Electronic Colloquium on Computational Complexity, Report No. 55 (2016)

On the Communication Complexity of Approximate Fixed Points

Tim Roughgarden* Omri Weinstein†

April 10, 2016

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 $\varepsilon = \Omega(1/\log^2 N)$-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 $\varepsilon = \Omega(1)$-Nash equilibrium in a k-player constant-action game is at least $2^{\Omega(k/\log k)}$ (while the nondeterministic communication complexity is only $O(k)$).

*Department of Computer Science, Stanford University. Supported in part by NSF grant CCF-1524062.
†Department of Computer Science, New York University. Research Supported by a Simons Society Junior Fellowship.

ISSN 1433-8092

1 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 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. Babichenko [Bab14] recently extended this lower bound to randomized query algorithms (decision-trees) as well. 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 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?

[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.

[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.

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 $\lim_{\alpha\to 0}\frac{\varepsilon}{\alpha} = \infty$). Since our lower bound trivially implies an exponential lower bound on the deterministic *query* complexity of computing an ε-approximate fixed point for $\lambda_1\lambda_2$-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 2.1.1).

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].[4] 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 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. [**?**] 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

[4] For query complexity, there is an exponential (in the number of players) lower bound even for ε-approximate Nash equilibria with constant ε [Bab14, CCT15]. The computational complexity remains open for the constant ε case; it is known to be quasi-polynomial-time solvable [LMM03] and there are plausible conjectures under which no faster algorithm exists [BPR16].

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.

2 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 4.)

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 2.1 (Deterministic communication lower bound for **AFPC**)**.** *There are universal constants $\alpha \in (0,1)$, $\lambda_1, \lambda_2 > 1$ such that for every $n \geq 3$ 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 **AFPC** to the problem of computing an approximate Nash equilibrium (**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 2.2 (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)*
  $\mathsf{R}^{\mathsf{CC}}{}_\rho\left(\mathbf{ANE}_{K,\Omega(1/\log^2 K)}\right) \geq \mathsf{R}^{\mathsf{CC}}{}_\rho\left(\mathbf{AFPC}_{\alpha,(n,\lambda_1),(m,\lambda_2),2\lambda_1\lambda_2\alpha}\right)$*, where* $K = (1/\alpha)^m$.
- *(k-player games)*
  $\mathsf{R}^{\mathsf{CC}}{}_\rho\left(k\text{-}\mathbf{ANE}_{1/\alpha,3\alpha^3/16}\right) \geq \mathsf{R}^{\mathsf{CC}}{}_\rho\left(\mathbf{AFPC}_{\alpha,(n,\lambda_1),(m,\lambda_2),2\lambda_1\lambda_2\alpha}\right)$*, where* $k = O_\alpha(m\log m)$.

Theorem 2.2 implies that a slightly stronger version of Theorem 2.1 (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 Corollary 6.4 for the formal statement and a more elaborate discussion on this direction in Section 7).

2.1 Overview of Proofs and Techniques

"Lifting": Communication Lower Bounds from Query Lower Bounds. To prove Theorem 2.1, 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 fashion (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 $\log(\min\{|\mathcal{X}|, |\mathcal{Y}|\}) \cdot (\text{query complexity}(S))$, 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$^+$15, GP14]). The gadget g plays a crucial role in such results, as it ensures Alice and Bob cannot take "short-cuts" by avoiding queries made by the decision tree.[5] Thus the gadget g must be a sufficiently "hard" function to rule out such manipulations. We elaborate more on this in Section 5.4. We remark that simulation theorems have recently led to breakthrough results in complexity theory, including the resolution of the long-standing "Clique vs. Independent Set" problem [Göö15, GPW15], separation theorems between various deterministic and non-deterministic communication measures [GPW15, ABB$^+$15], and 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 $k = \text{poly}(N)$, 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 2.3 ([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 subsections, we explain the relevance of this theorem to the distributed approximate fixed-point problem (**AFPC**), and provide a streamlined overview of the proofs of our main results (Theorem 2.2 and Theorem 2.1).

[5] 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!

2.1.1 A High-Level Proof Overview of Theorem 2.1

The approximate fixed-point problem that we study (**AFPC**) has a "geometric" aspect to it, in that both of the input functions are required to be $O(1)$-Lipschitz.[6] 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 **AFPC** is bounded from below by the deterministic *query complexity* of the 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 **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 and Figure 1) 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) function $h : [0,1]^n \mapsto [0,1]^n$ (i.e., $S :=$ **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 **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 2.3 cannot be invoked in a black-box fashion to conclude Theorem 2.1, 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 when $h := g \circ f$ is known to be Lipschitz, as in in the [HPV89] construction). We elaborate more on this in Section 5.2.

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).

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 (as claimed in

[6] The problem is not interesting without the Lipschitz requirement. Intuitively, if f and g are random functions (say), then there are no non-trivial communication protocols for the problem. It is not difficult to turn this intuition into a strong lower bound.

Theorem 2.1) 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 $[0,1]^m$.[7] 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. A more detailed description of our construction can be found in Section 5.3.

The lower bound we obtain in Theorem 2.1 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, even for $2\lambda_1\lambda_2\alpha$) 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, 86 or more). We discuss this direction further in Section 7 of the Appendix.[8]

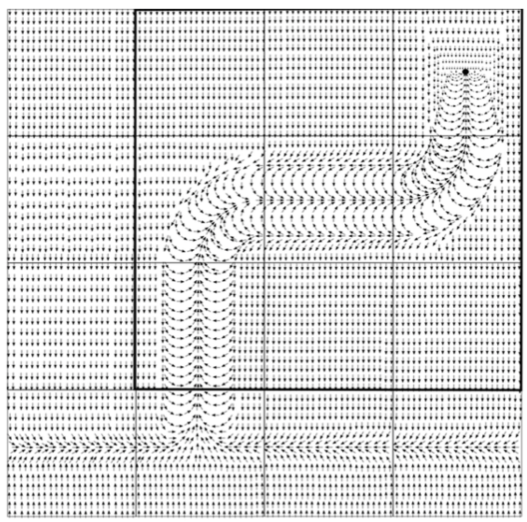

Figure 1: An illustration of the geometric construction of [HPV89] (cf. [Bab14]). The arrows in the figure correspond to displacements $h(x) - x$. Each function $h \in \mathcal{H}$ is a "continuous interpolation" of an (exponentially long) *discrete* path on a finite grid of $[0,1]^n$, whose endpoint 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 both f and g are both Lipschitz and have a product structure.

[7] 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 also Figure 4 in Section 5.3.

[8] 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": $\|h(x) - x\|_\infty \leq 5\varepsilon \;\; \forall x$.

2.1.2 From AFPC to ANE : A High-level Proof Overview of Theorem 2.2

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.[9]

Our reduction is inspired by a recent reduction due to Babichenko [Bab14] (in turn inspired by a blog post of Shmaya [Shm12]), 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:

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

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 analogues 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 6.1 for the definition), and then use an argument due to [Bab14] that allows us to convert it to a (standard) ANE (we remark that the analogues step for the k-player reduction involves a more sophisticated argument recently shown by [CCT15], which we show can be implemented in a distributed fashion). 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 $g \circ f$. The formal proof can be found in Section 6.3.

[9]The claim for k-player constant-action games follows a similar-in-spirit reduction from a *multiparty* variant of the **AFPC** problem which in turn admits an easy reduction from the two-party **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 for more details.

3 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.

3.1 Geometric Definitions and Notation

Our results involve geometric concepts and constructions. Since communication complexity is a discrete model, we consider (standard) discrete analogues of 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'.[10] 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 3.1 (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 3.2 (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.[11]

Lemma 3.3 (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 Section A of the Appendix. For subsets $A \subseteq B \subseteq \mathbb{R}^n$, we denote the MLE-extension of a function $f : A \mapsto G_{\delta,m}$ from A to B by $\bar{f}$. The MLE extension may produce real-valued points, and we define $\lceil\bar{f}\rceil$ as the "rounded MLE-extension" of $\bar{f}$ to $G_{\delta,m}$, obtained by rounding the coordinates of $\bar{f}$ up to the nearest multiples of δ. A standard triangle inequality argument implies that, if $B \subseteq G_{\alpha,n}$ and f is λ-Lipschitz on A, then $\lceil\bar{f}\rceil$ is $\leq (\lambda + \delta/\alpha)$-Lipschitz on B.

[10] 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).

[11] 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).

Conventions. 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.

3.2 Complexity Measures Notation

Definition 3.4 (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[12] 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* (NIH) non-deterministic, deterministic and randomized communication complexity of the k-party problem S, where the input of player $i \in [k]$ is $x_i \in \mathcal{X}_i$.

4 Two-Party Deterministic Communication Complexity of Approximate Fixed Points

We now formally define **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.

[12] 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]).

$\mathbf{AFPC}_{\alpha,(n,\lambda_1),(m,\lambda_2),\varepsilon}$
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 $g \circ f$.

Note that, whenever $\varepsilon \geq 2\lambda_1\lambda_2\alpha$, $\mathbf{AFPC}_{\alpha,(n,\lambda_1),(m,\lambda_2),\varepsilon}$ is a *total search problem*: Indeed, Proposition 3.2 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 A.1 (i.e., the (rounded) MLE extension of h) ensures it is possible to extend h to the entire *solid* cube $[0,1]^n$ in a way that it remains $(2\lambda_1\lambda_2)$-Lipschitz on the solid cube. By Brouwer's fixed-point theorem, the extended function must have an *exact* Brouwer fixed point $x \in [0,1]^n$, so rounding x to the closest grid point $x' \in G_{\alpha,n}$ ensures (via a standard triangle-inequality argument) that x' is a $2\lambda_2\lambda_1\alpha \leq \varepsilon$-fixed point of $g \circ f$. We conclude that such an approximate fixed-point must always exist. Notice that the *non-deterministic* communication complexity of **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 $\|x - g \circ f(x)\|_\infty \leq \varepsilon$). For $\varepsilon < 2\lambda_1\lambda_2\alpha$, **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 **AFPC** requires exponential communication (in the dimension n).

Theorem 4.1 (Deterministic Communication Lower bound for **AFPC**)**.** *There are universal constants $\alpha \in (0,1)$, $\lambda_1, \lambda_2 \geq 2$ such that for every $n \geq 3$ 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 4.1 is showing that the deterministic communication complexity of **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 for the formal definition). More formally, we shall prove

Lemma 4.2 (Geometric Simulation Lemma for **AFPC**)**.** *There are universal constants $\delta \in (0,1)$, $\lambda \geq 2$ and $D \geq 244$, such that for every $n \geq 3$ and $m = O_\alpha(n)$,*

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

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

Since the query complexity of $\mathbf{AFP}_{n,\alpha,\lambda,\varepsilon}$ was previously shown to be $2^{\Omega_\lambda(n)}$ (see Theorem 5.2 below), Lemma 4.2 will directly imply Theorem 4.1, 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 4.1 follows.

The main part of the proof of Theorem 4.1 is therefore devoted to the construction and proof of Lemma 4.2. This is the content of the next section.

5 Proof of Theorem 4.1 and the Geometric Simulation Lemma

In this section we prove Theorem 4.1, most of which is devoted to the proof of our geometric simulation lemma (Lemma 4.2). We begin by describing the approximate Brouwer fixed-point problem in the query model (**AFP**) and some important properties of the lower bound construction for this problem due to [HPV89]. We then describe the connection between the ("lifted" version of) the **AFP** problem and **AFPC**, via the Raz-McKenzie simulation theorem (Section 5.2), and finally provide the actual proof of Lemma 4.2 in Section 5.4. Throughout this section, we shall denote by

$$N := |G_{\alpha,n}| = (1/\alpha)^n \quad , \quad M := |G_{\alpha,m}| = (1/\alpha)^m$$

the respective sizes of the functions' domains.

5.1 Detour: The Query Complexity of Approximate Fixed Points

As discussed in the introduction, a central ingredient of our communication lower bound on **AFPC** is the following result of Hirsch et. al [HPV89] (recently strengthened by Babichenko [Bab14]), who settled the deterministic (resp. randomized) complexity of finding an approximate fixed point in the weaker *query complexity* model. In this section we state the results and properties of the construction of [HPV89] that will be relevant to our communication lower bound. We start by defining the approximate fixed-point search problem in the query oracle model (denoted **AFP**) in Figure 3.

$\mathbf{AFP}_{n,\alpha,\lambda,\varepsilon}$
INPUT: A λ-Lipschitz function $h : G_{\alpha,n} \mapsto G_{\alpha,n}$, and a parameter $\varepsilon > 0$. OUTPUT: $x \in G_{\alpha,n}$ such that $\|h(x) - x\|_\infty \leq \varepsilon$, i.e., an ε-fixed point of h. QUERIES: Each query is a point $x \in G_{\alpha,n}$ and the answer is $h(x)$.

Figure 3: The problem of finding an approximate fixed point of a (discrete) Lipschitz function in the query oracle model.

