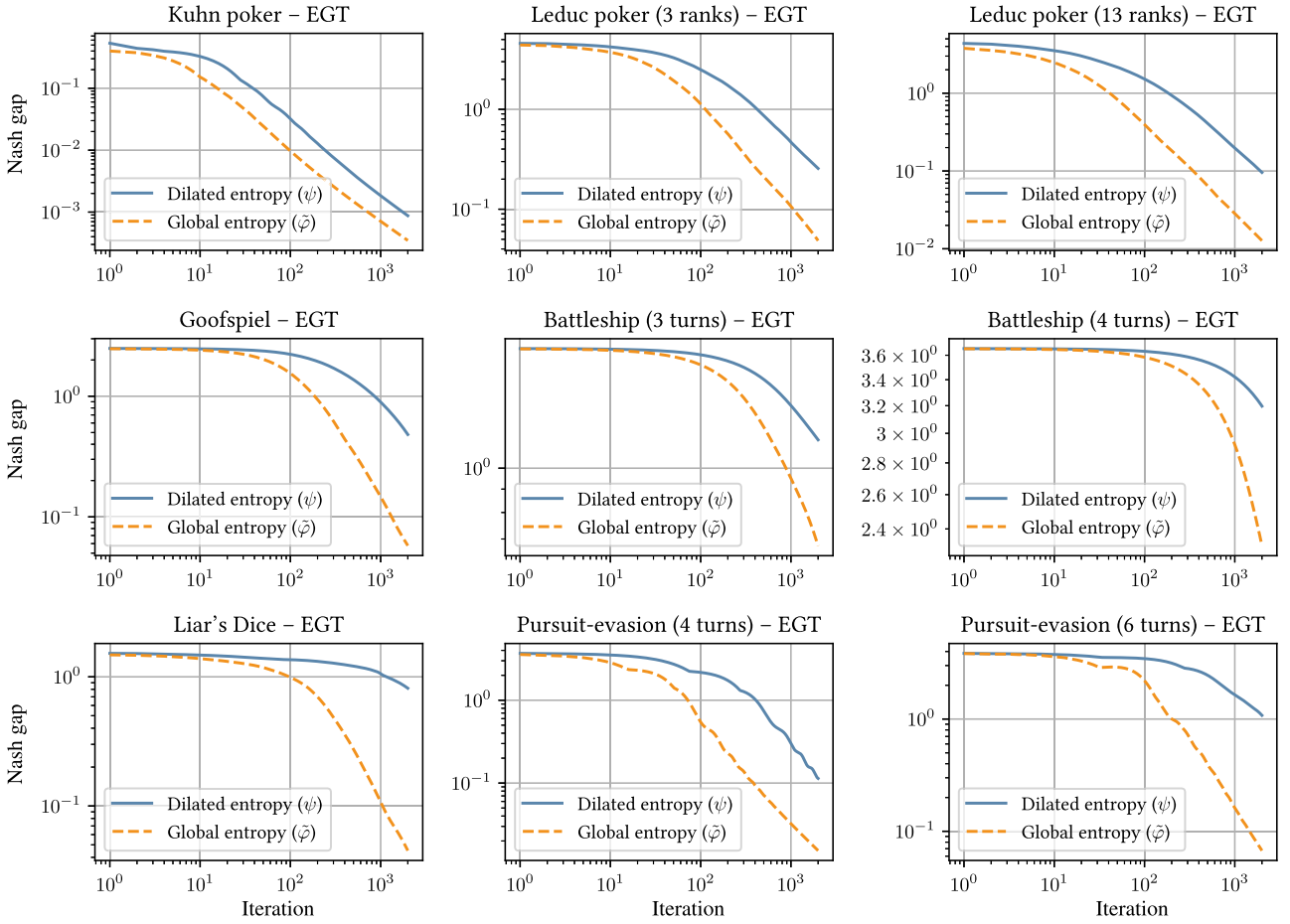
We will focus on comparing our new dilatable global entropy for the sequence-form polytope (Definition 4) to the prior state-of-the-art dilated entropy DGF (Definition 3) from Kroer et al. (2020).

Before we study the numerical performance, we look at the size of the DGF weights β and γ for each of the games. Table 1, column (b) shows the average and maximum size of the dilated entropy for player 1, and Table 1, column (c) shows the corresponding values for the DGE for the same player. We see that the DGE requires vastly less weight, especially in terms of the maximal weights near the root of each decision space.

First, we study the theoretically correct way to use the DGFs. In particular, we instantiate both EGT and mirror prox with the step sizes and DGFs as specified in Theorems 1 and 2 for the dilatable global entropy and dilated entropy. The results for EGT are shown in Figures 2 and 3. Across both algorithms and all nine games, we see that our new dilatable global entropy DGF performs better, sometimes by over an order of magnitude (e.g., in Liar's dice and pursuit evasion (six turns)). This is in line with the fact that our new DGF has a better strong-convexity modulus, which allows for a much smaller amount of smoothing while still guaranteeing correctness. This, in turn, allows the algorithms to safely take larger steps, thereby progressing faster.

