2016 IEEE 57th Annual Symposium on Foundations of Computer Science

On the Communication Complexity of Approximate Fixed Points

Tim Roughgarden
Department of Computer Science
Stanford University
Palo Alto, CA
tim@cs.stanford.edu

Omri Weinstein
Department of Computer Science
New York University
New York, NY
omri@cs.columbia.edu

***Abstract*—We study the two-party communication complexity of finding an approximate Brouwer fixed point of a composition of two Lipschitz functions $g \circ f : [0,1]^n \to [0,1]^n$, where Alice holds f and Bob holds g. We prove an exponential (in n) lower bound on the deterministic communication complexity of this problem. Our technical approach is to adapt the Raz-McKenzie simulation theorem (FOCS 1999) into geometric settings, thereby "smoothly lifting" the deterministic *query* lower bound for finding an approximate fixed point (Hirsch, Papadimitriou and Vavasis, Complexity 1989) from the oracle model to the two-party model. Our results also suggest an approach to the well-known open problem of proving strong lower bounds on the communication complexity of computing approximate Nash equilibria. Specifically, we show that a slightly "smoother" version of our fixed-point computation lower bound (by an absolute constant factor) would imply that:**

- **The deterministic two-party communication complexity of finding an $\epsilon = \Omega(1/\log^2 N)$-approximate Nash equilibrium in an $N \times N$ bimatrix game (where each player knows only his own payoff matrix) is at least N^γ for some constant $\gamma > 0$. (In contrast, the *nondeterministic* communication complexity of this problem is only $O(\log^6 N)$).**
- **The deterministic (Number-In-Hand) multiparty communication complexity of finding an $\epsilon = \Omega(1)$-Nash equilibrium in a k-player constant-action game is at least $2^{\Omega(k/\log k)}$ (while the nondeterministic communication complexity is only $O(k)$).**

I. Introduction

Brouwer's fixed-point theorem states that every continuous function h from a closed convex set C to itself has at least one fixed point — that $h(x) = x$ for some $x \in C$. This result, and generalizations thereof such as Kakutani's fixed-point theorem and the Borsuk-Ulam theorem, have countless applications in mathematics, logic and economics ([Bor85], [Mat07]). To give just one example, all known proofs of the existence of Nash equilibria in general finite games rely on such fixed-point theorems.

Due to its fundamental nature, the problem of computing (approximate) Brouwer fixed points has been studied for half a century, beginning with Scarf [Sca67], who adapted ideas of Lemke and Howson [LH64] to obtain an (exponential-time) algorithm for the problem. Previous work provides a fairly sharp understanding of the complexity of finding approximate fixed-points in two computational models: Hirsch, Papadimitriou, and Vavasis [HPV89] pioneered the study of the *query complexity* of the problem in the "black-box" oracle model, where an algorithm can only interact with the function h by (adaptively) querying it at different points in the domain (i.e., no explicit description is provided). The main result of [HPV89], which is a tour de force, is that every deterministic algorithm for computing an ε-approximate fixed point of a function h mapping the n-dimensional cube to itself has worst-case query complexity $(\frac{1}{\varepsilon})^{\Theta(n)}$, even when the function h has a Lipschitz constant arbitrarily close to 1. This lower bound was subsequently extended to the *randomized* query model for two-dimensional manifolds ([LNNW95]), and only recently by Babichenko [Bab14] for any dimension $d \geq 2$. In parallel to this line of work, Papadimitriou [Pap94] considered the *computational* complexity of computing approximate Brouwer fixed points for explicitly described functions,[1] i.e., in a "white-box" model, and proved that the problem is complete for the complexity class **PPAD** in 3 or more dimensions [Pap94]. The two-dimensional version of the problem also turned out to be **PPAD**-complete [CD09].

This paper initiates the study of the two-party (and multiparty) *communication complexity* of computing approximate Brouwer fixed points. That is, we study the problem in a "grey-box" model of computation. We consider the natural version of the problem in which Alice's input is an explicitly described function $f : C_1 \to C_2$, Bob's input is an explicitly described function $g : C_2 \to C_1$, and the task is to compute an approximate fixed point of the composed function $g \circ f : C_1 \to C_1$. Our lower bounds are for the case where C_1 and C_2 are discretized hypercubes (of possibly different dimensions), with every coordinate of

[1]For example, one can describe a function on a finite set of points, and use some canonical interpolation to define a continuous real-valued function.

0272-5428/16 $31.00 © 2016 IEEE
DOI 10.1109/FOCS.2016.32

every point a multiple of some (small) constant α. The goal is to compute some $x \in C_1$ with $\|h(x) - x\|_\infty \leq \varepsilon$ (if one exists).[2] We will generally think of ε as a small constant (e.g., 10^{-3}) and α as a much smaller constant (e.g., 10^{-6}).

The communication complexity of this problem varies with the approximation parameter ε and also with the geometry (amount of structure) imposed on the input functions f and g. We interpolate between easy and hard versions of the problem through *Lipschitz constraints* on the functions f and g. Specifically, we assume that Alice's function f is λ_1-Lipschitz (meaning $\|f(x) - f(y)\|_\infty \leq \lambda\|x - y\|_\infty$ for all $x, y \in C_1$) and Bob's function g is λ_2-Lipschitz. If we only constrain $\lambda_1, \lambda_2 = O(\frac{\varepsilon}{\alpha})$ and hence $\lambda_1\lambda_2 = \Theta(\frac{\varepsilon^2}{\alpha^2})$, then it is easy to prove strong lower bounds on the problem (e.g., via a reduction from set disjointness). On the other hand, if $\lambda_1\lambda_2 < 1$ (i.e., when $g \circ f$ is a "contraction" over the domain C_1), Alice and Bob can easily find an ε-approximate fixed point in C_1 whenever it exists[3] by iteratively evaluating the function using only $O(\log 1/\varepsilon)$ many rounds of communication. So the problem transitions from easy to hard as $\lambda_1\lambda_2$ varies from small to large — where does the transition occur?

Our main result is an exponential (in the dimension) lower bound on the deterministic communication complexity of computing an ε-approximate fixed point, even when $\lambda_1\lambda_2$ is as small as $43\frac{\varepsilon}{\alpha}$ (i.e., the values of $g \circ f$ on neighboring α-grid points differ by at most 43ε). Put differently, our lower bound applies to the regime where the approximation parameter ε is independent of the "discretization parameter" α (and in particular when $\lim_{\alpha \to 0} \frac{\varepsilon}{\alpha} = \infty$). [4]

An important feature of our lower bound is that it holds even under the promise that the composed function $h := g \circ f$ is $O(1)$-Lipschitz (irrespective of α !), even though the *marginal* Lipschitz constants of f and g in our lower bound only satisfy the weaker constraint $\lambda_1\lambda_2 = \Theta(1/\alpha)$ as stated above. Note that the former promise immediately guarantees the existence of an approximate fixed-point of h (for a small enough constant α), as h can always be extended to a *continuous* function on the solid cube. Informally, this feature of our construction means that the "bottleneck" of our lower bound is the two-party *decomposition* (aka factoring) of h, while a "smoother" decomposition (which still yields a hard communication problem) is plausible as h is itself very smooth (see Section VI for further discussion). In contrast, any "disjointness-based" reduction to the composed fixed-point problem would inherently yield a *discontinuous* function (due to the local nature of such a reduction when translating $x \in \{0, 1\}^n$ to displacements of a function). Since we always have $\lambda_g\lambda_f \geq \lambda_{g \circ f}$ (by transitivity), it is not clear how such a reduction can carry over to the total version of the problem (in which a fixed-point always exists). Acknowledging this qualitative difference is imperative to understand our result and its potential future implications.