In [HPV89], the authors constructed a particular family of *continuous* Lipschitz functions on the solid cube $[0,1]^n$ with the following important properties:

Definition 5.1 (The Class $\mathcal{H}_{\delta,\lambda,n}$, [HPV89])**.** *For every $\delta < 1, \lambda > 1$ satisfying $\lambda\delta \leq 1200$, and every $n \geq 2$, there is a family of Lipschitz continuous functions $\mathcal{H}_{\delta,\lambda,n} \subseteq \{h : [0,1]^n \mapsto [0,1]^n\}$, such that every $h \in \mathcal{H}_{\delta,\lambda,n}$ has the following properties:*

(1) *(Discrete encoding, Lemma 11 in [HPV89]) h is completely determined by its values on the grid $G_{\delta,n}$, i.e., by a truth table $h : G_{\delta,n} \mapsto [0,1]^n$.*

(2) *(Lipschitz continuity, Lemma 9 in [HPV89]) h is λ-Lipschitz.*

(3) *(Unique fixed point and bounded displacements, Lemma 7 in [HPV89]) Let $\varepsilon = \frac{\lambda\delta}{1200}$. For every $h \in \mathcal{H}_{\delta,\lambda,n}$, there is a* unique *cube $C^h_\varepsilon \subset G_{\delta,n}$ such that for all $x \notin \overline{C^h_\varepsilon}$, $2\varepsilon < \|h(x)-x\|_\infty \le 5\varepsilon$. In particular, $\overline{C^h_\varepsilon}$ contains the (unique) exact fixed point of h, and $\forall\, x \in [0,1]^n$, $\|h(x)-x\|_\infty \le 5\varepsilon$. The fact that the maximum "displacement" of h is 5ε will be important for our proof.*

An illustration of this construction (adapted from [Bab14]) can be found in Figure 1. Since our goal is to somehow adapt the construction of [HPV89] to the communication complexity setup, we shall work, as standard, with a discretized version of the family $\mathcal{H}_{\delta,\lambda,n}$. To avoid confusion between the continuous and discrete versions, we will denote by

$$\mathcal{H}_{\delta,\lambda,n}|_\alpha$$

the "discretized" class of functions $\mathcal{H}_{\delta,\lambda,n}$ where each function h is evaluated on $G_{\alpha,n}$ and $h(x)$ is rounded to the nearest point $\arg\min_{z \in G_{\alpha,n}} \|z - h(x)\|_\infty$. In other words, every $h \in \mathcal{H}_{\delta,\lambda,n}|_\alpha$ is represented by a truth table of size $|G_{\alpha,n}| = (1/\alpha)^n$. Note that since every $h \in \mathcal{H}_{\delta,\lambda,n}$ is λ-Lipschitz on $[0,1]^n$, it must have an (exact) Brouwer fixed point. Thus, for any discretization parameter

$$\alpha \le \frac{\delta}{1200},$$

a standard rounding argument[13] implies that every h must have a $\lambda\alpha \le \frac{\delta\lambda}{1200} \le \varepsilon$ fixed point $x_0 \in G_{\alpha,n}$. By property (3), $x_0 \in C^h_\varepsilon$.

[HPV89] proved an exponential lower bound on the number of *deterministic* queries needed to solve **AFP** under the restricted class of input functions $\mathcal{H}_{\delta,\lambda,n}|_\alpha$:

Theorem 5.2 (Deterministic Query Complexity of Approximate Fixed Point, [HPV89] Theorem 2)**.** *There are universal constants δ, λ such that for any $n \ge 3$ and $\alpha \le \delta/1200$,*

$$\mathsf{D}^{\mathsf{QC}}\left(\mathbf{AFP}_{n,\alpha,\lambda,\delta\lambda/1200}\right) \ge (1/\delta)^{\Omega_\lambda(n)}.$$

Moreover, the lower bound holds under the restricted class of input functions $\mathcal{H}_{\delta,\lambda,n}|_\alpha$.

Remark 5.3. *[HPV89] proved Theorem 5.2 in the continuous setting where each function $h \in \mathcal{H}_{\delta,\lambda,n}$ maps the solid cube to itself, and each query is a point $x \in [0,1]^n$, while we restrict to discrete inputs and queries as mentioned above. We note that, once again, for any choice $\alpha \le \delta/1200$, $\|h(x) - h'(x)\|_\infty \le \alpha = \delta/1200 \ll \varepsilon$, and therefore Property (3) in Definition 5.1 implies that $\|h'(x) - x\|_\infty \ge 1.5\varepsilon$ (say) for any $x \notin \overline{C^h_\varepsilon}$. Since the lower bound in [HPV89] applies for any (deterministic) query algorithm that finds the cube C^h_ε, any solution for the discretized version (**AFP**) can be used to determine C^h_ε, and thus the lower bound in the discrete case is implied by the original lower bound of [HPV89] (Clearly, restricting to discrete queries can only drive up the query complexity).*

We note that recently, Babichenko [Bab14] strengthened the result of [HPV89], showing that the exponential query bound continues to hold even for *randomized* query algorithms with exponentially small success probability.

[13] In fact, rounding h's values on $G_{\alpha,n}$ can increase the Lipschitz constant of h (w.r.t $G_{\alpha,n}$) from λ to $\lambda + 1$ (this can be directly proved via the triangle inequality), but since our choice of λ will be arbitrary, we will ignore this minor point.

5.2 From AFP to AFPC : Raz-McKenzie Simulation

The following definition will be central to understanding the connection between the **AFP** search problem (in the query model) and the **AFPC** problem (in the communication model):

Definition 5.4 (g-Lift of Search Problems)**.** *Let $S : \Sigma^N \mapsto \Sigma$ be a search problem. For a two-party communication function $g : \mathcal{X} \times \mathcal{Y} \mapsto \Sigma$, the g-*lift of S *is the two-party relation*

$$S \circ g^N(\mathbf{x}, \mathbf{y}) := S\left((g(x_1, y_1), g(x_2, y_2), \ldots, g(x_N, y_N))\right).$$

The "inner" two-party function $g(x, y)$ is sometimes called the *gadget.* Raz and McKenzie [RM99] showed that, for any "canonical"[14] search problem S, if the gadget g is chosen to be the *index function*

$$\mathsf{IND}(x, y) := y[x],$$

then any decision-tree lower bound for S is also a lower bound on the *deterministic communication complexity* of $S \circ \mathsf{IND}^N$. In a recent work, Goos, Pitassi and Watson simplified and generalized their result to any (possibly partial) relation S, proving the following theorem:

Theorem 5.5 (Simulation Theorem, [RM99, GPW15])**.** *Let $S : \Sigma^N \to \Sigma$ be a search problem, and let $\mathbf{x} \in [k]^N$, $\mathbf{y} \in (\Sigma^k)^N$ where $k = N^{20}$ and $|\Sigma| \leq k^{1/1000}$ (note that $\mathbf{y}_i[\mathbf{x}_i] \in \Sigma$ for all $i \in [N]$). Then*

$$\mathsf{D}^{\mathsf{CC}}\left(S \circ \mathsf{IND}^N(\mathbf{x}, \mathbf{y})\right) \geq \Omega(\log N) \cdot \mathsf{D}^{\mathsf{QC}}\left(S\right).$$

In other words, Theorem 5.5 states that, when the input to S is "decomposed" via the index gadget (of large enough dimension $k = \mathrm{poly}(N)$), the most efficient (deterministic) protocol for solving $S \circ \mathsf{IND}^N$ is essentially a simulation of the optimal decision tree for S, in which Alice and Bob sequentially use $\log k = O(\log N)$ bits to query the input bits $\mathbf{z}_i = \mathbf{y}_i[\mathbf{x}_i]$ as dictated by the decision tree for S (Raz and McKenzie called such protocols "sequential protocols").

In order to describe how Theorem 5.5 is relevant to the setting of Lemma 4.2, it will be useful to identify both the set of coordinates $[N]$ and the alphabet Σ in the statement of Theorem 5.5, with the set $G_{\alpha,n}$, and the set of "indices" $[k]$ with the set $[M] = G_{\alpha,m}$ (note that we can "artificially" set $k = N^{1000}$ instead of $k = N^{20}$ in Theorem 5.5 so that $|\Sigma| = N \leq k^{1/1000}$, losing only a constant factor in the lower bound). Hence, in what follows, we interchangeably use the bijections

$$[\Sigma] \longleftrightarrow G_{\alpha,n} \longleftrightarrow [N] \;\;,\quad [k] \longleftrightarrow G_{\alpha,m} \longleftrightarrow [M], \tag{1}$$

where we shall set the dimension $m \approx 1000 \cdot n$ (to meet the requirement that $M \geq \Sigma^{1000} = N^{1000}$). Notice that in this terminology, Alice's input $\mathbf{x} \in [M]^N$ can be interpreted as the truth table of a function $f : G_{\alpha,n} \mapsto G_{\alpha,m}$, defined by

$$\mathbf{x}_i = f(i) \in G_{\alpha,m} \quad \forall i \; \in G_{\alpha,n}.$$

Similarly, Bob's input $\mathbf{y} \in (\Sigma^M)^N$ encodes the truth table of N functions $g_i : G_{\alpha,m} \mapsto G_{\alpha,n}$, defined by

$$\mathbf{y}_i[j] = g_i(j) \in G_{\alpha,n} \quad \forall \; i \in [N], \;\; j \in G_{\alpha,m}.$$

[14] The proof of [RM99] applies to so called "structured" search problems (canonical search problem associated with DNF tautologies).

Suppose, for the sake of argument, that in Theorem 5.5 we had the further assumption that Bob has the *same* input $\mathbf{y}$ in each of the N coordinates, i.e., that

$$\mathbf{y}_1 = \mathbf{y}_2 = \ldots = \mathbf{y}_N. \tag{2}$$

(This assumption is almost "for free" since each $\mathbf{y}_i$ is an M-dimensional vector, and $M \gg N$ anyway, so one "might as well" consider, in each coordinate i, the single concatenated array $\mathbf{y} := \mathbf{y}_1 \circ \ldots \circ \mathbf{y}_N$. We shall formalize this intuition later.) In this case, Bob's input $\mathbf{y}$ encodes a single function $g(j) := \mathbf{y}[j] \quad \forall\, j \in G_{\alpha,m}$, in which case we observe that

$$\mathbf{y}[\mathbf{x}_i] = g(f(i)) \in G_{\alpha,n}.$$

In other words, if we denote

$$\overline{\mathsf{IND}}^N(\mathbf{x}, \mathbf{y}) := (\mathbf{y}[\mathbf{x}_1], \mathbf{y}[\mathbf{x}_2], \ldots, \mathbf{y}[\mathbf{x}_N])\,,$$

then $\overline{\mathsf{IND}}^N(\mathbf{x}, \mathbf{y})$ encodes (the truth-table of) the composed function $g \circ f$, i.e., *the index gadget can be interpreted as the evaluation of the composed function* $g \circ f$. Since every function $f : G_{\alpha,n} \mapsto G_{\alpha,m}$ is (trivially) $(1/\alpha)$-Lipschitz on $G_{\alpha,n}$ (recall that we are measuring the Lipschitz constant in ℓ_∞ sense) and the same goes for $g : G_{\alpha,m} \mapsto G_{\alpha,n}$, we have the following observation :

Observation 5.6. $\forall\, \alpha, \varepsilon \leq 1$, $\mathbf{AFPC}_{\alpha,(n,1/\alpha),(m,1/\alpha),\varepsilon} \equiv \mathbf{AFP}_{n,\alpha,1/\alpha^2,\varepsilon} \circ \left(\overline{\mathsf{IND}}^N(\mathbf{x}, \mathbf{y})\right)$.

Hence if in Theorem 5.5 we could replace the gadget IND^N with $\overline{\mathsf{IND}}^N$, applying it with the search problem $S := \mathbf{AFP}_{n,\alpha,\lambda,\varepsilon}$ (and the aforementioned parameters) would have directly implied a lower bound on the two-party communication problem $\mathsf{D}^{\mathsf{CC}}\left(\mathbf{AFPC}_{\alpha,(n,1/\alpha),(m,1/\alpha),\varepsilon}\right)$.

Of course, the problem with this "black-box" application of Theorem 5.5 is that it does not guarantee any nontrivial Lipschitz constraints on the input functions f, g to $\mathbf{AFPC}$, which are intrinsic to the $\mathbf{AFPC}$ problem. Indeed, the simulation theorem we shall prove (Lemma 4.2) requires that the input functions f, g to $\mathbf{AFPC}$ satisfy

$$\lambda_1 \lambda_2 \leq \frac{\varepsilon}{43\alpha}\ ,$$

while the decomposition produced by Theorem 5.5 of $h \in S = \mathbf{AFP}_{n,\alpha,\lambda,\varepsilon}$ into f and g ($\mathbf{x}$ and $\mathbf{y}$) may yield arbitrary functions (in which case one can only guarantee that $\lambda_1 \lambda_2 \leq 1/\alpha^2$). Facilitating this further geometric constraint is the main focus of the rest of this section, in which we show how one can adapt the simulation theorem above (Theorem 5.5) to our specific geometric setting, while exploiting the properties of the class of inputs $\mathcal{H}_{\delta,\lambda,n}|_\alpha$. A first important step to facilitate this approach is to replace the Lipschitz conditions on f and g with stronger "local" conditions that imply Lipschitzness, yet are more suitable for the proof of Lemma 4.2. This is the content of the next section.