Secondly, we investigate the numerical performance of the two entropy DGFs in the EGT/AS algorithm in

Figure 2. (Color online) Performance of the EGT Algorithm Instantiated with the Two Entropy DGFs Across Nine Games



Note. The x-axis shows the number of EGT iterations, and the y-axis shows the distance to Nash equilibrium.

Figure 4. Here, we see a smaller performance improvement. For most of the games, we get a small factor of improvement for the first 100 or so iterations, but then the performance is similar thereafter. For Liar's dice, there is a persistent improvement in using dilatable global entropy across all iterations.

7.2. Correlated and Team Equilibrium Computation

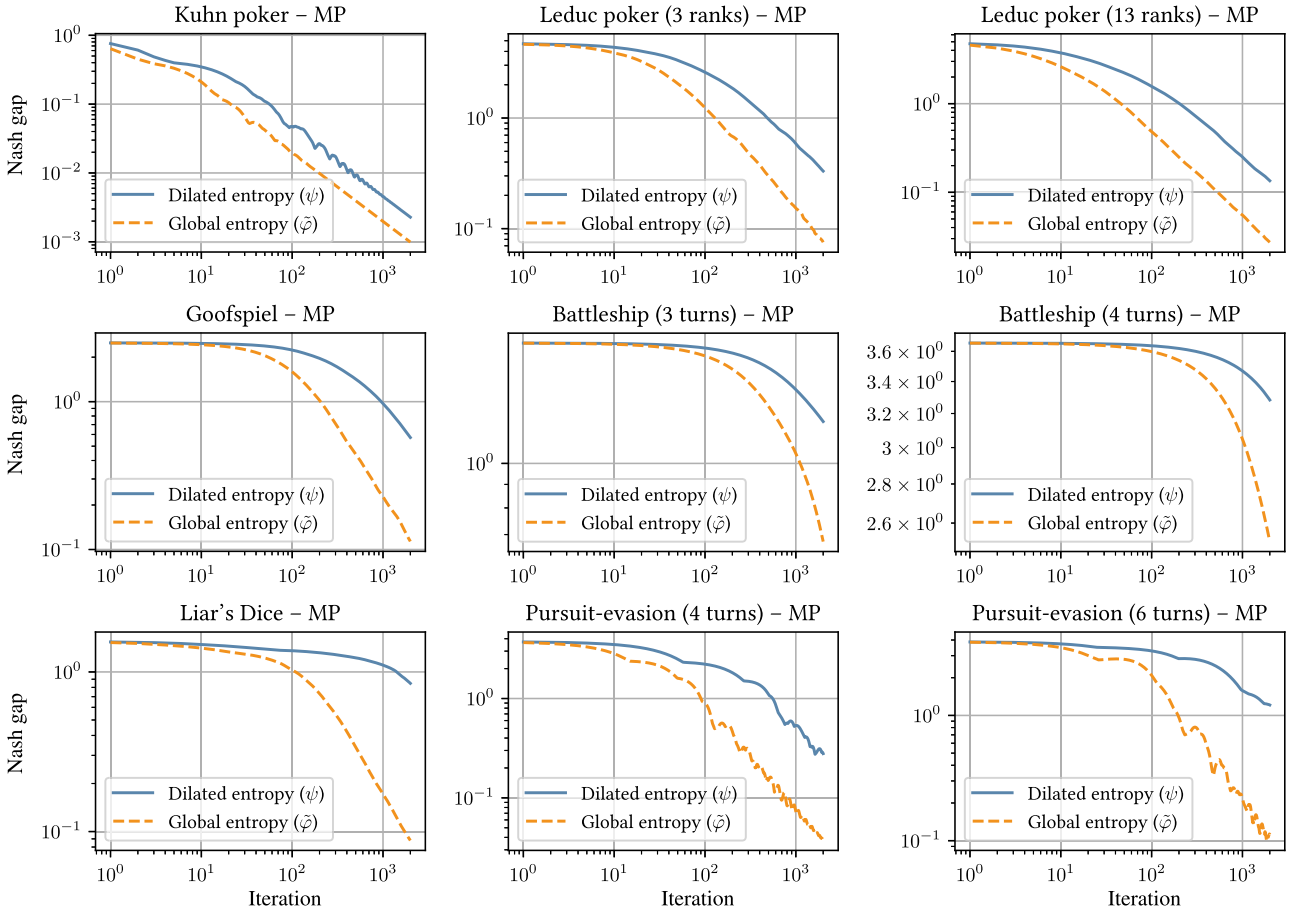
Next, we investigate the computational performance of our extension of both the dilated entropy DGF and the DGE DGF to the scaled extension operator. In particular, we will consider that the problem of computing an NFCCE, which we saw in Section 6.2, can be formulated as a BSPP via the scaled extension operator. Because the constructed polytope is the same for ex ante team-correlated equilibria, extensive form correlated, and extensive form coarse correlated, we restrict our attention to NFCCE and leave the numerical investigation on the other solution concepts for future work. We expect the takeaways to be similar.