Since our lower bound trivially implies an exponential lower bound on the deterministic *query* complexity of computing an ε-approximate fixed point for an $O(1)$-Lipschitz functions h — the main result in [HPV89] modulo polynomial factors — we do not expect a simple proof of this result (see our proof outline in Section II-B).

In addition to its basic nature, a second motivation for studying the problem of computing fixed points is its tight connections to other problems, such as computing a Nash equilibrium in strategic games. For both query and computational complexity, lower bounds for the former problem were crucial prerequisites to lower bounds for the latter problem [Bab14], [CCT15], [CDT09], [DGP09].[5] What about *distributed* computation of (approximate) Nash equilibria in the (realistic) scenario where each player only knows his own payoff matrix? This question was advocated before due to its implications on the rate of convergence of uncoupled market dynamics [HMC02], [HM10]. We stress that lower bounds in the communication complexity model isolate the *information-theoretic bottleneck* faced by such dynamics, as opposed to, e.g., conditional lower bounds based on "bounded-rationality"-type assumptions (see e.g., [Sha64] and Section IV in [HMC02]).

Conitzer and Sandholm [CS04] were the first to study the communication complexity of equilibria. In $N \times N$ bimatrix games, they proved an $\Omega(N^2)$ communication complexity lower bound for the problem of deciding whether or not a game has a *pure* Nash equilibrium (via a reduction from set disjointness). Hart and Mansour [HM10] focused on the search problem of finding a mixed Nash equilibrium in an n-player game with binary strategy sets and proved that the communication complexity of finding an *exact* Nash equilibrium is 2^n (note that the input size of each player is 2^n, as there

[2]A protocol is allowed to behave arbitrarily on inputs that have no approximate fixed points.

[3]When C_1 is an (equally spaced) grid over $[0, 1]^n$, it is easy to see that any "contracting" function must in fact be constant (and in particular must have an exact fixed point), so the argument for grids is trivial in this case.

[4]We assume α is equal to the minimum non-zero displacement of $g \circ f$, as otherwise, a finer grid can be chosen without modifying the function nor its fixed-points.

[5]For query complexity, there is an exponential (in the number of players) lower bound even for computing an ε-ANE with $\varepsilon = \Omega(1)$ ([Bab14], [CCT15]). The *computational complexity* of computing an ε-ANE in two-player games for $\varepsilon = \Omega(1)$ was only very recently settled ([Rub16], [BPR16]), where it was shown that the quasi-polynomial-time algorithm for this problem of [LMM03] is essentially optimal under a certain ("ETH"-type) conjecture for PPAD.

are 2^n joint strategy profiles). It is noteworthy that both of these lower bounds hold also for the *nondeterministic* communication complexity of the problem.

Almost nothing is known about the communication complexity of computing ε-*approximate* Nash equilibria (ε-ANE) for small positive values of ε. This is not a coincidence: In sharp contrast to the problems above, the nondeterministic communication complexity of this problem is only logarithmic in the size of the game description (and quadratic in $\frac{1}{\varepsilon}$) [LMM03]. Moreover, for ε sufficiently large, the problem turns out to be easy – Goldberg and Pastink [GP13] and subsequent improvements due to Czumaj et. al [CDF+15] show that finding an $\varepsilon = 0.382$-ANE in a bimatrix game can be done using only $poly \log(N)$ deterministic communication, suggesting that the problem is subtle (as any lower bound has to inherently rely on ε being sufficiently small). [GP13] proved strong lower bounds only for the *one-way* communication complexity of the problem, but there are no known non-trivial lower bounds in the unbounded-round communication model for any $\varepsilon > 0$ (for both the two-payer and the multi-player settings).

We propose a path to proving strong lower bounds on the communication complexity of computing ε-approximate Nash equilibria. Specifically, in both the bimatrix and multi-player cases, we show how to use a protocol for computing approximate Nash equilibria to compute ε-approximate fixed points for input functions f and g with Lipschitz constants that satisfy $\lambda_1\lambda_2 \leq \frac{1}{2}\frac{\varepsilon}{\alpha}$ (this reduction holds for both deterministic and randomized communication). Thus, a constant-factor (namely, 86) improvement in the Lipschitz constraint in our main result immediately implies strong deterministic communication complexity lower bounds for computing approximate Nash equilibria. As explained in [HM10], such a lower bound would rule out the fast convergence of any form of deterministic uncoupled dynamics that converges even to an *approximately* stable market state.

II. Overview of Results

Let $\mathbf{AFPC}_{\alpha,(n,\lambda_1),(m,\lambda_2),\varepsilon}$ denote the two-party search problem of finding an ε-fixed-point of $g \circ f$, where Alice holds (the truth table of) a λ_1-Lipschitz function $f : G_{\alpha,n} \mapsto G_{\alpha,m}$ and Bob holds a λ_2-Lipschitz function $g : G_{\alpha,m} \mapsto G_{\alpha,n}$. ($G_{\alpha,n}$ denotes the α-grid of the n-dimensional solid cube $[0,1]^n$, see the formal definition of the problem in Section V).

Our first and main result asserts that every deterministic communication protocol that finds a $(\lambda_1\lambda_2\alpha/43)$-fixed-point of the composed function $g \circ f$ requires exponential communication in the dimension n (with $m = O(n)$).

Theorem II.1 (Deterministic communication lower bound for $\mathbf{AFPC}$)**.** *There are universal constants* $\alpha \in (0,1)$*,* $\lambda_1, \lambda_2 > 1$ *such that for every* $n \geq 2$ *and* $m = O_\alpha(n)$*,*

$$\mathsf{D}^{\mathsf{CC}}\left(\mathbf{AFPC}_{\alpha,(n,\lambda_1),(m,\lambda_2),\frac{\lambda_1\lambda_2\alpha}{43}}\right) \geq 2^{\Omega_\alpha(n)}.$$

We stress that that parameters in the result above are such that $\lambda_1\lambda_2 = \Theta(1/\alpha)$, that is, the approximation parameter $\varepsilon = (\lambda_1\lambda_2\alpha/43)$ for which we prove the lower bound is an absolute constant *independent* of the grid size (i.e., the "discretization parameter") α, and in particular, ε can be much larger than α.

Our second contribution is a reduction from $\mathbf{AFPC}$ to the problem of computing an approximate Nash equilibrium ($\mathbf{ANE}$). This result shows that any communication lower bound (deterministic or randomized) on finding a $(2\lambda_1\lambda_2\alpha)$-fixed-point of $g \circ f$ translates to two different lower bounds: (i) on the two-party communication complexity of finding an $\Omega(1/\log^2 K)$-ANE in a 2-player bimatrix game with $K = \exp(n)$ actions ; (ii) on the k-party (Number-In-Hand) communication complexity of finding an $\Omega(1)$-ANE in a k-player constant-action game.