5.3 Replacing the Lipschitz condition with a local condition

Replacing the Lipschitz condition (on Alice's function f) with an (appropriately chosen) "local" condition has two important benefits:

(1) It will ensure that, on one hand, Alice's set of possible input functions f has a *product structure*, i.e., values to different coordinates can be chosen *independently* of eachother (this property will be important to maintain the simulation invariant of [GPW15]), while f remains Lipschitz ($\lambda_1 = O(1)$). Intuitively, such a property can be achieved by constraining f to map each $x \in G_{\alpha,n}$ into some large enough "local" neighborhood in $G_{\alpha,m}$ (as formalized below).

(2) Our notion of "locality" will actually imply that f is *bi-Lipschitz* (i.e, both f and f^{-1} are Lipschitz). This fact, together with the promise that the composed input function $h := g \circ f \in \mathcal{H}_{\delta,\lambda,n}|_\alpha$ has *local displacements*[15] will imply that g must automatically be "local" as well [16].

We turn to formalize the above intuition. Throughout the rest of the proof we assume without loss of generality $m/n \in \mathbb{N}$. The natural notion of "locality" that implies a bounded Lipschitz condition, is that of *bounded displacements*. Indeed, if $f : G_{\alpha,n} \mapsto [0,1]^m$ and $m = n$, then it is straightforward to see that the condition $\|f(x) - x\|_\infty \le r \ \forall x \in G_{\alpha,n}$ implies (by the triangle inequality) that $\|f(y) - f(x)\| \le ((2r/\alpha) + 1)\|x - y\|_\infty$, $\forall x, y$.

Since $m > n$ in our setup, it is not immediately clear how to define "displacements", as $f(x) - x$ is not a well defined quantity. Instead, we will define an equally useful notion of locality, which requires a certain *projection* of $f(x)$ to the original n-dimensional space, to be "close" to x (in an ℓ_∞ sense). There are many ways to define such projection; we choose the following natural one.

Definition 5.7 (Block partition τ)**.** *Let $m \ge n$ be two integers such that $m/n \in \mathbb{N}$. Let $\tau : [m] \mapsto [n]$ denote the map defined by $\tau(j) = \lceil \frac{n \cdot j}{m} \rceil$. We say that $j \in [m]$ belongs to block B_i if $\tau(j) = i$.*

Let $D \in \mathbb{N}$ be a constant. For every $x \in G_{\alpha,n}$, we denote by

$$L'_x := \{y \in G_{\alpha,m} \ \mid \ \forall\, j \in [m]\ |y_j - x_{\tau(j)}| \le D\alpha\} \tag{3}$$

the *D-local neighborhood* of x induced by the projection τ. Informally speaking, L'_x is the (exponentially large) set of points in $G_{\alpha,m}$ whose $[m]$ coordinates are (point-wise) $D\alpha$-additive approximations of the coordinates of x (with respect to the standard partition τ). Intuitively, the property that $f(x) \in L'_x$ means that $f(x)$ is an (approximate) "duplication" of x, and this will be useful to ensure that f is $\Theta(D)$-Lipschitz. On the other hand, in order to ensure simultaneously that Bob's input function g has a bounded Lipschitz constant , it will be useful that the points in the image (range) of f will also be sufficiently *far apart* from each other, ensuring that f is *bi-Lipschitz.* This can be easily achieved by selecting an (exponentially large) random subset of points inside L'_x, thus the range of $f(x)$ for each x can be thought of as a "code" embedded in some local geometric region (See Figure 4 for illustration). The following definition formalizes the above.

Definition 5.8 ((D, d, ρ)-local functions)**.** *Let $D \ge d \in \mathbb{N}$, and let $\rho < 1$. A family $\mathcal{F}$ of functions $f : G_{\alpha,n} \mapsto G_{\alpha,m}$ is said to be* (D, d, ρ)-local *if there exists a collection of disjoint subsets $\{L_x\}_{x \in G_{\alpha,n}}$, $L_x \subseteq L'_x$ such that, for any $f \in \mathcal{F}$, it holds that $f(x) \in L_x$ for each $x \in G_{\alpha,n}$, $|L_x| \ge 2^{\rho m}$, and*

$$\forall y, y' \in \bigcup_{x \in G_{\alpha,n}} L_x, \ \ \|y - y'\|_\infty \ge d\alpha.$$

The following standard probabilistic argument asserts the existence of a good family of local functions.

Lemma 5.9. *For any $D \ge 2$ and $m \ge 4\log(1/\alpha) \cdot n$, there exists a $(D, D/2, 1/4)$-local family of functions $f : G_{\alpha,n} \mapsto G_{\alpha,m}$.*

Proof. We use a standard probabilistic-method argument to show the existence of the sets L_x. To this end, for each $x \in G_{\alpha,n}$, pick $t = \frac{2^{m/2}}{(1/\alpha)^n}$ random points uniformly and independently from L'_x (for simplicity, let us assume the points are drawn with replacement – by the birthday paradox, the

[15] I.e., $\|h(x) - x\|_\infty \le 5\varepsilon \ \forall x$, see Property (3) in Definition 5.1.

[16] We remark that the analogous claim is not true for the Lipschitz property: if f is a "contraction" ($\lambda_f < 1$), then it is easy to construct a λ-Lipschitz $g \circ f$ where λ_g is arbitrarily large.

probability of drawing the same point twice is negligible hence we ignore this minor issue). Note that for $m \geq 4\log(1/\alpha) \cdot n$, $t \geq 2^{m/4}$ (say), so $|L_x| \geq 2^{m/4}$ for any $x \in G_{\alpha,n}$ as desired.

Fix any two points $y, y' \in \bigcup_{x \in G_{\alpha,n}} L_x$. We have

$$\Pr\left[\|y - y'\|_\infty < D\alpha/2\right] \leq \Pr_{y,y' \in L_x \text{ for some } x}\left[\|y - y'\|_\infty < D\alpha/2\right] < 2^{-m}$$

since in each coordinate $j \in [m]$, the probability that $|y_j - y'_j| < D\alpha/2$ is at most $1/2$ (since the j'th coordinate of y, y' is chosen uniformly and independently from the range $x_{\tau(j)} \pm D\alpha$ by definition of L_x). A union bound over all possible pairs of points in $\bigcup_{x \in G_{\alpha,n}} L_x$ implies that

$$\Pr\left[\exists y, y' \in \bigcup_{x \in G_{\alpha,n}} L_x \;\text{ s.t }\; \|y - y'\|_\infty < D\alpha/2\right] < \left(\sum_x |L_x|\right)^2 \cdot 2^{-m} \leq (t \cdot (1/\alpha)^n)^2 \cdot 2^{-m} = 1.$$

Hence we conclude that there exists a collection of subset $\{L_x\}_{x \in G_{\alpha,n}}$ with all desired properties. $\square$

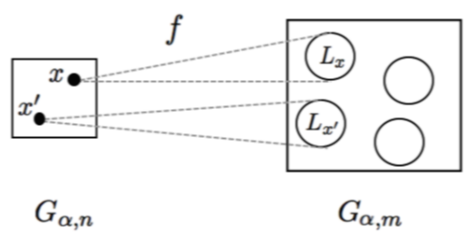

Figure 4: A schematic illustration of a "local" mapping $f : G_{\alpha,n} \mapsto G_{\alpha,m}$. Every function $f \in \mathcal{F}$ maps each point x in the domain $G_{\alpha,n}$ into a point in an exponentially large, disjoint, "local" subset $L_x \subseteq G_{\alpha,m}$. The range of f, $\mathcal{L} = \bigcup_x L_x$, forms a "code" in the sense that the (ℓ_∞) distance between any two points in this set is $\Omega(D\alpha)$.

Convention. Once we've established the existence of the desired sets $\{L_x\}_{x \in G_{\alpha,n}}$, from now we will say, by a slight abuse of notation, that a function $f : G_{\alpha,n} \mapsto G_{\alpha,m}$ is $(D, D/2, 1/4)$-local iff $f(x) \in L_x$. We henceforth use the notation

$$\mathcal{L} := \bigcup_{x \in G_{\alpha,n}} L_x \tag{4}$$

to denote the range of $f \in \mathcal{F}$. In the proof of Lemma 4.2, we shall consider the restricted class of $(D, D/2, 1/4)$-local functions as Alice's possible inputs, that is

$$\mathcal{F} := \{f : G_{\alpha,n} \mapsto G_{\alpha,m} \;\mid\; f(x) \in L_x\}. \tag{5}$$

We now turn to define Bob's input set. We shall use a similar (yet simpler) local condition on the function g, that again will be useful to establish the required Lipschitz condition on g. We start with the following definition.

Definition 5.10 (Locality of g w.r.t $\mathcal{L}$)**.** *We say that a function $g : G_{\alpha,m} \mapsto G_{\alpha,n}$ is* η-local with respect to $\mathcal{L}$ *if*

$$y \in L_x \implies \|g(y) - x\|_\infty \leq \eta. \tag{6}$$

Recall that in Lemma 4.2, the approximation parameter is chosen to be $\varepsilon = \lambda\alpha/1200$. We will restrict Bob's possible inputs functions to the class

$$\mathcal{G} := \{g : G_{\alpha,m} \mapsto G_{\alpha,n} \mid g \text{ is } 5\varepsilon\text{-local with respect to } \mathcal{L}\}. \tag{7}$$

The reason this restriction can be done without loss of generality is that, for any function $h \in \mathcal{H}_{\delta,\lambda,n}|_\alpha$, $\|h(x) - x\|_\infty \leq 5\varepsilon$ (Property (3) in Definition 5.1). Thus, if we denote by $\mathcal{B}_{5\varepsilon}(x) := \{z \in G_{\alpha,n} \mid \|z - x\|_\infty \leq 5\varepsilon\}$ the ℓ_∞-ball of radius 5ε around x, we have $h(x) \in \mathcal{B}_{5\varepsilon}(x)$ for all $x \in G_{\alpha,n}$. The following claim asserts that the local properties of the classes $\mathcal{F}$ and $\mathcal{G}$ imply respective Lipschitz conditions on these functions.

Claim 5.11 (Locality implies Lipschitzness)**.** *If f is $(D, D/2, 1/4)$-local (i.e., $f \in \mathcal{F}$), and g is 5ε-local (i.e., $g \in \mathcal{G}$), then*

- *f is $(2D + 1)$-Lipschitz on $G_{\alpha,n}$.*
- *g is $\left(\frac{21\varepsilon}{D\alpha}\right)$-Lipschitz on $\mathcal{L} = \bigcup_{x \in G_{\alpha,n}} L_x$.*

Proof. **First proposition:** This proposition follows easily from the fact that $L_x \subset L'_x$. Indeed, let $x, x' \in G_{\alpha,n}$ $(x \neq x')$. By the triangle inequality and the assumption that $f(x) \in L_x \subset L'_x$, we have that for any $j \in [m]$,

$$\begin{aligned}
|f(x)_j - f(x')_j| &\leq |f(x)_j - x_{\tau(j)}| + |x_{\tau(j)} - x'_{\tau(j)}| + |f(x')_j - x'_{\tau(j)}| \\
&\leq 2D\alpha + |x_{\tau(j)} - x'_{\tau(j)}| \leq 2D\alpha + \|x - x'\|_\infty \\
&\leq \frac{2D\alpha}{\alpha}\|x - x'\|_\infty + \|x - x'\|_\infty \qquad (\text{since } x, x' \in G_{\alpha,n}) \\
&\leq (2D + 1) \cdot \|x - x'\|_\infty.
\end{aligned}$$

Second proposition: We need to show that $\|g(y) - g(y')\|_\infty \leq (21\varepsilon/D\alpha) \cdot \|y - y'\|_\infty$ for every $y \neq y' \in \mathcal{L}$. Fix $y \in L_x$, $y' \in L_{x'}$ (where possibly $x = x'$). By the triangle inequality and the assumption that g is 5ε-local, we have that

$$\|g(y) - g(y')\|_\infty \leq \|g(y) - x\|_\infty + \|g(y') - x'\|_\infty + \|x - x'\|_\infty \leq 10\varepsilon + \|x - x'\|_\infty.$$

Applying the triangle inequality once again and using the fact that $x, y \in L_x \subset L'_x$ (and similarly $x', y' \in L_{x'} \subset L'_{x'}$), we have

$$\|x - x'\|_\infty \leq \|y - x\|_\infty + \|y' - x'\|_\infty + \|y - y'\|_\infty \leq 2D\alpha + \|y - y'\|_\infty.$$

Combining the last two inequalities gives $\|g(y) - g(y')\|_\infty \leq 10\varepsilon + 2D\alpha + \|y - y'\|_\infty$. Finally, since $\|y - y'\|_\infty \geq D\alpha/2$ for any $y \neq y' \in \mathcal{L}$ (by definition of $\mathcal{L}$), we conclude that

$$\begin{aligned}
\|g(y) - g(y')\|_\infty &\leq \frac{2 \cdot 10\varepsilon}{D\alpha} \cdot \|y - y'\|_\infty + \frac{2 \cdot 2D\alpha}{D\alpha} \cdot \|y - y'\|_\infty + \|y - y'\|_\infty \\
&\leq \left(\frac{20\varepsilon}{D\alpha} + 5\right) \cdot \|y - y'\|_\infty \leq \frac{21\varepsilon}{D\alpha} \cdot \|y - y'\|_\infty
\end{aligned}$$

so long as $5 \leq 21\varepsilon/(D\alpha)$. $\square$

Remark 5.12. *Notice that the class of functions $\mathcal{G}$ only specifies values of $g : G_{\alpha,m} \mapsto G_{\alpha,n}$ on the subset of points $\mathcal{L} \subsetneq G_{\alpha,m}$. To define $g \in \mathcal{G}$ on the entire domain $G_{\alpha,m}$, we use the* MLE-extension *$\bar{g}$ of g from $\mathcal{L}$ to $G_{\alpha,m}$ (see Lemma 3.3). As discussed in Section 3.1, this extension may produce real-valued points in $[0,1]^m$, but rounding these values to the nearest grid-point in $G_{\alpha,n}$ will incur at most an additive factor of 1 in the Lipschitz constant of $g \in \mathcal{G}$ guaranteed by Claim 5.11. Since $\bar{g}$ is determined by g, and since f's range is contained in $\mathcal{L}$, for the purpose of proving Lemma 4.2, we may continue to use Definition 5.10 (while $\mathcal{G}$ is actually the class of all* MLE-extensions *of 5ε-local functions on $\mathcal{L}$).*

5.4 A Streamlined Overview of the Proof of Lemma 4.2

We are now ready to prove Lemma 4.2. Morally speaking, the proof follows by a reduction from the simulation theorem of [GPW15] (Theorem 5.5 above), setting

- $S = \mathbf{AFP}_{n,\alpha,\lambda,\varepsilon}$, $\Sigma = G_{\alpha,n} = [N]$, $\mathcal{X} = \mathcal{F}$, $\mathcal{Y} = \mathcal{G}$.
- $m = \Theta(\log(1/\alpha) \cdot n)$, $k = 2^{m/4}$ (where $k = |L_x|$).