Figure 5 shows the results for instantiating the mirror prox algorithm with our two DGFs. We see that, similar to the case of zero-sum Nash equilibrium, the DGE DGF performs much better than the dilated entropy DGF, again likely because of the smaller weights needed in order to make the DGF strongly convex on the correlation-plan polytope.

8. Conclusions and Future Research

We introduced the dilatable global entropy as a distance-generating function for sequential decision-making polytopes such as those encountered in sequential games. We showed that the DGE function leads to better strong-convexity properties than prior DGFs for the sequence-form polytope, and it improves the associated polytope diameter, as well as the convergence rate of FOMs, by a factor of $2^{\mathfrak{D}_Q}$. Experiments confirmed that this leads to a superior notion of distance. We then extended the DGE, as well as the general dilation framework, to the scaled extension operation. We thereby showed how to construct suitable DGFs for the convex polytopes encountered when computing certain correlated equilibria, as well as team equilibria. Based on these extensions, we developed the first algorithm that

Figure 3. (Color online) Performance of the MP Algorithm Instantiated with the Two Entropy DGFs Across Nine Games



Note. The x-axis shows the number of MP iterations, and the y-axis shows the distance to Nash equilibrium.

achieves a $1/T$ convergence rate to the set of various correlated equilibria and ex ante team-coordinated equilibria while requiring only linear time (in the polytope size) for each iteration.

In future research, it would be interesting to investigate whether our new DGFs can be used to achieve numerical performance comparable to that of the currently practically fastest algorithms for game solving—concretely, the state-of-the-art CFR variants (Brown and Sandholm 2019a, Farina et al. 2021b, Xu et al. 2024, Zhang et al. 2024) for Nash equilibrium finding in two-player zero-sum games, which have worse theoretical convergence rates. In particular, we think that stochastic methods could be a promising line of research for this because it is harder to tune the step size in those methods in order to account for the weights previously used in the dilated entropy DGF.

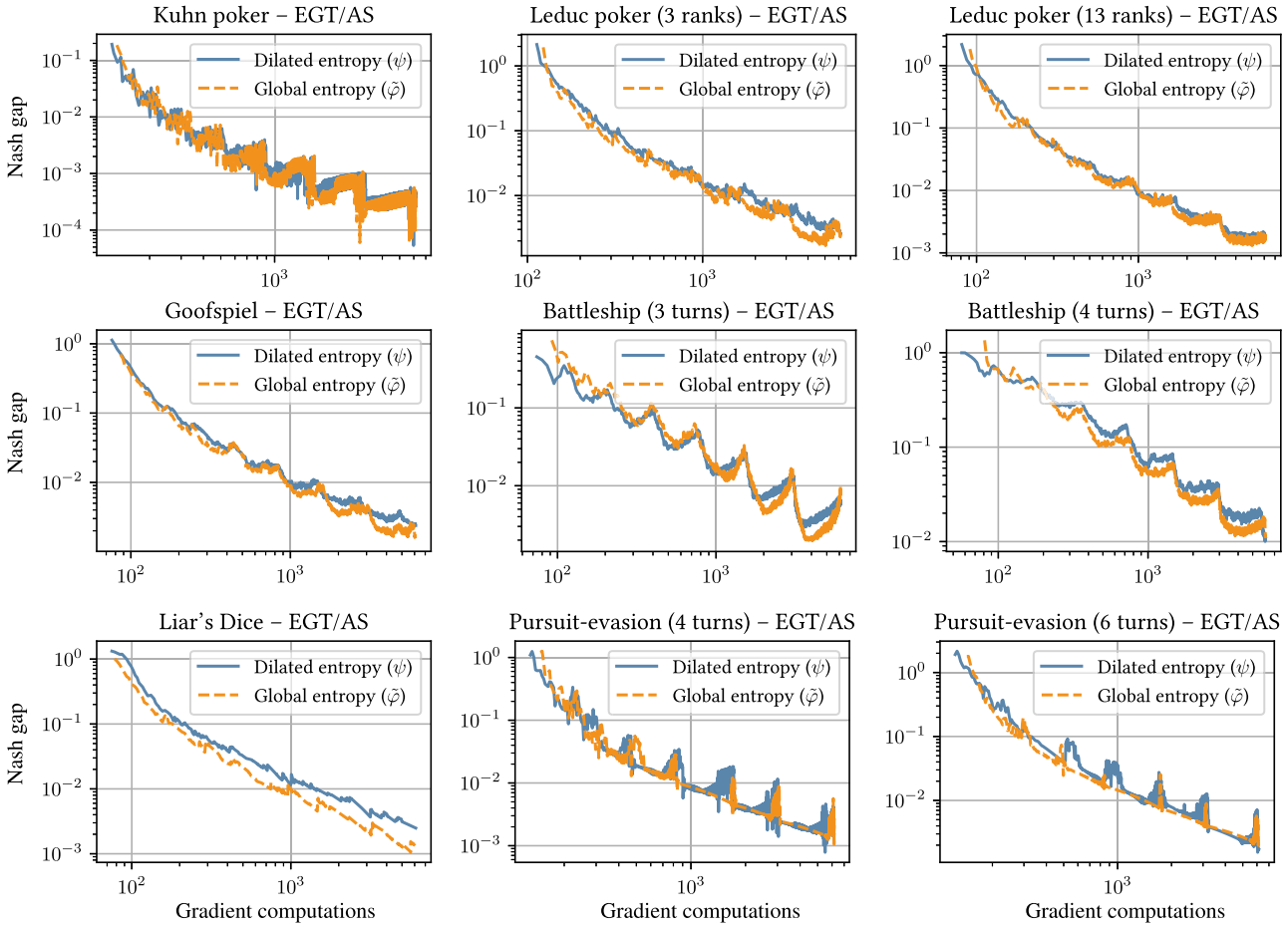
8.1. Work Subsequent to the Original Version of This Paper