Theorem II.2 (From approximate fixed points to approximate Nash, informal)**.** *For every* $m \geq n \in \mathbb{N}$*, any constants* λ_1, λ_2*,* $\alpha \in (0,1)$*, and any error parameter* $\rho \geq 0$*:*

- *(Two-player games) Setting* $K := (1/\alpha)^m$*,*
$$\mathsf{R}^{\mathsf{CC}}{}_{\rho}\left(\mathbf{ANE}_{K,\Omega(1/\log^2 K)}\right) \geq \mathsf{R}^{\mathsf{CC}}{}_{\rho}\left(\mathbf{AFPC}_{\alpha,(n,\lambda_1),(m,\lambda_2),2\lambda_1\lambda_2\alpha}\right).$$
- *(k-player games) Setting* $k := O_\alpha(m \log m)$*,*
$$\mathsf{R}^{\mathsf{CC}}{}_{\rho}\left(k\text{-}\mathbf{ANE}_{1/\alpha,3\alpha^3/16}\right) \geq \mathsf{R}^{\mathsf{CC}}{}_{\rho}\left(\mathbf{AFPC}_{\alpha,(n,\lambda_1),(m,\lambda_2),2\lambda_1\lambda_2\alpha}\right).$$

Theorem II.2 implies that a slightly stronger version of Theorem II.1 (where the approximation parameter is larger only by an absolute constant factor) would imply near-optimal deterministic communication lower bounds for finding approximate Nash equilibria in both two-player and k-player games. In turn, this would rule out any efficient distributed dynamics that converges even to an *approximately* stable state (see Proposition 6.4 in [RW16] for the formal statement, and a more elaborate discussion on this direction in Section VI below).

A. Streamlined Overview of Proofs and Techniques

"Lifting": Communication Lower Bounds from Query Lower Bounds. : To prove Theorem II.1, we follow an approach that converts lower bounds in the weaker (and simpler-to-understand) query complexity world ([BdW02]) into two-party lower bounds in the communication complexity world (e.g., [NW95],

[BdW02], [GLM+15], [RM99], [GPW15]). This approach is based on a technique known as "lifting," where the inputs to the (query) problem are distributed in some carefully chosen manner (using a 2-party "gadget") between Alice and Bob, who are then required to solve the resulting distributed search problem.

More formally, let $S : \Sigma^N \longrightarrow \Sigma$ be some search problem (sometimes called the "outer function"). The *g-lift* of S is the two-party communication problem defined by

$$S \circ g^N(\mathbf{x}, \mathbf{y}) := S(g(x_1, y_1), \ldots, g(x_N, y_N)),$$

where the gadget $g : \mathcal{X} \times \mathcal{Y} \longrightarrow \Sigma$ is typically some "small" two-party function. Clearly, the communication complexity of solving $S \circ g^N$ is at most $\log(\min\{|\mathcal{X}|, |\mathcal{Y}|\})\cdot$(query complexity($S$)), since Alice and Bob can always simulate any decision tree for S by sequentially having the player with the shorter input send his corresponding coordinate to the other, who then evaluates the query. Proving the other direction, namely, that such communication protocols are essentially *optimal*, is a highly nontrivial result, commonly referred to as a *simulation theorem* (e.g., [RM99], [GPW15], [GLM+15], [GP14]). The gadget g plays a crucial role in such results, as it ensures Alice and Bob cannot take "short-cuts" by avoiding queries made by the decision tree.[6] Thus the gadget g must be a sufficiently "hard" function to rule out such manipulations (for a more elaborate discussion, see Section 5.2 in the full version of this paper [RW16]). We remark that simulation theorems have recently led to breakthrough results in complexity theory, including the resolution of the long-standing "Clique vs. Independent Set" problem [Göö15], [GPW15], to separation theorems between various deterministic and non-deterministic communication measures [GPW15], [ABB+15], and, earlier, to the separation of the monotone circuit hierarchy [RM99].

The most relevant result to our problem is the simulation theorem of Raz and McKenzie ([RM99]) and its recent generalization due to Goos, Pitassi and Watson ([GPW15]), who showed that, for any search problem $S : \Sigma^N \mapsto \Sigma$, if the input $z = (z_1, \ldots, z_N)$ to S is "lifted" using the *index gadget*

$$\mathsf{IND}(x_i, y_i) := y_i[x_i]$$

(i.e., Alice's input is a set of indices $\mathbf{x} = \{x_i\}_{i=1}^N \in [k]^N$, Bob's input is a set of vectors $\mathbf{y} = \{y_i\}_{i=1}^N \in (\Sigma^k)^N$ for $k = \mathrm{poly}(N)$, such that $y_i[x_i] = z_i$ for every $i \in [N]$), then the "lifted" communication problem remains as hard as the corresponding query problem:

Theorem II.3 ([RM99], [GPW15], informal)**.** *For any search problem S, the deterministic communication complexity of the two-party problem $S \circ \mathsf{IND}^N(\mathbf{x}, \mathbf{y}) := S(y_1[x_1], \ldots, y_N[x_N])$ is at least $\Omega(\log k)$ times the deterministic query complexity of S.*

In the next subsection, we explain the relevance of this theorem to the distributed approximate fixed-point problem ($\mathbf{AFPC}$), and provide a streamlined overview of the proofs of our main results (Theorem II.2 and Theorem II.1).

B. A High-Level Proof Overview of Theorem II.1

The approximate fixed-point problem that we study ($\mathbf{AFPC}$) has a "geometric" aspect to it, in that both of the input functions are required to be $O(1)$-Lipschitz.[7] The Lipschitz condition implies that if, for example, Alice sends Bob a value $f(x)$, then Bob automatically learns information about the value of f on inputs close to x. Dealing with this geometric aspect of the problem is the most challenging and subtle aspect of the proof.

As mentioned above, the key step of the proof is showing that the deterministic communication complexity of $\mathbf{AFPC}$ is bounded from below by the deterministic *query complexity* of the (promise) search problem of finding an approximate fixed point of a λ-Lipschitz function $h : [0,1]^n \mapsto [0,1]^n$ (we denote this problem by $\mathbf{AFP}$). Fortunately, the query complexity of this problem was previously studied by Hirsch, Papadimitriou and Vavasis [HPV89], who showed (using a highly nontrivial geometric construction, see Section 5.1 in [RW16] and an illustration in Figure 1 below) that any (deterministic) decision tree solving this problem requires $2^{\Omega_\lambda(n)}$ queries, for any Lipschitz constant $\lambda > 1$. (This lower bound was recently generalized to the randomized query model by Babichenko [Bab14]).

A natural approach at this point is to try and use simulation theorems to "lift" the aforementioned lower bound from the query setup to the communication setup. Alas, as discussed above, simulation theorems rely on a carefully chosen gadget g, and thus the "lifted" communication problem $S \circ g^N$ typically corresponds to some contrived two-party problem, even when S is a natural problem. Fortunately, the lifting gadget in the Raz-McKenzie simulation theorem is (almost) exactly what we were looking for: Our simple but central observation is that, letting S denote the search problem of finding an approximate fixed point of a (discrete)

[6]For example, if S is the AND function $\bigwedge_{i=1}^N z_i$ and g is chosen as an AND-gadget itself, i.e., $g(x_i, y_i) = x_i \wedge y_i$, then it is easy to see that the deterministic query complexity of S is N, but $S \circ g^N = \bigwedge_i (x_i \wedge y_i) \equiv (\bigwedge_i x_i) \wedge (\bigwedge_i y_i)$ and therefore the communication complexity of $S \circ g^N$ is 0!