Under these definitions, Alice interprets her input (i.e., her set of indices) $\mathbf{x} = \{\mathbf{x}_i\}_{i \in N} \in [k]^N$ as (the truth table of) a mapping $f \in \mathcal{F}$, such that $f(x) \in L_x$ (note the this is well defined as $\mathcal{F} = \times_{x \in G_{\alpha,n}} L_x$ is a product set and we assume that $k = |L_x|$). Bob interprets his input $\mathbf{y} = \{\mathbf{y}_i\}_{i=1}^N$ as a function $g : \mathcal{L} \mapsto G_{\alpha,n} \in \mathcal{G}$, i.e., $\mathbf{y} \in \times_{x \in G_{\alpha,n}} (\mathcal{B}_{5\varepsilon}(x)^{L_x})$ (recall that $\mathcal{L} := \bigcup_{x \in G_{\alpha,n}} L_x$). In this terminology, any protocol for **AFPC** with the above parameters induces a protocol with the same communication for $\mathbf{AFP}_{n,\alpha,\lambda,\varepsilon} \circ \mathsf{IND}^N$.

However, there are several imprecise details in the above reduction that require some care and prevent us from applying Theorem 5.5 in a black-box fashion:

(1) Theorem 5.5 requires a "fresh" copy $\mathbf{y}_i$ in each coordinate, while in **AFPC**, Bob receives a single function g (represented by a single truth-table $\mathbf{y}$).

(2) In Theorem 5.5, both the set of Alice's inputs (indices) and the alphabet Σ are fixed across all coordinates $i \in [N]$, while in **AFPC**, each coordinate $x \in G_{\alpha,n}$ has its own distinct set of indices L_x (i.e., the range of $f(x)$ is different (in fact, disjoint) for for different x's), and each $x \in G_{\alpha,n}$ has its own set of "colors", since we are using the locality property of $h \in \mathcal{H}_{\delta,\lambda,n}|_\alpha$, which implies that $h(x) \in \mathcal{B}_{5\varepsilon}(x)$.

(3) Lemma 4.2 implicitly uses the promise that the input function $h \in \mathcal{H}_{\delta,\lambda,n}|_\alpha$ (in particular, this assumption is important to ensure that $\forall h \exists f \in \mathcal{F}, g \in \mathcal{G}$ s.t $h = g \circ f$), so the simulation theorem only needs to produce a decision tree that solves $\mathbf{AFP}_{n,\alpha,\lambda,\varepsilon}$ under the restricted class of inputs $\mathcal{H}_{\delta,\lambda,n}|_\alpha$. Theorem 5.5 has no such promise.

(4) [GPW15] remarks that Theorem 5.5 applies to arbitrary alphabets Σ, but provides a formal proof only for the Boolean case $\Sigma = \{0, 1\}$.

While none of the above issues are a major obstacle to the proof, for completeness, we provide the full streamlined proof of [GPW15], adapted to our specific setting. Below we provide a high-level description of the proof and modifications, while the formal proof is deferred to Section B of the Appendix.

Proof outline. The idea is to use a deterministic communication protocol π for $\mathbf{AFPC}(f, g)$ of communication cost C, operating over the input space $\mathcal{F} \times \mathcal{G}$, to produce a (deterministic) decision tree for solving $\mathbf{AFP}(h)$ of cost $O(C/\log k)$, where $g \circ f = h \in \mathcal{H}_{\delta,\lambda,n}|_\alpha$ and $k := |L_x|$ is the size of the "local" neigborhood x is mapped to in $G_{\alpha,m}$. The simulating decision tree proceeds by iterations, where each iteration either "descends" one level in the communication tree of π (by restricting the set of potential inputs h to a smaller set resulting from communicating a bit of π), or descends one level in the decision tree (by querying the value of h on a point $x \in G_{\alpha,n}$). To argue that the simulation is correct, an invariant is maintained ensuring that any leaf of π reached by the decision tree has the correct value of $\mathbf{AFP}(h)$ (i.e., the leaf corresponds to an approximate fixed points of h). To ensure the simulation is efficient, a potential argument is used in the analysis, showing that in each "communication iteration", the potential function (capturing the "relative size" of the remaining input sets compared to its original domain) increases by at most $O(1)$, and in each "query iteration", the potential decreases by at least $\Omega(\log k)$, hence the number of query iterations is at most $O(C/\log k)$ since there are at most C communication iterations.

The kind of iteration that needs to be performed at each step of the simulation is determined by measuring the *predictability* of all unqueried values of h so far, from values of the other unqueried values of h. If no value of h on any unqueried point $x \in G_{\alpha,n}$ is too "predictable", then we can safely perform a communication iteration, restricting the set of inputs to the "bigger" side according to the communicated bit of π. On the other hand, if some value $h(x)$ becomes too predictable, then it is in danger of becoming a fixed function of the remaining unqueried coordinates (which would violate our invariant), and therefore we query $h(x)$ *while it is still possible to accommodate any potential value (in $\mathcal{B}_{5\varepsilon}(x)$) for it* (this is an important place where the properties of the "index gadget" are used in the proofs of [GPW15, RM99], namely, that this function has large monochromatic rectangles). One can sense that the geometric Lipschitz constraints on $\mathcal{F}, \mathcal{G}$ pose challenges on implementing the above approach since the Lipschitz condition of the input functions imposes correlations between neighboring points $x \in G_{\alpha,n}$. Luckily, $\mathcal{F}$ and $\mathcal{G}$ were constructed so that *any value* in some large enough "local" neighborhood of a point x are indeed possible inputs for the players. Indeed, the product structure of the input sets $\mathcal{F}, \mathcal{G}$, and the fact that Alice's set of possible values ("indices") $f(x)$ for each $x \in G_{\alpha,n}$ are *disjoint* by construction ($L_x \cap L_{x'} = \emptyset$), enables maintaining the same invariants required for the [GPW15] proof under a *re-encoded* alphabet. The formal proof of Lemma 4.2 can be found in Section B of the Appendix.

6 Towards the Communication Complexity of Approximate Nash Equilibrium

In this Section we prove Theorem 2.2, showing that the two-party communication complexity of finding a $(2\lambda_1\lambda_2\alpha)$-approximate fixed point of $g \circ f$ is a lower bound on both the two-party and multiparty communication complexity of finding an approximate Nash equilibrium (in distributed two-player and k-player games, respectively). We begin by defining formally the two-party and multiparty approximate Nash equilibrium problems, and then provide the proof of Theorem 2.2 in Section 6.3.

6.1 The two-player setting

A two-player bimatrix game is defined by two payoff matrices $A, B \in \mathbb{R}^{N \times M}$, such that if the row player (Alice) and column player (Bob) choose pure strategies $i \in [N]$, $j \in [M]$, respectively, the player's payoffs are $A(i, j)$ and $B(i, j)$, respectively. We consider $N \times M$ bimatrix games with payoffs in $[-1,1]$ where N and M are polynomially related (hence we can assume that $N = M$ without loss of generality by a "padding" argument).

A *mixed strategy* for a player is a distribution over pure strategies (i.e. rows/columns), and for brevity we may refer to it simply as a strategy. An *ε-approximate Nash equilibrium* (or simply, ε-ANE) is a pair of mixed strategies (x, y) such that

$$\forall i \in [N]\ ,\ \mathbf{e}_i^\top Ay \leq x^\top Ay + \varepsilon, \quad \text{and} \quad \forall j \in [M]\ ,\ x^\top B\mathbf{e}_j \leq x^\top By + \varepsilon. \tag{8}$$

That is, the mixed strategy of each player is at most worse by ε than the (pure) best-response strategy to the opponent's strategy. An *ε-approximate well-supported Nash equilibrium* (or simply, ε-WSNE) is a pair of mixed strategies (x, y) such that

$$\forall i \in [N]\ \forall\ i' \in \mathsf{Supp}(x),\ \mathbf{e}_i^\top Ay \leq \mathbf{e}_{i'}^\top Ay + \varepsilon, \quad \text{and} \tag{9}$$
$$\forall j \in [M]\ \forall\ j' \in \mathsf{Supp}(y),\ x^\top B\mathbf{e}_j \leq x^\top B\mathbf{e}_{j'} + \varepsilon, \tag{10}$$

that is, *every* action played by each player (with nonzero probability) is an ε-best-response action to the opponent's mixed strategy. If $\varepsilon = 0$, the strategy pair (x, y) is called an (exact) *Nash equilibrium* (NE), in which case the two definitions above above coincide.

While the notions of ε-ANE and ε-WSNE are morally equivalent in our communication model (see Lemma 6.6 and Lemma 6.14), the notion of well-supported approximate equilibria will be more natural to work with when we reduce the approximate fixed point problem (**AFPC**) to that of finding an approximate Nash equilibrium, both in the two-party and multiparty communication settings (Section 6.3).

The two-party problem of finding an approximate Nash equilibrium in a two-player game ($\mathbf{ANE}_{K,\varepsilon}$) is defined in Figure 5.

$\mathbf{ANE}_{K,\varepsilon}$
INPUTS : Alice receives an $N \times M$ matrix A with entries $\in [-1, 1]$ (each encoded using $\ell = O(\log\ \max\{N, M\})$ bits), specifying her payoffs for any pair $(i, j) \in [N] \times [M]$ of pure actions. Similarly, Bob receives an $N \times M$ payoff matrix B. Denote by $\mathcal{G}_K = (A, B)$ the two-player game corresponding to the player's inputs, where $K := \max\{M, N\}$ denotes the size of the game. OUTPUT: An ε-ANE of $\mathcal{G}_K$. (Alternatively, two multi-sets $S \subseteq [N], T \subseteq [M]$, such that $(\mathcal{U}(S), \mathcal{U}(T))$ is an ε-ANE of $\mathcal{G}_K$.)

Figure 5: The two-party communication problem of finding an ε-approximate Nash equilibrium in a bimatrix game.

Analogously, denote by $\mathbf{AWNE}_{K,\varepsilon}$ the two-party communication problem in which Alice and Bob need to output an *ε-well-supported* Nash equilibrium (ε-WSNE) of the game $\mathcal{G}_K$ (our reduction below will involve this communication problem as an intermediate step).

Requiring the output of the protocol to be of the form $(\mathcal{U}(S), \mathcal{U}(T))$ where S, T are multi-sets (of $[N]$ and $[M]$ respectively), is essentially without loss of generality. Indeed, if Alice and Bob agree on some arbitrary ε-ANE (resp. ε-WSNE) (μ, ν), then the sub-sampling argument of [LMM03]

(see Lemma 6.2 below) applied to (μ, ν), produces a 2ε-ANE (resp. 2ε-WSNE) of $\mathcal{G}_K$. We favor this output form as: (*i*) it ensures that the output length of the protocol is short (see Lemma 6.2 below), hence excluding a trivial communication lower bound; (*ii*) it makes the problem purely discrete, so we do not have to worry about encoding issues.
We also remind here again, that throughout the paper, we will restrict our attention to games in which M and N are polynomially related (so in particular, $K = \text{poly}(N)$).