Because the original version of our paper was made available in 2021, there has been significant follow-up work on solving EFGs with various notions of dilated entropy regularization. Below, we survey each of these lines of work.

Bai et al. (2022) show that the online mirror descent algorithm setup with a different choice of coefficients than the dilatable global entropy DGF can achieve better dependence on EFG parameters. The analysis rests on showing iterate equivalence with a seemingly very different approach called the *kernelized optimistic multiplicative weights update* (KOMWU) algorithm (Farina et al. 2022b). KOMWU is a kernel-based generalization of the classic multiplicative weights update algorithm from simplex domains to sequence-form polytopes. Similar ideas were used by Fiegel et al. (2023) in the context of bandit algorithms for extensive-form games.

Chakrabarti et al. (2024) showed that our dilatable global entropy regularizer in a block-coordinate framework leads to efficient blockwise updates and strong practical performance in certain games. Moreover, they show that our regularizer combined with restarting leads to $O(2^{-\Omega(T)})$ convergence in some games (hiding game-dependent constants). There have also been recent results on last-iterate convergence in EFGs (Lee et al. 2021, Wei et al. 2021). Most related to our paper, Lee et al. (2021)

Figure 4. (Color online) Performance of the EGT/AS Algorithm Instantiated with the Two Entropy DGFs, as well as Aggressive Step Sizing, μ Balancing, and Initial μ Fitting



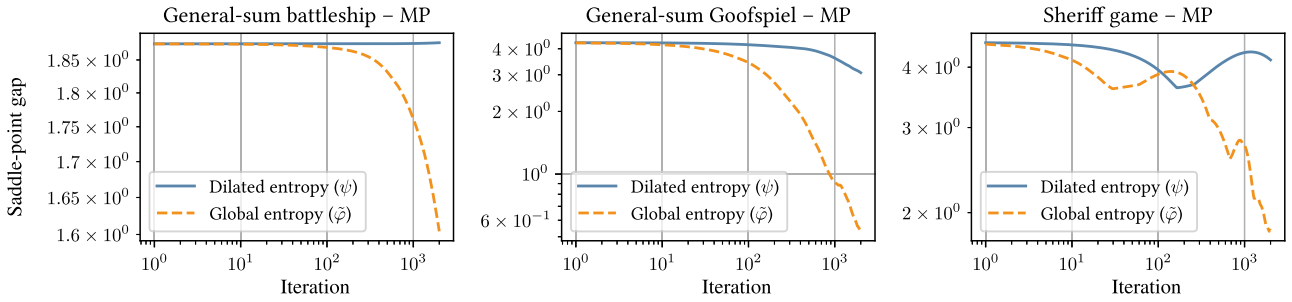
Note. The x-axis shows the number of iterations of EGT with Aggressive Stepsizing (EGT/AS), and the y-axis shows the distance to Nash equilibrium.

show that in two-player zero-sum EFGs with a unique Nash equilibrium, the dilated entropy regularizer achieves a linear-rate convergence, albeit with a dependence on certain game constants that may be large.

In the context of bandit linear optimization, Farina et al. (2021c) used dilated entropy with the weights defined by Kroer et al. (2020). It is likely that the new weights from the present paper would lead to an immediate improvement in the regret bound both in theory and practice.