[7]The problem is not interesting in the absence of the Lipschitz requirement. In this case, a simple reduction e.g. from set-disjointness can trivially establish the exponential lower bound, even in the *randomized* setting.

function $h : [0,1]^n \mapsto [0,1]^n$ (i.e., $S := \mathbf{AFP}$), and letting the domain $[N]$ denote (some finite discretization of) the domain $[0,1]^n$ (i.e., $N = 2^{O(n)}$), the "lifted" communication problem $\mathbf{AFP} \circ \mathsf{IND}^N(\mathbf{x}, \mathbf{y})$ essentially corresponds to $\mathbf{AFPC}$, albeit with *unbounded* Lipschitz constraints on f and g. That is:

Key Observation: When the input vectors $\mathbf{x}$ and $\mathbf{y}$ are interpreted as the truth tables of (discrete) functions $f : [0,1]^n \mapsto [0,1]^m$ and $g : [0,1]^m \mapsto [0,1]^n$ respectively, the index gadget $\mathsf{IND}(\mathbf{x}_i, \mathbf{y}) = \mathbf{y}[\mathbf{x}_i] = g(f(i))$ encodes the truth table of the composed function $h := g \circ f$.

Unfortunately, Theorem II.3 cannot be invoked in a black-box fashion to conclude Theorem II.1, the main reason being that the decomposition $h(x) = g(f(x))$ produced by these proofs does not obey any (nontrivial) Lipschitz constraints on f and g (even though $h := g \circ f$ is known to be Lipschitz).

Embedding these geometric Lipschitz constraints into the Raz-Mackenzie simulation theorem is a substantial conceptual and technical obstruction, since the simulation argument (of both [GPW15], [RM99]) heavily relies on the invariant that the *unqueried coordinates* in the simulating decision-tree can retain *any* potential value (intuitively, this invariant ensures that there's enough remaining "entropy" in the inputs so that the simulating decision tree does not get "stuck"). This property essentially requires the set of inputs of Alice and Bob to have a *product* structure (which in our context means that f, g assign *independent* values to each point in their domain, i.e., $f \in \times_x \mathcal{B}(x)$ and $g \in \times_y \mathcal{B}(y)$, where $\mathcal{B}(x)$ (resp. $\mathcal{B}(y)$) are some predetermined sets of values to which each x (resp. y) is mapped to). For further elaboration on this, see Section 5.4 of the full version of this paper.

We show how to modify the [GPW15] simulation argument so that the decomposition (lifting) of h into $g \circ f$ accommodates simultaneously the Lipschitz constraints on f and g claimed in Theorem II.1, *and* the product-structure constraint on f, g, at the price of slightly increasing the dimension m of the "intermediate" domain (i.e., the range of f) so that $m \gg n$ yet still $m = O(n)$.

Indeed, increasing the dimension of f's range enables us to replace the global Lipschitz constraints on f with *local* "displacement-like" conditions of the form $f(x) \in \mathcal{B}(x)$ where $\mathcal{B}(x)$ is some large enough "local" neighborhood of x in $[0,1]^m$.[8] Replacing the Lipschitz constraint on f with the above local-displacement constraint has another important feature, namely, it ensures that f is in fact *bi-Lipschitz*, which is necessary to facilitate the desired Lipschitz constraint on g. To accommodate the Lipschitz property of g in a similar product-structure fashion, we rely on the local-displacement property of the *composed* function h of [HPV89] and on so-called *Lipschitz-extension* arguments, which allow us to extend any partial Lipschitz function g from any subset of points to its entire domain ($[0,1]^m$) without increasing g's Lipschitz constant. The formal construction can be found in Section 5.3 of the full version of this paper [RW16].

The lower bound we obtain in Theorem II.1 holds for (the promise problem of) finding a $\lambda_1\lambda_2\alpha/43$-fixed-point of $g \circ f$. The constant-factor loss is the cost that we pay to retain the product structure necessary for a simulation theorem. Improving our lower bound further so that it holds for larger approximation parameters (ideally, for $\varepsilon = \lambda_1\lambda_2\alpha + 1$) requires decomposing h into $g \circ f$ in a slightly "smoother" fashion, so that $\lambda_1\lambda_2$ is smaller by an absolute constant factor (ideally, by a factor > 43). We discuss this direction further in Section VI of the Appendix.[9]

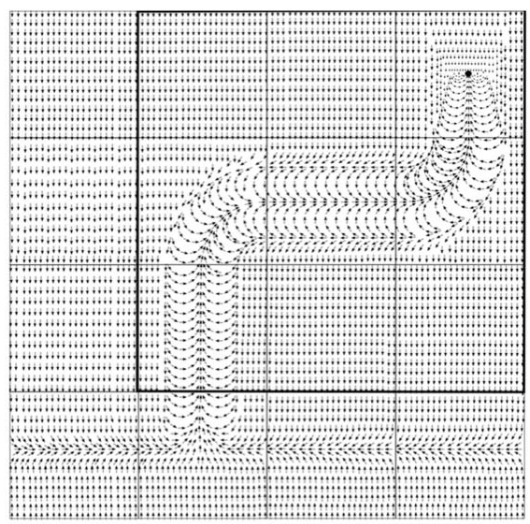

Figure 1. An illustration of the geometric construction of [HPV89] (cf. [Bab14]). The arrows in the figure correspond to displacements $h(x) - x$. Each function $h \in \mathcal{H}$ is a "continuous interpolation" of an (exponentially long) *discrete* path on a finite grid of $[0,1]^n$, whose endpoint is the *unique* hard-to-find fixed-point of h. The "geometric simulation theorem" we prove decomposes h (i.e., each arrow in the picture) into $g \circ f$ using the "index" gadget, in a way that ensures f and g are both Lipschitz and have a product structure.

[8] Intuitively, since distances are measured in the ℓ_∞ norm, allowing the dimension of the range of f to be $\gg n$ allows us to "embed" exponentially large local balls into $[0,1]^m$, one for each x in the domain of f, and these disjoint local neighborhoods form the range of all possible functions f Alice may receive. See Section 5.3 and Figure 4 in [RW16].

[9] In short, the main reason we believe such improvement is plausible is that our current proof does not make direct use of the premise that the "lifted" function $h = g \circ f$ is itself guaranteed to be λ-Lipschitz (for a constant $\lambda > 1$ arbitrarily close to 1), but only uses a weaker property, namely, that h has "local displacements": $\|h(x) - x\|_\infty \leq 5\varepsilon \;\; \forall x$.