We observe that the *non-deterministic* communication complexity of finding an ε-ANE in a bimatrix game is small as long as $\varepsilon \geq (1/\text{poly}\log K)$:

Proposition 6.1. $\mathsf{ND}^{\mathsf{CC}}\left(\mathbf{ANE}_{K,\Omega\left(1/\log^2 K\right)}\right) \leq O\left(\log^6 K\right).$

Proof. We use the following well known lemma:

Lemma 6.2 (Existence of Small-Support Approximate Equilibrium, [LMM03])**.** *Every $K \times K$ bimatrix game with payoffs in $[-1, 1]$ has an ε-ANE* [17] *of the form $(\mathcal{U}(S), \mathcal{U}(T))$, where $S, T \subseteq [K]$ are multi-sets of size $O\left(\frac{\log K}{\varepsilon^2}\right)$, for every $\varepsilon > 0$.*

Applying the lemma for the game $\mathcal{G}_K$ (setting $K := \max\{N, M\}$ and $\varepsilon = 1/O(\log^2 K)$) implies that there exist multi-sets S, T of size $O(\log K/\varepsilon^2) = O(\log^5 K)$ such that $(\mathcal{U}(S), \mathcal{U}(T))$ forms an ε-ANE of $\mathcal{G}_K$. Merlin can specify these sets using $\log K$ bits per element, so the total size of the advice is $O(\log^6 K)$. Alice and Bob can now *privately* verify that $(\mathcal{U}(S), \mathcal{U}(T))$ is indeed an ε-ANE (as the verification of condition (9) only depends on Alice's payoff matrix A, and vice versa for Bob with condition (10)). Therefore, there is an $O(\log^6 K)$ total communication non-deterministic protocol for this problem.

□

6.2 The k-Player Setting

A k-player game $\mathcal{G} = (u_1, u_2, \ldots, u_k)$ over a fixed action set C is defined by k payoff functions $u_i : [C]^k \to \mathbb{R}$ (represented by a matrix $\in \mathbb{R}^{C^k}$), specifying the respective utility of each player $i \in [k]$ for each (pure) action profile $\mathbf{a} := (a_1, a_2, \ldots, a_k) \in [C]^k$. For a mixed action profile $\mathbf{x}$ (i.e., a product distribution over $[C]^k$), let $br_i(\mathbf{x}) := \max_{a_i \in [C]} u_i(a_i, \mathbf{x}_{-i})$ denote the maximum expected utility player i can obtain against the opponents mixed strategy $\mathbf{x}_{-i} \in [C]^{k-1}$, where in this case $u_i(\mathbf{x})$ denotes the *expected* utility of player i over the joint mixed strategy $\mathbf{x}$.

In analogy with condition (8) in the two-player case, an *ε-approximate Nash equilibrium* (ε-ANE) of $\mathcal{G}$ is a mixed strategy profile $\mathbf{x} = (x_1, \ldots, x_k)$, satisfying

$$br_i(\mathbf{x}) \leq u_i(\mathbf{x}) + \varepsilon \quad \forall \;\; i \in [k].$$

Similarly, an *ε-approximate well-supported Nash equilibrium* (ε-WSNE) is a mixed strategy profile satisfying

$$br_i(\mathbf{x}) \leq u_i(a_i, \mathbf{x}_{-i}) + \varepsilon \quad \forall \;\; a_i \in \mathsf{Supp}(x_i) \quad \text{and } \forall \;\; i \in [k],$$

that is, *every* action played by each player i (with nonzero probability) is an ε-best-response action to the opponents' mixed strategy $(\mathbf{x}_{\text{-}i})$. In this paper we consider k-player games with a *constant*

[17] We remark that the sampling argument in [LMM03] actually produces an ε-WSNE (not just an ε-ANE), as the proof guarantees that if x^* is the empirical distribution of $O(\log N/\varepsilon^2)$ samples from the strategy profile x (with replacements), then $\|Ax - Ax^*\|_1 \leq \varepsilon$, which is all that is needed to guarantee the well-supportedness of the subsampled strategy profile.

number of actions C and payoffs in $[-1,1]$.

In analogy with the two-player setting, let k-$\mathbf{ANE}_{C,\varepsilon}$ denote the k-party (Number-In-Hand) communication problem in which each of the k players receives a payoff matrix $u_i \in [-1,1]^{C^k}$ describing her own payoff for each k-tuple of pure actions from a constant action set $[C]$ (so the input size of each player is $|u_i| = O(C^k)$), and the players need to output an ε-ANE of the k-player constant-action game $\mathcal{G}_k := (u_1, u_2, \ldots, u_k)$. Similarly, let k-$\mathbf{AWNE}_{C,\varepsilon}$ denote the communication problem of finding an ε-WSNE of $\mathcal{G}_k$.

We assume that players communicate in the "shared blackboard" communication model [KN97], in which the message of each player is viewable to all k players, though we remark that our results apply to the "message-passing" model as well (in which players have private pair-wise communication channels).

It is noteworthy that the *non-determinisitc* communication complexity of both k-$\mathbf{ANE}_{C,\varepsilon}$ and k-$\mathbf{AWNE}_{C,\varepsilon}$ is only $O_\varepsilon(Ck)$, since this is the number of bits required to describe any joint mixed strategy profile of the k players up to precision ε (In fact, this can be improved to $O_\varepsilon(k \cdot \log C)$ ([BBP14])).

6.3 From Approximate Fixed Points to Approximate Nash

In this section we prove Theorem 2.2, which we restate below.

Theorem 6.3 (From Approximate Fixed Points to Approximate Nash)**.** *For every $m \geq n \in \mathbb{N}$, constants $\alpha \in (0,1)$, $\lambda_1, \lambda_2 \geq 2$, and error parameter ρ,*

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

In particular, Theorem 6.3 implies that a slightly "smoother" version of Theorem 2.1 (in which the Lipschitz constants of f and g satisfy $\lambda_1\lambda_2\alpha \leq \varepsilon$ instead of $\lambda_1\lambda_2\alpha = 86\varepsilon$), would imply strong (deterministic) communication lower bounds on finding approximate Nash equilibria in both two-player and k-player games. This is the content of the following claim.

Proposition 6.4. *Suppose there are constants $\alpha \in (0,1)$, $\lambda_1, \lambda_2 \geq 2$ such that for any n and $m = O(n)$,*

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

Then,

- *For large enough K,* $\quad \mathsf{D^{CC}}\left(\mathbf{ANE}_{K,\Omega(1/\log^2 K)}\right) \geq K^{\Omega_\alpha(1)}$.
- *For large enough k,* $\quad \mathsf{D^{CC}}\left(k\text{-}\mathbf{ANE}_{1/\alpha,3\alpha^3/16}\right) \geq 2^{\Omega_\alpha(k/\log k)}$.

Proof. The first claim follows from the first proposition of Theorem 6.3, observing that whenever $m = O(n)$, $2^{\Omega(n)} = (1/\alpha)^{\Omega(m/\log(1/\alpha))} = K^{\Omega_\alpha(1)}$, since $K = (1/\alpha)^m$. Similarly, the second claim follows from the second proposition of Theorem 6.3, observing that whenever $m = O(n)$,

$$k = O_\alpha(m \log m) = O_\alpha(n \log n) \iff n = \Omega_\alpha(k/\log k),$$

and thus $2^{\Omega(n)} = 2^{\Omega_\alpha(k/\log k)}$. $\square$

In Section 7 we elaborate on the potential modifications required in the proof of our simulation theorem (Lemma 4.2) that would facilitate the stronger ("smoother") lower bound as required in the premise of Proposition 6.4.

We now turn to prove Theorem 6.3. The proof of the first proposition (for the two-player case) is given in the next section (Section 6.3.1). The proof of the second proposition (for the k-player case) is given in the following section (Section 6.3.2).

6.3.1 Proof of the reduction for two-player games

We shall prove the following lemma that proves the first proposition of Theorem 6.3.

Lemma 6.5. *For any $m, n \in \mathbb{N}$ and any constant parameters $\alpha \in (0,1)$, $\lambda_1, \lambda_2 \geq 2$, it holds that*

$$\mathsf{R^{CC}}_\rho\left(\mathbf{ANE}_{(1/\alpha)^m, \Omega(\alpha^4/m^2)}\right) \geq \mathsf{R^{CC}}_\rho\left(\mathbf{AFPC}_{\alpha,(n,\lambda_1),(m,\lambda_2),2\lambda_1\lambda_2\alpha}\right).$$

Indeed, as the size of the game of the **ANE** instance above is $K = (1/\alpha)^m$ and α is assumed to be a universal constant, we have that $m = \Omega_\alpha(\log K)$, hence Lemma 6.5 asserts a communication lower bound on finding an $\Omega(\alpha^4/m^2) = \Omega_\alpha(1/\log^2 K)$ in a K-action bimatrix game.

Our proof can be viewed as a generalization of a reduction used recently in the simpler query-complexity model by Babichenko [Bab14], The proof actually reduces **AFPC** to **AWNE**, i.e., to the two-party problem of finding an approximate *well-supported* equilibrium. This is sufficient due to the following lemma, which guarantees that in any bimatrix game (with payoffs in $[-1,1]$), Alice and Bob can always construct, with no extra communication, an ε-WSNE from an $\Omega(\varepsilon^2)$-ANE, so in particular :

Claim 6.6 (Converting ANE to AWNE, [Bab14] Appendix 6)**.** *For any $\varepsilon \geq 0$,*

$$\mathsf{R^{CC}}_\rho\left(\mathbf{ANE}_{K,\Omega(\varepsilon^2)}\right) \geq \mathsf{R^{CC}}_\rho\left(\mathbf{AWNE}_{K,\varepsilon}\right).$$

With Claim 6.6 in hand, Lemma 6.5 would follow from the following lemma:

Lemma 6.7. *For any $m, n \in \mathbb{N}$ and any constant parameters $\alpha \in (0,1)$, $\lambda_1, \lambda_2 \geq 2$,*

$$\mathsf{D^{CC}}\left(\mathbf{AWNE}_{(1/\alpha)^m, 3\alpha^2/(4m)}\right) \geq \mathsf{D^{CC}}\left(\mathbf{AFPC}_{\alpha,(n,\lambda_1),(m,\lambda_2),2\lambda_1\lambda_2\alpha}\right).$$

Proof. We show how Alice and Bob can (privately) translate their inputs for the fixed-point problem (**AFPC**) to an input for a two-player game (an instance of **AWNE**) with no extra communication. Given the two Lipschitz functions $f : G_{\alpha,n} \mapsto G_{\alpha,m}$ and $g : G_{\alpha,m} \mapsto G_{\alpha,n}$, define the following two-player game. The set of pure actions of Alice is $G_{\alpha,m}$, and the set of pure actions of Bob is $G_{\alpha,n}$ (so the size of the resulting game is $K = \max\{|G_{\alpha,m}|, |G_{\alpha,n}|\} = (1/\alpha)^m$). For any $(x,y) \in G_{\alpha,m} \times G_{\alpha,n}$, Alice's payoff is given by

$$u_1(x,y) = -\frac{1}{m} \cdot \|x - f(y)\|_2^2.$$

Bob's payoff is given by

$$u_2(x,y) = -\frac{1}{n} \cdot \|g(x) - y\|_2^2.$$

In other words, Alice's goal is to try and match the image of Bob's action under her input function f, while Bob should try and match Alice's action under g. Notice that the above payoffs are in the range $[-1,0]$ (as the ℓ_2^2-norm of a vector in $[0,1]^k$ is at most k). Furthermore, Alice and Bob can unilaterally (privately) define these payoff matrices, as Alice's payoff function is determined solely by f and Bob's payoff is determined solely by g, hence this step does not require any communication. Let us denote by $\mathcal{G}_{f,g}$ the resulting two-player game.

Consider a $\frac{3\alpha^2}{4m}$-WSNE (μ, σ) of $\mathcal{G}_{f,g}$. Note that Alice's payoff from any pure action x against Bob's (mixed) strategy σ can be written as

$$u_1(x,\sigma) = -\frac{1}{m} \cdot \mathbb{E}_{y\sim\sigma}\left[\|x - f(y)\|_2^2\right] = -\frac{1}{m}\sum_{i=1}^{m}(x_i - \mathbb{E}_{y\sim\sigma}[f(y)_i])^2 - \frac{1}{m}\sum_{i=1}^{m} Var[\sigma_i]. \tag{11}$$

Let $z = (z_1, z_2, \ldots, z_m) \in G_{\alpha,m}$ be the closest grid point to the point $\mathbb{E}_{y\sim\sigma}[f(y)] := \mathbb{E}[f(\sigma)]$, i.e., $z_i \le \mathbb{E}[f(\sigma)_i] \le z_i + \alpha$, and without loss of generality assume $|\mathbb{E}[f(\sigma)_i] - z_i| \le \alpha/2$ for all $i \in [m]$ (that is, $\mathbb{E}[f(\sigma)_i]$ is closer to z_i than to $z_i + \alpha$). By convexity of the ℓ_2^2 norm, it is clear from equation (11) that Alice's best response to σ is z, in which case the condition $|\mathbb{E}[f(\sigma)_i] - z_i| \le \alpha/2$ ensures that her payoff is at least

$$u_1(z,\sigma) = -\frac{1}{m} \cdot \mathbb{E}_{y\sim\sigma}\left[\|z - f(y)\|_2^2\right] = -\frac{1}{m}\cdot\left(m\cdot\frac{\alpha^2}{4}\right) - \frac{1}{m}\sum_{i=1}^{m} Var[\sigma_i] = -\frac{\alpha^2}{4} - C, \tag{12}$$

where $C := \frac{1}{m}\sum_{i=1}^{m} Var[\sigma_i]$ is a quantity that does not depend on Alice's action. On the other hand, suppose that in μ, with non-zero probability Alice plays a strategy $w \notin C(z)$, where

$$C(z) := \{x \in G_{\alpha,m} :\ x_i \in \{z_i, z_i + \alpha\}\}$$

is the set of nearest lattice points of $G_{\alpha,m}$ to Alice's best response z. Then there is some $i^* \in [M]$ for which $|\mathbb{E}[f(\sigma)_{i^*}] - w_{i^*}| \ge \alpha$, thus

$$u_1(z,\sigma) - u_1(w,\sigma) \ge -\frac{1}{m}\cdot\left(\frac{\alpha^2}{4} - \alpha^2\right) = \frac{3\alpha^2}{4m}, \tag{13}$$

which contradicts the assumption that (μ, σ) is a $\frac{3\alpha^2}{4m}$-WSNE. Hence, Alice's strategy μ must be entirely supported on $C(z)$. Denoting

$$w = (w_1, w_2, \ldots, w_n) \in G_{\alpha,n}$$

the closest lattice point to the point $\mathbb{E}_{x\sim\mu}[g(x)]$ (and assuming once again without loss of generality $|\mathbb{E}_{x\sim\mu}[g(x)_i] - w_i| \le \alpha/2$ for all $i \in [n]$), a symmetric argument asserts that in every $\frac{3\alpha^2}{4m}$-WSNE, Bob's strategy σ is entirely supported on $C(w)$, where $C(w) := \{y :\ \ y_i \in \{w_i, w_i + \alpha\}\}$. We therefore have the following.