There has been a series of papers using the dilated entropy function as a regularizer of the game (rather than as a regularizer of the algorithm). This leads to a strongly convex-strongly concave saddle-point problem, which has nice properties such as uniqueness of the regularized saddle point (this is generally *not* a

Figure 5. (Color online) Performance of the MP Algorithm for Finding Normal-Form Coarse Correlated Equilibria in Three General-Sum Games



Note. The x-axis shows the number of iterations of MP, and the y-axis shows the distance to a coarse correlated equilibrium.

Nash equilibrium of the original game) and attractive last-iterate convergence properties, even with simple mirror descent algorithms (Sokota et al. 2023).

In terms of experimental performance in solving two-player zero-sum EFGs, it is well-known that the state-of-the-art CFR-based algorithms DCFR (Brown and Sandholm 2019a), PCFR$^+$ (Farina et al. 2021b), DD-CFR (Xu et al. 2024), HS-DCFR (Zhang et al. 2024), and HS-PCFR$^+$ (Zhang et al. 2024) lead to the fastest practical performance in spite of an inferior $1/\sqrt{T}$ rate of convergence in the worst case, as compared with the $1/T$ rate achieved by, for example, EGT or mirror prox using a dilated DGF. Here, we focus on follow-up experimental work relating to the use of dilated DGFs for EFG solving and, generally, DGF-based methods. Chakrabarti et al. (2024) show that when our regularizer is combined with restarting in a block-coordinate algorithm, it leads to better performance than the classical dilated entropy and dilated Euclidean distance DGFs across several games (though notably, dilated entropy with all weights set to one performs better in Liar's dice variants). That said, a restarting variant of the PCFR$^+$ algorithm still performs best in the experiments of Chakrabarti et al. (2024), despite its worse theoretical convergence guarantees. To our knowledge, restarting with EGT, using all the practical tricks employed in Hoda et al. (2010), the method of Kroer et al. (2018) has not been tried; it is unknown whether this method could be restarted and how it would perform. For the problem of computing correlated equilibria in general-sum games, experimental comparisons between our regularizer and follow-up methods such as the KOMWU algorithm have not been performed. Farina et al. (2022b) compare CFR and CFR$^+$ to KOMWU in terms of regret cumulated in several general-sum games and find that KOMWU cumulates lower regret than the CFR methods, though the break-even point for CFR$^+$ is after several thousand iterations. A thorough investigation of experimental performance across different algorithms and DGFs for computing correlated equilibria is a promising future direction.

Acknowledgments

The authors gratefully acknowledge the insightful comments from the anonymous reviewers at Operations Research and the ACM Conference on Economics and Computation. These comments led to a greatly improved manuscript.

This paper first appeared as an extended abstract at the ACM Conference on Economics and Computation in 2021.

Endnotes

[1] Techniques have been developed for sparsifying the LP in advance of solving it, and those techniques dramatically increase the scalability of the LP-based approach in games with certain structures (Gilpin and Sandholm 2007, Hoda et al. 2010, Zhang and Sandholm 2020, Farina and Sandholm 2022). Even then, however, the state of the art is first-order methods (Farina and Sandholm 2022).

[2] The $1/T$ rate can also be achieved with *predictive* (aka, *optimistic*) variants of *regret-based methods* (Syrgkanis et al. 2015, Farina et al. 2019b). A better rate of $\log(1/T)$ can be achieved, but with a game-dependent multiplicative condition number (Tseng 1995, Gilpin et al. 2012, Wei et al. 2021).

[3] A Bregman divergence need not be symmetric and thus might not be a metric in the technical sense.

[4] In this paper, we let $0\log(0) = 0\log(0/0) = 0$. Because the dilated entropy DGF is a Legendre function, it is guaranteed that all iterates and prox-steps will remain in the relative interior of the optimization domain at all times, thus avoiding the nondifferentiability issue of the entropy function at the boundary of Q.

[5] The triangle-free condition is rather technical, and so we omit its exact definition here, as it is beyond the scope of the paper. The most natural class of games that it captures is the set of EFGs where all chance moves are public, that is, observed by all players.

[6] We recall that we use the convention $0\log(0) = 0\log(0/0) = 0$.

References

Abernethy JD, Rakhlin A (2009) Beating the adaptive bandit with high probability. *COLT 2009 - 22nd Conf. Learn. Theory (Montreal, Quebec)*, 280.

Bai Y, Jin C, Mei S, Song Z, Yu T (2022) Efficient phi-regret minimization in extensive-form games via online mirror descent. *NIPS'22 Proc. 36th Internat. Conf. Neural Inform. Processing Systems* (Curran Associates Inc., Red Hook, NY), 22313–22325.

Beck A, Teboulle M (2003) Mirror descent and nonlinear projected subgradient methods for convex optimization. *Oper. Res. Lett.* 31(3):167–175.

Ben-Tal A, Nemirovski A (2001) *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Application*, vol. 2 (Society for Industrial and Applied Mathematics, Philadelphia).