C. From $\mathbf{AFPC}$ to $\mathbf{ANE}$: A High-level Proof Overview of Theorem II.2

We sketch the proof of the reduction for the two-player case, which shows that any (deterministic or randomized) two-party communication protocol that finds an $\Omega(1/\log^2 N)$-ANE in an $N \times N$ game, can be used, with no extra communication, to recover a $(2\lambda_1\lambda_2\alpha)$-approximate Brouwer fixed-point of $g \circ f$, assuming f and g are λ_1 and λ_2 Lipschitz, respectively.[10]

Our reduction is inspired by a recent reduction due to Babichenko [Bab14] (in turn inspired by a blog post of Shmaya [Shm]), who used it to relate the approximate Nash problem to the approximate fixed-point problem in the weaker query oracle model. The basic idea behind the reduction is that Alice and Bob can translate their respective input functions (f, g resp.) to the fixed-point problem, into convex payoff functions in which Alice's goal is to match the image of Bob's action under her function f, and similarly Bob's goal is to match the image of Alice's action under his function g, where "pure actions" are points in some finite ($\alpha > 0$) grid of the m-dimensional (resp. n-dimensional) cubes. More formally, Alice and Bob can use their respective input functions to define (using no communication at all!) a two-player game with the following payoff functions:

$$u_A(x, y) = -\frac{1}{m} \cdot \|x - f(y)\|_2^2 \quad ,$$
$$u_B(x, y) = -\frac{1}{n} \cdot \|g(x) - y\|_2^2.$$

Crucially, defining these payoff functions requires no interaction, since Alice's payoff only depends on f, and similarly for Bob (note that the size of the game is $N = (1/\alpha)^m$ as this is the number of α-grid points in the m (resp. n) dimensional cube, and that the normalization by m (n) ensures that payoffs are in $[-1, 1]$).

Now consider, for the sake of simplicity, that Alice and Bob have some protocol π that finds an *exact* Nash equilibrium (μ, σ) of the above game. Intuitively, (μ, σ) must be a *pure* equilibrium: Indeed, by definition of Alice's payoff and the convexity of the ℓ_2 norm, it is easy to see that for any equilibrium strategy σ played by Bob, Alice has a *unique* best response $x^* := \mathbb{E}_{y\sim\sigma}[f(y)]$ (this is essentially the well-known fact that expectation is the minimizer of the variance). An analogous argument shows that Bob's unique best response to any strategy μ played by Alice is $y^* := \mathbb{E}_{x\sim\mu}[g(x)]$. Since x^* and y^* are pure strategies, this means that any (exact) equilibrium must have the form $x^* = f(y^*)$ and $y^* = g(x^*)$. Combining the two together, we have $y^* = g(x^*) = g(f(y^*))$, so y^* is an exact fixed-point of $g \circ f$.

Alas, the argument above has a subtle flaw: the point $x^* := \mathbb{E}_{y\sim\sigma}[f(y)]$ might not lie on the (α) grid, in which case it is not a legitimate pure strategy of Alice (similarly for Bob's best response y^*), so the argument above is not precise (this is no surprise, as $g \circ f$ need not have an *exact* fixed-point on the discrete grid). However, what *does* turn out to be true is that any "good enough" ($\approx 1/n^2 = \Theta_\alpha(1/\log^2 N)$) approximate Nash equilibrium (μ, σ) of the above game, must be *entirely supported on the unique grid cubes* $C(x^*), C(y^*)$ that *contain* the points x^*, y^* respectively. In fact, we show this more generally for any good enough approximate *well-supported* (mixed) equilibrium (see Section 6 in [RW16] for formal definitions), and then use an argument due to [Bab14] that allows us to convert it to a (standard) ANE. (We remark that the analogous step for the k-player reduction involves a more sophisticated argument recently employed by [CCT15], which we show can be implemented in a distributed manner to facilitate the reduction). One can then use the Lipschitz properties of f and g to argue that "rounding" the exact fixed-point $y^* := \mathbb{E}_{x\sim\mu}[g(x)]$ on Bob's corresponding grid-cube (found by the protocol π), incurs an additive precision-loss of $\approx \lambda_1\lambda_2\alpha$, hence π can be used to recover a $(2\lambda_1\lambda_2\alpha)$-approximate fixed-point of $g \circ f$. For the formal proof, we refer the reader to Section 6 in the full version of this paper [RW16].

III. Organization

Due to space constraints, this manuscript contains only the high-level structure of the proof of our main result (Section V). The formal construction and proofs of both Theorem II.1 and Theorem II.2 can be found in the full version of this paper ([RW16]). We conclude with some interesting open problems and a discussion of our techniques in Section VI.

IV. Preliminaries

We denote by $\|x\|_\infty := \max_i |x_i|$ the ℓ_∞ (max) norm, and by $\|x\|_2$ the ℓ_2 (Euclidean) norm. For a multi-set S of $[n], \mathcal{U}(S)$ denotes the uniform distribution over S. The family of all distributions over a set S is denoted $\Delta(S)$ (for example, $\Delta([n])$ is the family of all distributions over $[n]$, and $\mathcal{U}([n]) \in \Delta([n])$). We let $\mathbf{e}_i$ denote the i'th vector in the standard n-dimensional basis.

A. Geometric Definitions and Notation

Our results involve geometric concepts and constructions. Since communication complexity is a discrete model, we consider (standard) discrete analogous of

[10] The claim for k-player constant-action games follows a similar-in-spirit reduction from a *multiparty* variant of the $\mathbf{AFPC}$ problem which in turn admits an easy reduction from the two-party $\mathbf{AFPC}$ problem, but this time the reduction applies even to k-party protocols that merely find an $\Omega(1)$-ANE in k-player constant-action games. See Section 6.3.2 in the full version of this paper.

continuous geometric concepts, and make a recurring use of discretization throughout the paper. We denote by

$$G_{\delta,n} : \{x \in [0,1]^n \ : \ x_i \in \delta\mathbb{N}\}$$

the δ-grid on the n-dimensional solid cube. A set $C \subseteq G_{\delta,n}$ is called a *δ-grid-cube* (or simply *cube*) if there is some $x \in G_{\delta,n}$ such that $C = \{x + \delta \cdot \mathbf{e}_i \mid i \in [n]\}$. For a point $x' \in [0,1]^n$, we sometimes use the shorthand $C_\delta(x')$ to denote the (unique) δ-grid-cube containing x'.[11] We denote by $\overline{C} := \times_{i\in[n]} \ [x, x + \delta \cdot \mathbf{e}_i]$ the (continuous) subcube of the solid cube $[0,1]^n$ induced by C.

Definition IV.1 (Lipschitz functions)**.** *We say that a mapping $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ is λ-Lipschitz if for every $x, y \in [0,1]^n$,*

$$\|f(x) - f(y)\|_\infty \leq \lambda \|x - y\|_\infty.$$

Note that the above condition is well defined even when $m \neq n$. When the domain of f is discrete, say $f : G_{\delta,n} \mapsto [0,1]^m$, the condition above ranges over all points $(x, y) \in G^2_{\delta,n}$, and in this case (whenever not clear from context) we will say that f is λ-Lipschitz on $G_{\delta,n}$. The following simple proposition follows directly from the triangle inequality.