Corollary 6.8. *Let $C(z), C(w)$ be the sub-cubes defined above. In every $\frac{3\alpha^2}{4m}$-WSNE (μ, σ) of $\mathcal{G}_{f,g}$, Alice's strategy μ is entirely supported on $C(z)$, and Bob's strategy σ is entirely supported on $C(w)$.*

We now argue that the point w ($\in G_{\alpha,n}$) is an approximate fixed point of $g \circ f$. To this end, recall our shorthands

$$\mathbb{E}_{y\sim\sigma}[f(y)] := \mathbb{E}[f(\sigma)] \quad \text{and} \quad \mathbb{E}_{x\sim\mu}[g(x)] := \mathbb{E}[g(\mu)].$$

We first claim that

$$\|\mathbb{E}[g(\mu)] - g(z)\|_\infty \le \lambda_2\alpha \quad \text{and} \quad \|\mathbb{E}[f(\sigma)] - f(w)\|_\infty \le \lambda_1\alpha. \tag{14}$$

To see why the first inequality holds, observe that for any $x \in C(z)$,

$$\|g(x) - g(z)\|_\infty \le \lambda_2\|x - z\|_\infty \le \lambda_2\alpha$$

since g is λ_2-Lipschitz, and by definition of $C(z)$. By Corollary 6.8, the same statement holds for any x in the support of μ, and thus in particular $\|\mathbb{E}[g(\mu)] - g(z)\|_\infty \leq \lambda_2\alpha$. An analogous argument asserts that $\|\mathbb{E}[f(\sigma)] - f(w)\|_\infty \leq \lambda_1\alpha$.

By the triangle inequality and since we assume $\lambda_1, \lambda_2 \geq 2$, we can now write

$$\|w - g(z)\|_\infty \leq \|w - \mathbb{E}[g(\mu)]\|_\infty + \|\mathbb{E}[g(\mu)] - g(z)\|_\infty \leq \frac{\alpha}{2} + \lambda_2\alpha \leq \frac{3}{2}\lambda_2\alpha, \tag{15}$$

where the second before last transition follows from (14) and the definition of w. Similarly,

$$\|z - f(w)\|_\infty \leq \|z - \mathbb{E}[f(\sigma)]\|_\infty + \|\mathbb{E}[f(\sigma)] - f(w)\|_\infty \leq \frac{\alpha}{2} + \lambda_1\alpha, \tag{16}$$

where the second before last transition follows from (14) and the definition of z.
We conclude that

$$\begin{aligned}
&\|w - g(f(w))\|_\infty \\
&\leq \|w - g(z)\|_\infty + \|g(z) - g(f(w)\|_\infty && \text{(by the triangle inequality)} \\
&\leq \frac{3}{2}\lambda_2\alpha + \lambda_2\|z - f(w)\|_\infty && \text{(by (15) and since } g \text{ is } \lambda\text{-Lipschitz)} \\
&\leq \frac{3}{2}\lambda_2\alpha + \lambda_2 \cdot (\lambda_1\alpha + \alpha/2) && \text{(by (16))}. \\
&\leq 2\lambda_2\alpha + \lambda_2\lambda_1\alpha \\
&\leq 2\lambda_1\lambda_2\alpha,
\end{aligned}$$

since we assumed $\lambda_1, \lambda_2 \geq 2$. Thus, w is a $(2\lambda_1\lambda_2\alpha)$-fixed point of $g \circ f$, as desired.

□

Lemma 6.5 now follows directly from Lemma 6.6 (applied with $\varepsilon := 3\alpha^2/4m$).

Remark 6.9. *The convexity argument below equation* (11)*, namely that expectation is the* unique *minimizer of the ℓ_2^2 norm, is crucial to our argument and is the primary reason for switching to the ℓ_2 norm in the payoff definition. It is also the reason we only manage to prove our result for sub constant values of ε: The normalization of payoffs in $\mathcal{G}_{f,g}$ by a factor of $\Theta(n)$ (to ensure payoffs are $\in [-1, 1]$) rescales the (additive) contribution of each coordinate by $\approx 1/n$, consequently forcing us to consider small deviation equilibria ($\varepsilon \approx 1/n \approx 1/\log K$); This argument can be viewed as the price we pay for the transition from $\|\cdot\|_2 \to \|\cdot\|_\infty$, which seems inevitable in the above reduction.*

6.3.2 Proof of the reduction for k-player constant-action games

We shall prove the following lemma that proves the second proposition of Theorem 6.3.

Lemma 6.10 (From Two-Party **AFPC** to Multiparty **ANE**)**.** *For any $m \geq n \in \mathbb{N}$, any constants $\alpha \in (0, 1)$, λ_1, λ_2, and any error parameter ρ, it holds that*

$$\mathsf{R^{CC}}_\rho \left(O_\alpha(m \log m)\text{-}\mathbf{ANE}_{1/\alpha, 3\alpha^3/16}\right) \geq \mathsf{R^{CC}}_\rho \left(\mathbf{AFPC}_{\alpha,(n,\lambda_1),(m,\lambda_2),2\lambda_1\lambda_2\alpha}\right).$$

The proof has three stages. The first and main stage is reducing the multiparty approximate *well-supported* Nash problem (k-**AWNE**) to a *multiparty* variant of **AFPC**, in analogy with the two-player setting. In the second stage we observe that the multiparty communication complexity of the latter k-party fixed-point problem (**MAFPC**) is at least that of the *two-party* **AFPC** problem. Finally, we use a recent lemma due to Chen et. al [CCT15] that, once again, allows us to convert (with some small loss in the parameters) any k-player ANE to a somewhat weaker WSNE, completing the entire reduction.

We turn to implement the above program. To this end, let us define the following multi-party variant of the approximate fixed point problem, which we henceforth denote by (n,m)-$\mathbf{MAFPC}_{\alpha,\lambda_1,\lambda_2,\varepsilon}$:

(n,m)-$\mathbf{MAFPC}_{\alpha,\lambda_1,\lambda_2,\varepsilon}$
INPUTS : Let $\alpha \in (0,1)$, $m \geq n$, $\lambda_1, \lambda_2 \geq 0$, and $\varepsilon \geq 0$ be publicly known parameters. There are $k = m+n$ players who are divided into two groups, $A := [n]$, and $B = [m]$ respectively. • Each player $i \in [n]$ receives a (truth table of a) λ_2-Lipschitz *scalar* function $g_i : G_{\alpha,m} \mapsto G_{\alpha,1}$ (i.e., a function mapping the m-dimensional α-grid to $[0,1]$). • Each player $j \in [m]$ receives a (truth table of a) λ_1-Lipschitz function *scalar* $f_j : G_{\alpha,n} \mapsto G_{\alpha,1}$ (i.e., a function mapping the n-dimensional α-grid to $[0,1]$). OUTPUT: Let $f : G_{\alpha,n} \mapsto G_{\alpha,m}$ be defined as $f(x) = (f_1(x), \ldots, f_m(x))$, and $g : G_{\alpha,m} \mapsto G_{\alpha,n}$ be defined as $g(y) = (g_1(y), \ldots, g_n(y))$. The players need to output $x \in G_{\alpha,n}$ such that $\|g(f(x)) - x\|_\infty \leq \varepsilon$, i.e., an ε-fixed point of $g \circ f$.

Figure 6: The $(n+m)$-party communication problem of finding an approximate fixed point of $g \circ f$.

The following simple simulation argument asserts that the multiparty communication complexity of the approximate fixed-point problem above, is lower bounded by the *two party* communication complexity of **AFPC**:

Proposition 6.11. *For any error parameter $\rho \geq 0$,*

$$\mathsf{R^{CC}}_\rho\left((n,m)\text{-}\mathbf{MAFPC}_{\alpha,\lambda_1,\lambda_2,\varepsilon}\right) \geq \mathsf{R^{CC}}_\rho\left(\mathbf{AFPC}_{\alpha,(n,\lambda_1),(m,\lambda_2),\varepsilon}\right)$$

Proof. Let τ be a multiparty communication protocol for solving (n,m)-$\mathbf{MAFPC}_{\alpha,\lambda_1,\lambda_2,\varepsilon}$ (either in the shared-blackboard model or in the message-passing model). Let f, g be, respectively, the inputs of Alice and Bob in the two-party communication problem of **AFPC**. The proof follows from the standard observation that Alice and Bob can simulate the k-party protocol τ: Given her input $f : G_{\alpha,n} \mapsto G_{\alpha,m}$, Alice can simulate all the players $j \in [m]$ in group B, and similarly Bob can simulate all the players $i \in [n]$ in group A (notice that the assumption that f (resp. g) is λ_1 (λ_2) Lipschitz, implies that *each* $f_j : G_{\alpha,n} \mapsto G_{\alpha,1}$ ($g_i : G_{\alpha,m} \mapsto G_{\alpha,1}$) is λ_1 (λ_2) Lipschitz, since distances are measured in the ℓ_∞ norm). Let π be the two-party protocol obtained by having Alice and Bob simulate the protocol τ. By assumption, τ finds an ε-fixed point of $g \circ f$ (w.p $1-\rho$), and clearly, the communication going between players in group A and group B is at most the communication of τ, therefore the communication cost of π is at most that of τ. $\square$

In analogy with Lemma 6.7 in the two-party setting, we now argue that the multiparty communication complexity of finding an ε-WSNE in a constant-action multiplayer game (this time for a *constant* value of $\varepsilon = \Omega_\alpha(1)$) is at least as large as the communication complexity of the multiparty approximate fixed-point problem:

Lemma 6.12. $\mathsf{R^{CC}}_\rho\left((n+m)\text{-}\mathbf{AWNE}_{1/\alpha,3\alpha^2/4}\right) \geq \mathsf{R^{CC}}_\rho\left((n,m)\text{-}\mathbf{MAFPC}_{\alpha,\lambda_1,\lambda_2,2\lambda_1\lambda_2\alpha}\right)$.

Proof. The reduction is similar to that of Theorem 6.7, namely, we show how the $k = m+n$ parties can (privately) construct, with no extra communication, a k-player game in which every approximate WSNE induces an $\varepsilon = (2\lambda_1\lambda_2\alpha)$-fixed point of the input function $g \circ f$ to the k-party **AFPC** problem. To this end, the players use their input functions to define (with no extra communication) the following k-player game:

- The action set of each player is the set of points $G_{\alpha,1} = [1/\alpha]$. Let $\mathbf{a} = (a_1, \ldots, a_n) \in G_{\alpha,n}$ denote a joint pure action profile of players in group $A = [n]$, and $\mathbf{b} = (b_1, \ldots, b_m) \in G_{\alpha,m}$ denote a joint pure action profile of players in group $B = [m]$.
- Each player $i \in [n]$ in group A has the payoff
$$u_i(\mathbf{a}, \mathbf{b}) := -|g_i(\mathbf{b}) - a_i|^2.$$
- Each player $j \in [m]$ in group B has the payoff
$$u_j(\mathbf{a}, \mathbf{b}) := -|b_j - f_j(\mathbf{a})|^2.$$

Denote the resulting game by $\mathcal{G}$. Notice that each player's payoff in $\mathcal{G}$ is in the range $[-1, 0]$ as desired. The following analysis is very similar to that of theorem 6.7, only this time the arguments are "per-coordinate".