Ben-Tal A, Nemirovski A (2005) Non-Euclidean restricted memory level method for large-scale convex optimization. *Math. Programming* 102(3):407–456.

Bowling M, Burch N, Johanson M, Tammelin O (2015) Heads-up limit hold'em poker is solved. *Science* 347(6218):145–149.

Brown N, Sandholm T (2017) Superhuman AI for heads-up no-limit poker: Libratus beats top professionals. *Science* 359(6374):418–424.

Brown N, Sandholm T (2019a) Solving imperfect-information games via discounted regret minimization. *Proc. AAAI Conf. Artificial Intelligence*, vol. 33 (Association for the Advancement of Artificial Intelligence, Washington, DC), 1829–1836.

Brown N, Sandholm T (2019b) Superhuman AI for multiplayer poker. *Science* 365(6456):885–890.

Celli A, Gatti N (2018) Computational results for extensive-form adversarial team games. *Proc. AAAI Conf. Artificial Intelligence*, vol. 32 (Association for the Advancement of Artificial Intelligence, Washington, DC), 965–972.

Chakrabarti D, Diakonikolas J, Kroer C (2024) Block-coordinate methods and restarting for solving extensive-form games. *NIPS'23 Proc. 37th Internat. Conf. Neural Inform. Processing Systems*, vol. 36 (Curran Associates Inc., Red Hook, NY), 10692–10720.

Condat L (2016) Fast projection onto the simplex and the ℓ_1 ball. *Math. Programming* 158(1):575–585.

Farina G, Sandholm T (2020) Polynomial-time computation of optimal correlated equilibria in two-player extensive-form games with public chance moves and beyond. *NIPS '20 Proc. 34th Internat. Conf. Neural Inform. Processing Systems* (Curran Associates Inc., Red Hook, NY), 19609–19619.

Farina G, Sandholm T (2022) Fast payoff matrix sparsification techniques for structured extensive-form games. *Proc. 36th AAAI*

Conf. Artificial Intelligence (Association for the Advancement of Artificial Intelligence, Washington, DC), 4999–5007.

Farina G, Bianchi T, Sandholm T (2020a) Coarse correlation in extensive-form games. *Proc. AAAI Conf. Artificial Intelligence*, vol. 34 (Association for the Advancement of Artificial Intelligence, Washington, DC), 1934–1941.

Farina G, Kroer C, Sandholm T (2019a) Online convex optimization for sequential decision processes and extensive-form games. *Proc. AAAI Conf. Artificial Intelligence*, vol. 33 (Association for the Advancement of Artificial Intelligence, Washington, DC), 1917–1925.

Farina G, Kroer C, Sandholm T (2019b) Optimistic regret minimization for extensive-form games via dilated distance-generating functions. *Proc. 33rd Internat. Conf. Neural Inform. Processing Systems (NeurIPS)* (Curran Associates Inc., Red Hook, NY), 5222–5232.

Farina G, Kroer C, Sandholm T (2020b) Stochastic regret minimization in extensive-form games. *Proc. 37th Internat. Conf. Machine Learn.*, Proceedings of Machine Learning Research, vol. 119 (PMLR, New York), 3018–3028.

Farina G, Kroer C, Sandholm T (2021b) Faster game solving via predictive Blackwell approachability: Connecting regret matching and mirror descent. *Proc. AAAI Conf. Artificial Intelligence*, vol. 35 (Association for the Advancement of Artificial Intelligence, Washington, DC), 5363–5371.

Farina G, Schmucker R, Sandholm T (2021c) Bandit linear optimization for sequential decision making and extensive-form games. *Proc. AAAI Conf. Artificial Intelligence*, vol. 35 (Association for the Advancement of Artificial Intelligence, Washington, DC), 5372–5380.

Farina G, Celli A, Gatti N, Sandholm T (2021a) Connecting optimal ex-ante collusion in teams to extensive-form correlation: Faster algorithms and positive complexity results. *Proc. 38th Internat. Conf. Machine Learn.*, Proceedings of Machine Learning Research, vol. 139 (PMLR, New York), 3164–3173.

Farina G, Lee C-W, Luo H, Kroer C (2022b) Kernelized multiplicative weights for 0/1-polyhedral games: Bridging the gap between learning in extensive-form and normal-form games. *Internat. Conf. Machine Learning, ICML 2022*, Proceedings of Machine Learning Research, vol. 162 (PMLR, New York), 6337–6357.

Farina G, Ling CK, Fang F, Sandholm T (2019c) Correlation in extensive-form games: Saddle-point formulation and benchmarks. *Proc. 33rd Internat. Conf. Neural Inform. Processing Systems (NeurIPS)* (Curran Associates Inc., Red Hook, NY), 9233–9243.

Farina G, Ling CK, Fang F, Sandholm T (2019d) Efficient regret minimization algorithm for extensive-form correlated equilibrium. *Proc. 33rd Internat. Conf. Neural Inform. Processing Systems (NeurIPS)* (Curran Associates Inc., Red Hook, NY), 5186–5196.