Proposition IV.2 (Transitivity of Lipschitz continuity)**.** *If $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ is λ_1-Lipschitz, $g : \mathbb{R}^m \mapsto \mathbb{R}^n$ is λ_2-Lipschitz, then the composed function $g \circ f : \mathbb{R}^n \mapsto \mathbb{R}^n$ is $(\lambda_1\lambda_2)$-Lipschitz.*

Lipschitz Extensions: The following known lemma asserts that it is possible to extend any (ℓ_∞) Lipschitz function from an arbitrary subset of points in its domain to any superset containing it, in a continuous fashion *without increasing* the Lipschitz constant of the function.[12]

Lemma IV.3 (Lipschitz Extension, essentially [Whi33])**.** *Let $A \subset \mathbb{R}^n$ be a non-empty set. If $f : A \mapsto \mathbb{R}^m$ is λ-Lipschitz on A (in the ℓ_∞ sense), then the function $\bar{f} : \mathbb{R}^n \mapsto \mathbb{R}^m$ whose coordinates are defined by*

$$\bar{f}_i(x) := \inf_{z \in A} \{f_i(z) + \lambda \cdot \|x - z\|_\infty\}$$

is λ-Lipschitz on $\mathbb{R}^n$.

An immediate proof of this lemma using [Whi33] can be found in Appendix A of [RW16]).

[11] If a coordinate x_i is a multiple of δ, associate it with the subcube for which x_i is the minimum value of the ith coordinate (for example).

[12] Analogous extension theorems for arbitrary metric spaces in $\mathbb{R}^m$ are generally false, in the sense that the Lipschitz constant resulting from any extension might strictly increase (see [ACJ04] for a survey on extension theorems).

Convention. Every discrete function $f : G_{\alpha,n} \mapsto G_{\alpha,m}$ (i.e., a mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$) can be encoded using a vector $\in G_{\alpha,m}^{G_{\alpha,n}}$. Throughout the paper, we shall refer to this vector of values as the *truth table* of f.

B. Complexity Measures Notation

Definition IV.4 (Search Problems (Relations))**.** *A search problem $S(x)$ is defined by a subset $S \subseteq \mathcal{X} \times \mathcal{Z}$. A search problem is called* total *if for all $x \in \mathcal{X}$ there is at least one $z \in \mathcal{Z}$ for which $(x, z) \in S$ (otherwise, S is a* promise *search problem). We say that a decision tree solves $S(x)$ if for any input x, it outputs some $z \in \mathcal{Z}$ such that $(x, z) \in S$.*

Similarly, a two-party search-problem $S(x, y)$ is defined by a subset $S \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$, and S is a total *search problem if for all x, y there is at least one z for which $(x, y, z) \in S$. We say that a communication protocol solves a total relation $S(x, y)$ if for any input pair (x, y), it outputs some $z \in \mathcal{Z}$ such that $(x, y, z) \in S$. An analogous definition applies to k-party relations $S \subseteq \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_k \times \mathcal{Z}$.*

We will be interested in the following complexity measures for a search problem $S \subseteq \mathcal{X} \times \mathcal{Z}$:

- $\mathsf{D}^{\mathsf{QC}}(S)$ denotes the deterministic query complexity of S, i.e., the smallest depth of a decision tree that outputs a correct solution for S on every input.
- $\mathsf{R}^{\mathsf{QC}}{}_\rho(S)$ denotes the (worst-case) depth of a randomized decision tree that outputs a correct solution for S with probability $\geq 1 - \rho$ for every input.

For a two-party search problem $S \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$,

- $\mathsf{ND}^{\mathsf{CC}}(S)$ denotes the cheapest non-deterministic communication protocol[13] which solves S.
- $\mathsf{D}^{\mathsf{CC}}(S)$ denotes the cheapest deterministic communication protocol which solves S.
- $\mathsf{R}^{\mathsf{CC}}{}_\rho(S)$ denotes the (worst-case) communication cost of the cheapest randomized two-party communication protocol which outputs a correct solution for $S(x, y)$ with probability $\geq 1 - \rho$ for all inputs $(x, y) \in \mathcal{X} \times \mathcal{Y}$, over the randomness of the protocol.

By abuse of notation, for a k-party relation $S \subseteq \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_k \times \mathcal{Z}$, we use the same communication complexity measures ($\mathsf{ND}^{\mathsf{CC}}(S)$, $\mathsf{D}^{\mathsf{CC}}(S)$ and $\mathsf{R}^{\mathsf{CC}}{}_\rho(S)$) to denote, respectively, the k-party *Number-In-Hand*

[13] A non-deterministic communication protocol for S is a protocol π in which a referee (Merlin) who has access to both player's inputs (x, y), can initially give Alice and Bob an *advice* $a = a(x, y)$, and after this step the protocol π proceeds as usual. The protocol should output a valid solution z to S (s.t $(x, y, z) \in S$ or $\perp$ if no such z exists) for any input pair (x, y). The cost of the protocol is the sum of bits communicated in both a and π. (For a more formal definition and a thorough overview of non-deterministic communication complexity and its importance and relations to other models of computation, see [KN97]).

(NIH) non-deterministic, deterministic and randomized communication complexity of the k-party problem S, where the input of player $i \in [k]$ is $x_i \in \mathcal{X}_i$.

V. The Two-Party Deterministic Communication Complexity of $\mathbf{AFPC}$

We now formally define $\mathbf{AFPC}$, the two-party problem of finding an approximate Brouwer fixed point of a composition of two Lipschitz functions. The problem is defined in Figure 2.

It is important to note that, whenever $\varepsilon \geq (\lambda_1\lambda_2 + 1)\alpha$, $\mathbf{AFPC}_{\alpha,(n,\lambda_1),(m,\lambda_2),\varepsilon}$ is a *total search problem*: Indeed, Proposition IV.2 guarantees that the composed function $h := g \circ f : G_{\alpha,n} \mapsto G_{\alpha,n}$ is $\lambda_1\lambda_2$-Lipschitz on $G_{\alpha,n}$, hence Lemma IV.3 ensures it is possible to extend h to the entire *solid* cube $[0,1]^n$ in a way that it remains $(\lambda_1\lambda_2)$-Lipschitz on the solid cube. By Brouwer's fixed-point theorem, the extended function must have an *exact* Brouwer fixed point $x \in [0,1]^n$, and rounding x to the closest grid point $x' \in G_{\alpha,n}$ ensures (via a standard triangle-inequality argument) that x' is a $(\lambda_1\lambda_2+1)\alpha \leq \varepsilon$-fixed point of $g \circ f$. In particular, we conclude that a $(2\lambda_1\lambda_2\alpha)$ fixed-point must always exist. [14]

For $\varepsilon < (\lambda_1\lambda_2 + 1)\alpha$, $\mathbf{AFPC}$ is a *promise problem*, where the players are guaranteed that the ε-fixed point exists. (A protocol can behave arbitrarily on inputs with no ε-fixed point.)

Our main result states that any two-party deterministic communication protocol solving the following promise version of $\mathbf{AFPC}$ requires exponential communication (in the dimension n).