Consider any joint (mixed) strategy profile $(\mathbf{x}, \mathbf{y}) := (x_1, \ldots, x_n, y_1, \ldots, y_m)$ which forms a $(3\alpha^2/4)$-WSNE of $\mathcal{G}$. Let
$$z = (z_1, z_2, \ldots, z_m) \in G_{\alpha,m}$$
be the closest grid point to the point $\mathbb{E}_{\mathbf{a}\sim\mathbf{x}}[f(\mathbf{a})] := \mathbb{E}[f(\mathbf{x})]$, i.e., $z_j \leq \mathbb{E}[f(\mathbf{x})_j] \leq z_j + \alpha$, and without loss of generality assume $|\mathbb{E}[f(\mathbf{x})_j] - z_j| \leq \alpha/2$ for all $j \in [m]$ (that is, $\mathbb{E}[f(\mathbf{x})_j]$ is closer to z_j than to $z_j + \alpha$). Similarly, let let
$$w = (w_1, w_2, \ldots, w_n) \in G_{\alpha,n}$$
be the closest grid point to the point $\mathbb{E}_{\mathbf{b}\sim\mathbf{y}}[g(\mathbf{b})] := \mathbb{E}[g(\mathbf{y})]$, i.e., $w_i \leq \mathbb{E}[g(\mathbf{y})_i] \leq w_i + \alpha$, and without loss of generality assume $|\mathbb{E}[g(\mathbf{y})_i] - w_i| \leq \alpha/2$ for all $i \in [m]$.

The same argument as in Corollary 6.8 ensures that in every $(3\alpha^2/4)$-WSNE of $\mathcal{G}$, each $\mathbf{y}_j$ is entirely supported on $\{z_j, z_j + \alpha\}$, and each $\mathbf{x}_i$ is entirely supported on $\{w_i, w_i + \alpha\}$. Repeating the same analysis as in equations (14),(15) and (16) for each individual coordinate $i \in [n]$ ($j \in [m]$ respectively), and using the assumption that the f_j's are λ_1-Lipschitz and that the g_i's are λ_2-Lipschitz, we conclude as in Theorem 6.7, that the point $w = (w_1, w_2, \ldots, w_n)$ satisfies
$$\|g(f(w)) - w\|_\infty \leq 2\lambda_2\alpha + \lambda_1\lambda_2\alpha \leq 2\lambda_1\lambda_2\alpha,$$
as desired.

$\square$

To complete the proof of Lemma 6.10, we use a recent result due to Chen et. al [CCT15], which asserts that, for any constant-action k-player game, an ε-ANE can be converted to an $O(\varepsilon)$-WSNE, at the expense of a slight blowup in the size of the game (the number of players):

Claim 6.13 (essentially [CCT15]). *For any $k \in \mathbb{N}, C, \varepsilon \geq 0$,*

$$\mathsf{R^{CC}}_\rho\left(k'\text{-}\mathbf{ANE}_{C,\varepsilon}\right) \geq \mathsf{R^{CC}}_\rho\left(k\text{-}\mathbf{AWNE}_{C,4C\varepsilon}\right),$$

where $k' = 2C^2 k \ln(k/\varepsilon)$.

Proof. Let $\mathcal{G}_k$ denote the k-player game instance of the **AWNE** problem. The following lemma asserts that the k players can use their inputs to define a slightly larger game $\mathcal{G}'_{k'}$, in which every ε-ANE can be translated with no further communication, to an $O(\varepsilon)$-WSNE of $\mathcal{G}_k$:

Lemma 6.14 (From ANE to AWNE, [CCT15], Section 3). *Let $\mathcal{G}_k = (u_1, u_2, \ldots, u_k)$ be a k-player game with a constant number of actions C. Then each player $i \in [k]$ can use her own input (payoff function) u_i to define additional utility functions (over the same action set $[C]$) of $s = 2C^2\ln(k/\varepsilon)$ additional "dummy" players $(u_{i,j},\quad j \in [s])$, so that any ε-ANE of the resulting k'-player game $g'_{k'}$ (where $k' = ks = 2C^2 k\ln(k/\varepsilon)$) can be converted, with no extra communication, into a $(4C\varepsilon)$-WSNE of the original game $\mathcal{G}_k$.*

Claim 6.13 now follows by observing that the k players in $\mathcal{G}_k$ can simulate any k'-party protocol τ for finding an ε-ANE in $\mathcal{G}'_{k'}$, by having each player $i \in [k]$ simulate his own group of "dummy" players.

□

Proof of Theorem 6.10. Let $m \geq n \in \mathbb{N}$, and let $\alpha \in (0,1)$, λ_1, λ_2 be absolute constants. Theorem 6.12, Proposition 6.11 and Theorem 6.13 (applied with $k = m+n, C = 1/\alpha, \varepsilon = 4\alpha^3/16$) together imply that

$$\begin{aligned}&\mathsf{R^{CC}}_\rho\left(O_\alpha(m\log m)\text{-}\mathbf{ANE}_{1/\alpha,3\alpha^3/16}\right) \geq \mathsf{R^{CC}}_\rho\left((n+m)\text{-}\mathbf{AWNE}_{1/\alpha,3\alpha^2/4}\right)\\ &\geq \mathsf{R^{CC}}_\rho\left((n,m)\text{-}\mathbf{MAFPC}_{\alpha,\lambda_1,\lambda_2,2\lambda_1\lambda_2\alpha}\right) \geq \mathsf{R^{CC}}_\rho\left(\mathbf{AFPC}_{\alpha,(n,\lambda_1),(m,\lambda_2),2\lambda_1\lambda_2\alpha}\right),\end{aligned}$$

where the first transition follows from Theorem 6.13 by observing that for the above choice of parameters,

$$k' = 2C^2 k\ln(k/\varepsilon) = 2(1/\alpha)^2(n+m)\ln\left((n+m)/(3\alpha^2/4)\right) = O_\alpha(m\log m)$$

since $m \geq n$ by assumption. □

7 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 $\varepsilon = 2\lambda_1\lambda_2\alpha$, 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 4.2 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 4.2 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 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[18] 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 (**AFP**) was recently shown to be exponential even in the *randomized* query model ([Bab14]), so a *randomized* analogue of our simulation theorem (Lemma 4.2) would have implied an exponential *randomized* communication lower bound for **AFPC**. While the Raz-McKenzie simulation theorem and our adapted geometric variant of it (Lemma 4.2) 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 2.1) 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 (in Section 6.3.2).

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.

[18] 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.

[BBP14] Yakov Babichenko, Siddharth Barman, and Ron Peretz. Simple approximate equilibria in large games. In *ACM Conference on Economics and Computation, EC '14, Stanford , CA, USA, June 8-12, 2014*, pages 753–770, 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.

[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.

[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.

[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.

[Shm12] 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/`, 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.

A Lipschitz extensions: Proof of Lemma 3.3

The lemma follows from the following known Lipschitz extension lemma for scalar functions.

Lemma A.1 (Minimal Lipschitz extension (MLE), special case [Whi33])**.** *Let $A \subset \mathbb{R}^n$ be a non-empty set. If $f : A \mapsto \mathbb{R}$ is λ-Lipschitz on A, then the extended function*

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

is λ-Lipschitz on $\mathbb{R}^n$. The function $\bar{f}$ is called the MLE-extension (minimal Lipschitz extension) of f.

It can directly be verified that $\bar{f}$ is indeed an extension of f, i.e., $\bar{f}|_A \equiv f$. To prove Lemma 3.3 for a function $f : \mathbb{R}^n \mapsto \mathbb{R}^m$, observe that the condition $\|\bar{f}(x) - \bar{f}(y)\|_\infty \leq \lambda \|x - y\|_\infty$ is satisfied iff for every coordinate $i \in n$, $|\bar{f}_i(x) - \bar{f}_i(y)| \leq \lambda \|x - y\|_\infty$, the statement follows from applying Lemma A.1 separately to each coordinate of $\bar{f}$.

B Proof of Lemma 4.2

In this section we provide the formal proof of Lemma 4.2, following the proof outline of [GPW15]. We begin by setting up some notation and definitions that will facilitate the proof.

B.0.3 Notation and main lemmas

Throughout the proof, we continue to denote $[N] = G_{\alpha,n}$ and $[M] = G_{\alpha,m}$, where

$$m = 4000 \log(1/\alpha) \cdot n := c_\alpha \cdot n.$$

Alice's set of possible input functions ($\mathcal{F}$) is the (restricted) set of all $(D, D/2, 1/4)$-local functions whose image is contained in $\mathcal{L} = \bigcup_{x \in G_{\alpha,n}} L_x$:

$$\mathcal{F} := \{f : G_{\alpha,n} \mapsto G_{\alpha,m} \ \mid \ f(x) \in L_x\}.$$

Simply written, $\mathcal{F} = \times_{x \in G_{\alpha,n}} L_x$. Bob's input set is the set of 5ε-local functions (w.r.t $\mathcal{L}$), that is

$$\mathcal{G} := \{g : \mathcal{L} \mapsto G_{\alpha,n} \ \mid \ y \in L_x \implies g(y) \in \mathcal{B}_{5\varepsilon}(x)\},$$

were we recall that $\mathcal{B}_{5\varepsilon}(x)$ is the ℓ_∞-ball of radius 5ε around x (notice that the above definition is indeed well defined, since for every $y \in \mathcal{L}$, there is a unique x s.t $y \in L_x$, by disjointness of the sets L_x). Recalling that $k := |L_x|$, we can simply write $\mathcal{G} = \times_{x \in G_{\alpha,n}} (\mathcal{B}_{5\varepsilon}(x))^k$. We assume (w.l.o.g) that for every $x \in G_{\alpha,n}$, $|\mathcal{B}_{5\varepsilon}(x)|$ is the same across all x's, and we henceforth denote this size by r, so r is the number of "colors" each element in $\mathcal{L}$ can obtain. We caution that, unlike the proofs of [RM99, GPW15], in our setting each coordinate $x \in G_{\alpha,n}$ has its *own* set of r possible colors. Note that we trivially have that $r \leq |G_{\alpha,n}| = N$ (since $\mathcal{B}_{5\varepsilon}(x) \subseteq G_{\alpha,n}$). In this terminology, initially $|\mathcal{G}| = \prod_{x \in G_{\alpha,n}} |\mathcal{B}_{5\varepsilon}(x)|^k = r^{kN}$. Recall that by definition of locality of $\mathcal{F}$, we have $|\mathcal{F}| = \prod_{x \in G_{\alpha,n}} |L_x| \geq 2^{mN/4}$, and in particular

$$k = |L_x| \geq 2^{m/4} > N^{1000} \geq r^{1000} \tag{17}$$

by choice of m and definition of r. For a subset of points $I \subseteq G_{\alpha,n}$, denote

$$\mathcal{L}_I := \bigcup_{x \in I} L_x$$

the range of all points in I under f.

Let π be a deterministic communication protocol for **AFPC**. For each node v of the protocol tree π, let $R_v := \mathcal{F}^v \times \mathcal{G}^v$ denote the rectangle associated with the node v, and let $\mathcal{F}^{v,b} \subseteq \mathcal{F}^v$ denote the set of Alice's inputs on which the bit $b \in \{0, 1\}$ would be sent if Alice is the speaker at node v, and similarly let $\mathcal{G}^{v,b} \subseteq \mathcal{G}^v$ denote the set of Bob's inputs on which the bit $b \in \{0, 1\}$ would be sent if Bob is the speaker at node v.

Let $A \subseteq \mathcal{F}$ and $B \subseteq \mathcal{G}$. The following definitions will be central to the proof.

- **Size of sets :** For $|A|, |B| > 0$, let $a(A)$ be such that $|A| = 2^{-a(A)} \cdot |\mathcal{F}| = 2^{-a(A)} \cdot k^N$, and $b(B)$ be such that $|B| = b(B) \cdot (r^k)^N$.

- **Projections :** For a subset of points $I \subseteq G_{\alpha,n}$, let

$$A_I := \{\{f(x)\}_{x\in I} \mid \exists \{f(x')\}_{x'\in G_{\alpha,n}\setminus I} \ s.t \ \{f(x)\}_{x\in I} \cup \{f(x')\}_{x'\in G_{\alpha,n}\setminus I} \in A\} \quad \subseteq [M]^{|I|}$$

denote the projection of A onto the subset I. Accordingly, let

$$B_{\mathcal{L}_I} := \{\{g(y)\}_{y\in\mathcal{L}_I} : \exists \{g(y')\}_{y'\in\mathcal{L}\setminus\mathcal{L}_I} \ s.t \ \{g(y)\}_{y\in\mathcal{L}_I} \cup \{g(y')\}_{y'\in\mathcal{L}\setminus\mathcal{L}_I} \in B\} \quad \subseteq [N]^{|\mathcal{L}_I|}$$

denote the projection of B onto the subset of points $\mathcal{L}_I \subseteq [M] = G_{\alpha,m}$.

- **Pruning :** Let $x \in G_{\alpha,n}$. For a subset $U \subseteq L_x$ of possible f-values in the "local neighborhood" of x, define $A^{x,U} := \{f \in A \mid f(x) \in U\}$, and for a subset $V \subseteq (\mathcal{B}_{5\varepsilon}(x))^k$ of possible g-values on the set L_x, define $B^{x,V} := \{g \in B \mid g(L_x) \in V\}$ (where the shorthand $g(L_x) \in V$ means that there is some $v \in V$ such that $g(y) = v_y$ for every $y \in L_x$).

- **Auxiliary Graph :** For every $x \in G_{\alpha,n}$, let $\mathsf{Graph}_x(A)$ be the bipartite graph defined as follows: The set of left nodes is L_x, the set of right nodes is $\mathcal{L}_{-x} := \times_{x'\neq x} L_{x'}$, and each tuple of values $\{f(x)\}_{x\in G_{\alpha,n}} \in A$ is viewed as an edge between the left node $f(x)$ and the right node $\{f(x')\}_{x'\in\neq x}$. Note that $A_{G_{\alpha,n}\setminus\{x\}}$ is the set of all right nodes of $\mathsf{Graph}_x(A)$ with non-zero degree.