Farina G, Anagnostides I, Lee C-W, Luo H, Kroer C, Sandholm T (2022a) Near-optimal no-regret learning for general convex games. *NIPS'22 Proc. 36th Internat. Conf. Neural Inform. Processing Systems (NeurIPS)* (Curran Associates Inc., Red Hook, NY), 39076–39089.

Fiegel C, Ménard P, Kozuno T, Munos R, Perchet V, Valko M (2023) Adapting to game trees in zero-sum imperfect information games. Krause A, Brunskill E, Cho K, Engelhardt B, Sabato S, Scarlett J, eds. *Internat. Conf. Machine Learn., ICML 2023*, Proceedings of Machine Learning Research, vol. 202 (PMLR, New York), 10093–10135.

Gibson R, Lanctot M, Burch N, Szafron D, Bowling M (2012) Generalized sampling and variance in counterfactual regret minimization. *Proc. AAAI Conf. Artificial Intelligence*, vol. 26 (Association for the Advancement of Artificial Intelligence, Washington, DC), 1355–1361.

Gilpin A, Sandholm T (2007) Lossless abstraction of imperfect information games. *J. ACM* 54(5):25.

Gilpin A, Peña J, Sandholm T (2012) First-order algorithm with $\mathcal{O}(\ln(1/\epsilon))$ convergence for ϵ-equilibrium in two-person zero-sum games. *Math. Programming* 133(1–2):279–298.

Held M, Wolfe P, Crowder HP (1974) Validation of subgradient optimization. *Math. Programming* 6(1):62–88.

Hoda S, Gilpin A, Peña J, Sandholm T (2010) Smoothing techniques for computing nash equilibria of sequential games. *Math. Oper. Res.* 35(2):494–512.

Koller D, Megiddo N, von Stengel B (1996) Efficient computation of equilibria for extensive two-person games. *Games Econom. Behav.* 14(2):247–259.

Kroer C, Farina G, Sandholm T (2018) Solving large sequential games with the excessive gap technique. *NIPS'18 Proc. 32nd Internat. Conf. Neural Inform. Processing Systems* (Curran Associates Inc., Red Hook, NY), 872–882.

Kroer C, Waugh K, Il Inç-Karzan FK, Sandholm T (2015) Faster first-order methods for extensive-form game solving. *EC'15 Proc. 16th ACM Conf. Econom. Comput.* (Association for Computing Machinery, New York), 817–834.

Kroer C, Waugh K, Il Inç-Karzan FK, Sandholm T (2017) Theoretical and practical advances on smoothing for extensive-form games. *EC'17 Proc. 2017 ACM Conf. Econom. Comput.* (Association for Computing Machinery, New York), 693.

Kroer C, Waugh K, Il Inç-Karzan FK, Sandholm T (2020) Faster algorithms for extensive-form game solving via improved smoothing functions. *Math.* Programming 179:385–417.

Lanctot M, Waugh K, Zinkevich M, Bowling M (2009) Monte Carlo sampling for regret minimization in extensive games. *NIPS'09 Proc. 23rd Internat. Conf. Neural Inform. Processing Systems* (Curran Associates Inc., Red Hook, NY), 1078–1086.

Lee C-W, Kroer C, Luo H (2021) Last-iterate convergence in extensive-form games. *NIPS'21 Proc. 35th Internat. Conf. Neural Inform. Processing Systems* (Curran Associates Inc., Red Hook, NY), 14293–14305.

Moravčík M, Schmid M, Burch N, Lisý V, Morrill D, Bard N, Davis T, Waugh K, Johanson M, Bowling M (2017) DeepStack: Expert-level artificial intelligence in heads-up no-limit poker. *Science* 356(6337):508–513.

Moulin H, Vial J-P (1978) Strategically zero-sum games: The class of games whose completely mixed equilibria cannot be improved upon. *Internat. J. Game Theory* 7(3–4):201–221.

Nemirovski A (2004) Prox-method with rate of convergence O(1/t) for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems. *SIAM J. Optim.* 15(1):229–251.

Nesterov Y (2005a) Excessive gap technique in nonsmooth convex minimization. *SIAM J. Optim.* 16(1):235–249.

Nesterov Y (2005b) Smooth minimization of non-smooth functions. *Math. Programming* 103:127–152.

Romanovskii I (1962) Reduction of a game with complete memory to a matrix game. *Soviet Math.* 3:678–681.

Schmid M, Burch N, Lanctot M, Moravcik M, Kadlec R, Bowling M (2019) Variance reduction in Monte Carlo counterfactual regret minimization (VR-MCCFR) for extensive form games using baselines. *Proc. AAAI Conf. Artificial Intelligence*, vol. 33 (Association for the Advancement of Artificial Intelligence, Washington, DC), 2157–2164.