Theorem V.1 (Deterministic Communication Lower bound for $\mathbf{AFPC}$). *There are universal constants $\alpha \in (0,1)$, $\lambda_1, \lambda_2 \geq 2$ such that for every $n \geq 2$ and $m = O_\alpha(n)$,*

$$\mathsf{D}^{\mathsf{CC}}\left(\mathbf{AFPC}_{\alpha,(n,\lambda_1),(m,\lambda_2),\frac{\lambda_1\lambda_2\alpha}{43}}\right) \geq 2^{\Omega_\alpha(n)}.$$

The key step in the proof of Theorem V.1 is showing that the deterministic communication complexity of $\mathbf{AFPC}$ with the above parameters is bounded below by the *deterministic query complexity* of $\mathbf{AFP}_{n,\alpha,\lambda,\varepsilon}$, the search problem of finding an ε-approximate fixed point of a λ-Lipschitz function $h : G_{\alpha,n} \mapsto G_{\alpha,n}$ (see Section 5.1 in the full version [RW16] for the formal definition). More formally, we prove

Lemma V.2 (Geometric Simulation Lemma for $\mathbf{AFPC}$). *There are universal constants $\delta \in (0,1)$, $\lambda \geq 2$ and $D \geq 244$, such that for every $n \geq 2$ and $m = O_\alpha(n)$,*

$$\mathsf{D}^{\mathsf{CC}}\left(\mathbf{AFPC}_{\alpha,(n,2D+1),(m,\frac{21\varepsilon}{D\alpha}),\varepsilon}\right) \geq \Omega(m) \cdot \mathsf{D}^{\mathsf{QC}}\left(\mathbf{AFP}_{n,\alpha,\lambda,\varepsilon}\right),$$

where $\alpha = \delta/1200, \varepsilon = \lambda\delta/1200$.

Since the query complexity of $\mathbf{AFP}_{n,\alpha,\lambda,\varepsilon}$ was previously shown to be $2^{\Omega_\lambda(n)}$ (for the formal statement, see e.g. Theorem 5.2 in [RW16]), Lemma V.2 will directly imply Theorem V.1, by setting $\lambda_1 := 2D+1, \lambda_2 := 21\varepsilon/(D\alpha)$, and observing that for this choice of parameters, we have $\lambda_1\lambda_2\alpha/43 \leq \varepsilon$, so Theorem V.1 follows.

The main part of the proof of Theorem V.1 is therefore devoted to the construction and proof of Lemma V.2. Due to space constraints, we defer the proof to the full version of this paper (see Section 5 in [RW16]).

VI. Discussion and Open Problems

This paper initiates the study of distributed computation of approximate fixed-point problems ($\mathbf{AFPC}_{\alpha,(n,\lambda_1),(m,\lambda_2),\varepsilon}$). We prove that finding an $\varepsilon = (\lambda_1\lambda_2\alpha/43)$-fixed point of a composition of two Lipschitz functions $g \circ f$ requires exponential communication in the dimension n, at least for deterministic protocols. While this is a highly nontrivial approximation parameter, an intriguing question is whether the same lower bound applies for the slightly looser approximation parameter $\varepsilon = (\lambda_1\lambda_2 + 1)\alpha$ (or even $\varepsilon = \lambda_1\lambda_2\alpha$), at which the problem becomes a total search problem and reduces to the (two-party and multiparty) problems of finding approximate Nash equilibria.

One plausible approach for "bridging" this constant gap in Lemma V.2 is to perform the "lifting" argument (the decomposition $h = g \circ f$) in a slightly "smoother" manner, so that the Lipschitz constants of f and g satisfy $2\lambda_1\lambda_2\alpha \leq \varepsilon$ instead of $2\lambda_1\lambda_2\alpha \approx 43\varepsilon$, as our current construction provides. This is essential for the lower bound to go through, since the maximum displacement of the composed function $h = g \circ f \in \mathcal{H}_{\delta,\lambda,n}|_\alpha$ is 5ε, and therefore finding a 43ε-fixed point of $g \circ f$ is trivial (as opposed to finding an ε-fixed point). In fact, our current proof only exploits this property, i.e., that the "lifted" function h has bounded-displacements, while the geometric construction of [HPV89] guarantees much more than that, namely, that every $h \in \mathcal{H}_{\delta,\lambda,n}|_\alpha$ is λ-Lipschitz for (an absolute constant λ). A natural idea is to redefine the class $\mathcal{G}$ in the proof of Lemma V.2 to be the class of all $O(\lambda)$-Lipschitz functions.

Unfortunately, it is not clear how to exploit this further property in the simulation argument of of [RM99], [GPW15], since the simulation invariants we maintain

[14] We remark that the *non-deterministic* communication complexity of $\mathbf{AFPC}$ in this regime is only $O(\log|G_{\alpha,n}|) = O_\alpha(n)$ (since Alice and Bob can exchange these many bits to *verify* that a given x satisfies $\|x - g \circ f(x)\|_\infty \leq \varepsilon$).

$\mathbf{AFPC}_{\alpha,(n,\lambda_1),(m,\lambda_2),\varepsilon}$
Let $\alpha \in (0,1)$, $m \geq n$, $\lambda_1, \lambda_2 \geq 0$, and $\varepsilon \in (0,1]$ be publicly known parameters. INPUTS : Alice receives a truth table of a λ_1-Lipschitz function $f : G_{\alpha,n} \mapsto G_{\alpha,m}$. Bob receives a truth table of a λ_2-Lipschitz function $g : G_{\alpha,m} \mapsto G_{\alpha,n}$. OUTPUT: $x \in G_{\alpha,n}$ such that $\|g(f(x)) - x\|_\infty \leq \varepsilon$, i.e., an ε-fixed point of $g \circ f$ (or $\perp$ if such point doesn't exist).

Figure 2. The two-party communication problem of finding an approximate fixed point of $g \circ f$.

require the input sets $\mathcal{F}, \mathcal{G}$ to be *product sets* (i.e., that values to different coordinates $f(x), f(x')$ can be chosen *independently* from some predefined set of values). Indeed, a simple calculation[15] shows that the stronger condition we seek ($2\lambda_1\lambda_2\alpha \leq \varepsilon$) requires *breaking the product structure of* $\mathcal{F}, \mathcal{G}$. While we believe this modification should be possible to implement in our specific settings (again, using the promise that the function h is guaranteed to be λ-Lipschitz), this seems to require further geometric insights and a new simulation invariant (ensuring that the "Thickness lemma" and the "Projection lemma" go through).

Finally, we recall that the query complexity of the approximate fixed-point problem ($\mathbf{AFP}$) was recently shown to be exponential even in the *randomized* query model ([Bab14]), so a *randomized* analogue of our simulation theorem (Lemma V.2) would have implied an exponential *randomized* communication lower bound for $\mathbf{AFPC}$. While the Raz-McKenzie simulation theorem and our adapted geometric variant of it (Lemma V.2) rely on an "adversarial" argument which currently applies only to the deterministic communication complexity model, a recent line of work has been focused on *randomized simulation theorems* ([GP13], [GLM+15]). Alas, these theorems require a lower bound on stronger measures than randomized query complexity. Notwithstanding, we believe that proving a randomized analogue of the Raz-McKenzie simulation theorem (and hence of Theorem II.1) is a natural and fascinating open problem.

ACKNOWLEDGEMENT

We would like to thank Yakov Babichenko for pointing us to the composed Brouwer fixed-point problem and for many insightful conversations, including the reduction from the two-player to the multi-player Nash problem. We would also like to thank Aviad Rubinstein for his valuable inputs.

[15]This follows by observing that $\lambda_1 \geq \frac{\|f(x)-f(x')\|_\infty}{\alpha}$ for neighboring points x, x' s.t $\|x = x'\|_\infty = \alpha$, and on the other hand, $\lambda_2 \geq \frac{\|g(f(x))-g(f(x'))\|_\infty}{\|f(x)-f(x')\|_\infty}$, so $\lambda_1\lambda_2\alpha \geq \|g(f(x))-g(f(x'))\|_\infty$. If $\mathcal{F}, \mathcal{G}$ are a product sets, it is not hard to see that the RHS has to be at least as large as $\max_x \|h(x) - h(x')\|_\infty = 5\varepsilon$, which is not good enough.

REFERENCES

[ABB+15] Andris Ambainis, Kaspars Balodis, Aleksandrs Belovs, Troy Lee, Miklos Santha, and Juris Smotrovs. Separations in query complexity based on pointer functions. *CoRR*, abs/1506.04719, 2015.

[ACJ04] Gunnar Aronsson, Michael G. Crandall, and Petri Juutinen. A tour of the theory of absolutely minimizing functions. *Bulletin of the AMS*, 41(4):439–505, 2004.

[Bab14] Yakov Babichenko. Query complexity of approximate nash equilibria. In *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 535–544, 2014.

[BdW02] Harry Buhrman and Ronald de Wolf. Complexity measures and decision tree complexity: a survey. *Theor. Comput. Sci.*, 288(1):21–43, 2002.

[Bor85] Kim Border. *Fixed point theorems with applications to economics and game theory*. Cambridge [Cambridgeshire], Cambridge University Press., 1985.

[BPR16] Y. Babichenko, C. H. Papadimitriou, and A. Rubinstein. Can almost everybody be happy? PCP for PPAD and the inapproximability of Nash. In *Proceedings of the 7th Annual Conference on Innovations in Theoretical Computer Science (ITCS)*, 2016.

[CCT15] Xi Chen, Yu Cheng, and Bo Tang. Well-supported versus approximate nash equilibria: Query complexity of large games. *CoRR*, abs/1511.00785, 2015.

[CD09] X. Chen and X. Deng. On the complexity of 2D discrete fixed point problem. *Theoretical Computer Science*, 410:4448–4456, 2009.

[CDF$^+$15] Artur Czumaj, Argyrios Deligkas, Michail Fasoulakis, John Fearnley, Marcin Jurdzinski, and Rahul Savani. Distributed methods for computing approximate equilibria. *CoRR*, abs/1512.03315, 2015.

[CDT09] X. Chen, X. Deng, and S.-H. Teng. Settling the complexity of computing two-player Nash equilibria. *Journal of the ACM*, 56(3):14, 2009.

[CS04] Vincent Conitzer and Tuomas Sandholm. Communication complexity as a lower bound for learning in games. In *Machine Learning, Proceedings of the Twenty-first International Conference (ICML 2004), Banff, Alberta, Canada, July 4-8, 2004*, 2004.

[DGP09] C. Daskalakis, P. W. Goldberg, and C. H. Papadimitriou. The complexity of computing a Nash equilibrium. *SIAM Journal on Computing*, 39(1):195–259, 2009.

[GLM$^+$15] Mika Göös, Shachar Lovett, Raghu Meka, Thomas Watson, and David Zuckerman. Rectangles are nonnegative juntas. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14-17, 2015*, pages 257–266, 2015.

[Göö15] Mika Göös. Lower bounds for clique vs. independent set. *Electronic Colloquium on Computational Complexity (ECCC)*, 22:12, 2015.

[GP13] Paul W. Goldberg and Arnoud Pastink. On the communication complexity of approximate nash equilibria. *CoRR*, abs/1302.3793, 2013.

[GP14] Mika Göös and Toniann Pitassi. Communication lower bounds via critical block sensitivity. In *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 847–856, 2014.

[GPW15] Mika Göös, Toniann Pitassi, and Thomas Watson. Deterministic communication vs. partition number. *Electronic Colloquium on Computational Complexity (ECCC)*, 22:50, 2015.

[HM10] Sergiu Hart and Yishay Mansour. How long to equilibrium? the communication complexity of uncoupled Nash equilibrium procedures. *Games and Economic Behavior*, 69(1):107–126, 2010.

[HMC02] Sergiu Hart and Andreu Mas-Colell. Uncoupled dynamics cannot lead to nash equilibrium. Discussion paper series, The Federmann Center for the Study of Rationality, the Hebrew University, Jerusalem, 2002.

[HPV89] Michael D. Hirsch, Christos H. Papadimitriou, and Stephen A. Vavasis. Exponential lower bounds for finding Brouwer fixed points. *Journal of Complexity*, 5(4):379–416, 1989.

[KN97] E. Kushilevitz and N. Nisan. *Communication Complexity*. Cambridge University Press, 1997.

[LH64] C. E. Lemke and J. T. Howson, Jr. Equilibrium points of bimatrix games. *SIAM Journal*, 12(2):413–423, 1964.

[LMM03] Richard J. Lipton, Evangelos Markakis, and Aranyak Mehta. Playing large games using simple strategies. In *Proceedings of the 4th ACM Conference on Electronic Commerce*, EC '03, pages 36–41, New York, NY, USA, 2003. ACM.

[LNNW95] László Lovász, Moni Naor, Ilan Newman, and Avi Wigderson. Search problems in the decision tree model. *SIAM J. Discrete Math.*, 8(1):119–132, 1995.

[Mat07] Jiri Matousek. *Using the Borsuk-Ulam Theorem: Lectures on Topological Methods in Combinatorics and Geometry*. Springer Publishing Company, Incorporated, 2007.

[NW95] Noam Nisan and Avi Wigderson. On rank vs. communication complexity. *Combinatorica*, 15(4):557–565, 1995.

[Pap94] C. H. Papadimitriou. On the complexity of the parity argument and other inefficient proofs of existence. *Journal of Computer and System Sciences*, 48(3):498–532, 1994.

[RM99] Ran Raz and Pierre McKenzie. Separation of the monotone NC hierarchy. *Combinatorica*, 19(3):403–435, 1999.

[Rub16] Aviad Rubinstein. Settling the complexity of computing approximate two-player nash equilibria. *CoRR*, abs/1606.04550, 2016.

[RW16] Tim Roughgarden and Omri Weinstein. On the communication complexity of approximate fixed points. *Electronic Colloquium on Computational Complexity (ECCC)*, 23:55, 2016.

[Sca67] H. Scarf. The approximation of fixed points of a continuous mapping. *SIAM Journal*, 15:1328–1343, 1967.

[Sha64] Lloyd S. Shapley. Some Topics in Two-person Games. Technical report, RAND CORP SANTA MONICA CA, June 1964.

[Shm] E. Shmaya. Brouwer implies Nash implies Brouwer. The Leisure of the Theory Class blog post, http://theoryclass.wordpress.com/2012/01/05/brouwer-implies-nash-implies-brouwer/, year = 2012.

[Whi33] Hassler Whitney. Abstract for "Analytic extensions of differentiable functions defined on closed sets". *Bull. Am. Math. Soc.*, 39(1):31, 1933. Abstract only; published in *Trans. Am. Math. Soc.* **36**:1 (1934), pp. 63–89. JFM:59.0260.04.