Sokota S, D'Orazio R, Kolter JZ, Loizou N, Lanctot M, Mitliagkas I, Brown N, Kroer C (2023) A unified approach to reinforcement learning, quantal response equilibria, and two-player zero-sum games. *Eleventh Internat. Conf. Learn. Representations, ICLR 2023 (Kigali, Rwanda)*.

Syrgkanis V, Agarwal A, Luo H, Schapire RE (2015) Fast convergence of regularized learning in games. *NIPS'15 Proc. 29th Internat. Conf. Neural Inform. Processing Systems* (Curran Associates Inc., Red Hook, NY), 2989–2997.

Tseng P (1995) On linear convergence of iterative methods for the variational inequality problem. *J. Comput. Appl. Math.* 60(1–2):237–252.

von Stengel B (1996) Efficient computation of behavior strategies. *Games Econom. Behav.* 14(2):220–246.

von Stengel B, Forges F (2008) Extensive-form correlated equilibrium: Definition and computational complexity. *Math. Oper. Res.* 33(4):1002–1022.

Wei C-Y, Wei Lee C, Zhang M, Luo H (2021) Linear last-iterate convergence in constrained saddle-point optimization. *9th Internat. Conf. Learn. Representations, ICLR 2021 (Vienna, Austria)*.

Xu H, Li K, Fu H, Fu Q, Xing J, Cheng J (2024) Dynamic discounted counterfactual regret minimization. *Twelfth Internat. Conf. Learn. Representations, ICLR 2024 (Vienna, Austria).*

Zhang B, Sandholm T (2020) Sparsified linear programming for zero-sum equilibrium finding. *ICML'20 Proc. 37th Internat. Conf. Machine Learning*, Proceedings of Machine Learning Research, vol. 119 (PMLR, New York), 11256–11267.

Zhang BH, Farina G, Sandholm T (2023) Team belief DAG: Generalizing the sequence form to team games for fast computation of correlated team max-min equilibria via regret minimization. *Internat. Conf. Machine Learn., ICML 2023*, Proceedings of Machine Learning Research, vol. 202 (PMLR, New York), 40996–41018.

Zhang N, McAleer S, Sandholm T (2024) Faster game solving via hyperparameter schedules. Preprint, submitted April 13, https://doi.org/10.48550/arXiv.2404.09097.

Zhang BH, Farina G, Celli A, Sandholm T (2022) Optimal correlated equilibria in general-sum extensive-form games: Fixed-parameter algorithms, hardness, and two-sided column-generation. *EC'22 Proc. 23rd ACM Conf. Econom. Comput.* (Association for Computing Machinery, New York), 1119–1120.

Zinkevich M, Bowling M, Johanson M, Piccione C (2007) Regret minimization in games with incomplete information. *NIPS'07 Proc. 21st Internat. Conf. Neural Inform. Processing Systems* (Curran Associates Inc., Red Hook, NY), 1729–1736.

Gabriele Farina is an assistant professor in the Department of Electrical Engineering and Computer Science at the Massachusetts Institute of Technology (MIT), and a member of MIT's Laboratory for Information and Decision Systems. His research combines techniques and notions of strategic behavior from game theory together with modern tools from machine learning, optimization, and statistics to construct state-of-the-art methods to compute optimal strategies for multiagent interactions. He received his PhD in computer science from Carnegie Mellon University, and his work has been recognized with several awards, including a Best Paper Award at NeurIPS'20 and an Outstanding Paper Honorable Mention at ICLR'23. His dissertation was recognized with the ACM SIGecom Doctoral Dissertation Award and one of the two ACM Dissertation Award Honorable Mentions, among others. He is the recipient of an NSF CAREER award and an AI2050 fellow.

Christian Kroer is an associate professor of industrial engineering and operations research at Columbia University, as well as a member of the Data Science Institute at Columbia. His research interests are at the intersection of operations research, economics, and computation, with a focus on how optimization and artificial intelligence methods enable large-scale economic solution concepts. He obtained his PhD in computer science from Carnegie Mellon University and spent a year as a postdoc with the Economics and Computation team at Facebook Research. He is the recipient of an Office of Naval Research Young Investigator award and a National Science Foundation CAREER award.

Tuomas Sandholm is Angel Jordan University Professor of Computer Science at Carnegie Mellon University and a serial entrepreneur. His research focuses on the convergence of artificial intelligence, economics, and operations research. He has published over 500 peer-reviewed papers and holds 29 patents, and his h-index is 97. In addition to his main appointment in the Computer Science Department, he has appointments in the machine learning department, PhD Program in Algorithms, Combinatorics, and Optimization (ACO), and CMU/UPitt Joint PhD Program in Computational Biology. He is a fellow of INFORMS, The Association for Computing Machinery, The Association for the Advancement of Artificial Intelligence, and The American Association for the Advancement of Science.