- **Average/Minimum degree :** Let $\mathsf{AvgDeg}_x(A) := |A|/|A_{G_{\alpha,n}\setminus\{x\}}|$ and $\mathsf{MinDeg}_x(A)$ be, respectively, the average and minimum degree of the non-zero degree right nodes of $\mathsf{Graph}_x(A)$. Intuitively, these quantities measure how "predictable" the value $f(x)$ is from $\{f(x')\}_{x'\in G_{\alpha,n}\setminus x}$.

- **Thickness :** We say that A is *thick* iff $\mathsf{MinDeg}_x(A) \geq k^{17/20}$ for all $x \in G_{\alpha,n}$.

Remark B.1. *All the above definitions apply also to projected subsets $A \subseteq \mathcal{F}_I$, $B \subseteq \mathcal{G}_{\mathcal{L}_I}$ for a subset $I \subseteq G_{\alpha,n}$, with the parameter N replaced by $|I|$ and $x \in G_{\alpha,n}$ replaced with $x \in I$. For example, $a(A)$ is now adjusted so that $|A| = 2^{-a(A)} \cdot k^{|I|}$, $|B| = b(B) \cdot (r^k)^{|I|}$ and $\mathsf{Graph}_x(A)$ is now defined for any $x \in I$ (instead of for every $x \in G_{\alpha,n}$). The proof shall make a recursive use of these extended definitions.*

The following lemma will be useful for the case in which we need π to communicate a bit.

Lemma B.2 (Thickness Lemma, [GPW15])**.** *Let $I \subseteq G_{\alpha,n}$. If $A \subseteq \mathcal{F}_I$ is such that* $\mathsf{AvgDeg}_x(A) \geq d$ *for all $x \in I$, then there exists an $A' \subseteq A$ such that:*

- $\mathsf{MinDeg}_x(A') \geq \frac{d}{2|I|}$ *for all $x \in I$.*

- $a(A') \leq a(A) + 1$.

The (short) proof of the Thickness Lemma is precisely the same as that of [GPW15] (see Lemma 6 and the proof in Section 3.5), hence we omit it. The following lemma will be useful for the case in which we need to have our decision tree query a value $h(x)$ of the input function to **AFP**.

Lemma B.3 (Projection Lemma, essentially [GPW15])**.** *Let $I \subseteq G_{\alpha,n}$, and suppose $A \subseteq \mathcal{F}_I$ is* thick*, and $B \subseteq \mathcal{G}$ is such that $b(B) \leq k^{2/20}$. Then for any $x \in I$ and every $z \in \mathcal{B}_{5\varepsilon}(x)$, there is a z-monochromatic rectangle*[19] *$U \times V \subseteq L_x \times (\mathcal{B}_{5\varepsilon}(x))^{|L_x|}$ such that:*

[19] That is, $\forall (f, g) \in U \times V$, $g(f(x)) \equiv z$.

- $A^{x,U}_{I\setminus\{x\}}$ *is thick,*

- $a(A^{x,U}_{I\setminus\{x\}}) \leq a(A) - \log k + \log \mathsf{AvgDeg}_x(A)$,

- $b(B^{x,V}_{\mathcal{L}_I\setminus\{L_x\}}) \leq b(B) + 1$.

The above lemma is the main technical lemma in the simulation argument of [GPW15]. We reprove this lemma in our setting in Section B.3.

B.1 The decision tree simulation of π

We now describe the [GPW15] algorithm (adapted to our setting), which shows how to construct a decision tree $\mathcal{T}$ that solves the approximate fixed-point problem **AFP** on an unknown input function $h = \{h(x)\}_{x\in G_{\alpha,n}} \in \mathcal{H}_{\delta,\lambda,n}|_\alpha$. This decision tree is obtained by simulating the execution of the hypothesized protocol π for **AFPC** (under the assumption that $(f,g) \in \mathcal{F}\times\mathcal{G}$). The decision tree $\mathcal{T}$ is described in Algorithm 1.

We briefly describe here the intuition behind the algorithm (for a more detailed overview see Section 3.3 in [GPW15]). On input $h = \{h(x)\}_{x\in G_{\alpha,n}} \in \mathcal{H}_{\delta,\lambda,n}|_\alpha$, the node v traces a root-to-leaf path (of length at most C) of π, which is used to determine which values $h(x)$ to query, and when. We maintain the invariant that every $(f,g) \in A\times B$ is consistent with the query history so far (i.e., $g(f(x)) = h(x)$ for every previously queried x). The interesting structure of $A\times B$ is what they look like on the unqueried points x, i.e., on the projected set $A_I \times B_{\mathcal{L}_I}$: By construction, we maintain the property that all possible values in $(\mathcal{B}_{5\varepsilon}(x))^I$ are still possible for the unqueried points $x \in I$ (in fact, a stronger property is maintained, namely, that A_I is thick, and $B_{\mathcal{L}_I}$ is "large" as measured by $b(B_{\mathcal{L}_I})$). The potential function is $a(A_I)$, namely, the size of the set of all projections of elements of A to the unqueried coordinates in I, relative to the original domain $\mathcal{F} := \times_x L_x$. The type of iteration is determined by $\min_{x\in I}\mathsf{AvgDeg}_x(A_I)$, which captures how much the values of h on the set of unqueried points I are "predictable" from each other in $A\times B$.

In a communication iteration (lines 5 and 11), the current set of inputs is restricted to the "larger" rectangle obtained by (either Alice or Bob) sending a bit $b \in \{0,1\}$. This ensures our potential does not increase too much (note that larger potential corresponds to a smaller set) if Alice is the sender, and that $B_{\mathcal{L}_I}$ stays large enough if Bob is the sender. If Alice is the sender, the restriction may result in violation of the thickness invariant, in which case we employ the Thickness Lemma (Lines 7-9).

In a query iteration, we query a value $h(x)$ whenever $\mathsf{AvgDeg}_x(A_I)$ drops and becomes too small. We can then use the Projection Lemma (Lines 17-21) to restrict $A\times B$ to an $h(x)$-monochromatic sub-rectangle (for *any* value in $\mathcal{B}_{5\varepsilon}(x)$, using the thickness invariant). The fact that $\mathsf{AvgDeg}_x(A_I)$ is small ensures at least an $\Omega(\log k)$ decrease in potential.

B.2 Analysis of the simulation algorithm

The analysis of the simulation tree $\mathcal{T}$ is the same as the one in [GPW15] (with the appropriate reparametrization of the parameters), hence we omit it and refer the reader to Section 3.4 in [GPW15]. To complete the proof of Lemma 4.2, it therefore remains to prove the Projection Lemma. This is the content of the next section.

B.3 Proof of the Projection Lemma

In this section we show how to adapt the [GPW15] proof of the Projection Lemma to our setting of parameters. Essentially all arguments are the same as in the [GPW15] proof, albeit with the

ALGORITHM 1: The simulating decision-tree $\mathcal{T}$ for solving **AFP** (cf. [GPW15]).

input : A truth table $\{h(x)\}_{x \in G_{\alpha,n}}$ of a function $h : G_{\alpha,n} \mapsto G_{\alpha,n} \in \mathcal{H}_{\delta,\lambda,n}|_\alpha$.
output: $x \in G_{\alpha,n}$ such that $\|x - h(x)\|_\infty \leq \varepsilon$ (i.e., a solution to $\mathbf{AFP}_{n,\alpha,\lambda,\varepsilon}(h)$).

Initialize $v =$ root of π, $I = G_{\alpha,n}$, $A = \mathcal{F} = \times_{x \in G_{\alpha,n}} L_x$, $B = \mathcal{G} = \times_{x \in G_{\alpha,n}} (\mathcal{B}_{5\varepsilon}(x))^k$.
while *v is not a leaf of π* **do**
  if $\mathsf{AvgDeg}_x(A) \geq k^{19/20}$ *for all $x \in I$* **then**
    let v_0, v_1 be the children of v ;
    if *Alice sends a bit at v* **then**
      let $b \in \{0,1\}$ be such that $a((A \cap \mathcal{F}^{v,b})_I) \leq a(A_I) + 1$;
      let $A' \subseteq (A \cap \mathcal{F}^{v,b})_I$ be such that:
      (1) A' is thick ;
      (2) $a(A') \leq a((A \cap \mathcal{F}^{v,b})_I) \leq a(A_I) + 1$;
      update $A = \left\{f \in A \cap \mathcal{F}^{v,b} : \ \{f(x)\}_{x \in I} \in A'\right\}$ and $v = v_b$ (so now $A_I = A'$) ;
    else
      (**if** Bob sends a bit at v) let $b \in \{0,1\}$ be such that $b((B \cap \mathcal{G}^{v,b})_{\mathcal{L}_I}) \leq b(B_{\mathcal{L}_I}) + 1$;
      update $B = B \cap \mathcal{G}^{v,b}$ and $v = v_b$;
    end
  else
    if $\mathsf{AvgDeg}_x(A) < k^{19/20}$ for some $x \in I$, **then** query $h(x)$;
    let $U \times V \subseteq L_x \times (\mathcal{B}_{5\varepsilon}(x))^k$ be an $(h(x))$-monochromatic rectangle such that:
    (1) $A^{x,U}_{I \setminus \{x\}}$ is thick ;
    (2) $a(A^{x,U}_{I \setminus \{x\}}) \leq a(A_I) - \log k$;
    (3) $a(B^{x,V}_{\mathcal{L}_I \setminus \{L_x\}}) \leq b(B_{\mathcal{L}_I}) + 1$;
    update $A = A^{x,U}, B = B^{x,V}$ and $I = I \setminus \{x\}$;
  end
  output the same value that v does.
end

re-encoding of the parameters, as defined in the reduction at the beginning of Section 5.4. In places where no modification is required we simply refer to the proofs of [GPW15, RM99] (we remark that the [GPW15] proof is written (for simplicity) for the binary $|\Sigma| = 2$ case while we need the "multi-color" version (arbitrary alphabets), which is explicitly treated in the proof of [RM99], so we shall make an occasional reference to the latter proof below).

Proof of Lemma B.3. Fix $x \in I$, and for simplicity of notation let us henceforth denote $A_{-x} := A_{I \setminus \{x\}}$ and similarly $A^{x,U}_{-x} := A^{x,U}_{I \setminus \{x\}}$. Recall that $A^{x,U}_{-x}$ is the set of right nodes $f_{-x} := \left\{\{f(x')\}_{x' \in I \setminus \{x\}}\right\}$ of $\mathsf{Graph}_x(A)$ that have a neighbor in $U \subseteq L_x$. Similarly, let

$$B^{x,V}_{-x} := B^{x,V}_{\mathcal{L}_I \setminus \{L_x\}} = \left\{\{g(y)\}_{y \in \mathcal{L}_I \setminus \{L_x\}} : \ \{g(y)\}_{y \in \mathcal{L}_I} \in B \text{ for some } g(L_x) \in V\right\}.$$

We claim that if we take a uniformly random subset $U \subseteq_R L_x$ of size $k^{7/20}$ and let $V := \{g \in \mathcal{G} : g(y) \equiv z \ \ \forall \, y \in U\}$, then:

(1) $A^{x,U}_{-x} = A_{-x}$ with probability at least $1 - 2^{-k^{3/20}}$;

(2) A_{-x} is thick ;

(3) $a(A_{-x}) \leq a(A) + \log \mathsf{AvgDeg}_x(A) - \log k$;

(4) $b(B_{-x}^{x,V}) \leq b(B) + 1$ with probability at least $2^{-k^{2/20}}$.

The Projection Lemma then follows from a union bound.

Property (1). For every non-zero degree right node $f_{-x} = \left\{\{f(x')\}_{x' \in I \setminus \{x\}}\right\} \in A_{-x}$ of $\mathsf{Graph}_x(A)$, let $F_{f_{-x}} := \{f(x) \in L_x : f_{-x} \cup \{f(x)\} \in A\}$ denote the set of all left nodes adjacent to it. Since A is thick by assumption, we have $|F_{f_{-x}}| \geq \mathsf{MinDeg}_x(A) \geq k^{17/20}$, and $f_{-x} \in A_{-x}^{x,U}$ iff U intersects $L_{f_{-x}}$. Since $|U| = k^{7/20}$, the probability that U does not intersect $F_{f_{-x}}$ is at most

$$(1 - k^{17/20}/|L_x|)^{k^{7/20}} \leq (1 - k^{17/20}/k)^{k^{7/20}} \leq e^{-k^{4/20}}.$$

Recall that $k \geq N^{1000}$. Since the number of elements $f_{-x} \in A_{-x}$ is at most $\prod_{x' \in I \setminus \{x\}} |L_{x'}| \leq k^{|I|-1} \leq k^{N-1} \leq 2^{k^{1/1000} \cdot \log k}$, by a union bound the probability that one of these elements is not in $A_{-x}^{x,U}$ is at most $2^{k^{1/1000} \cdot \log k} \cdot e^{-k^{4/20}} < 2^{-k^{3/20}}$.

Property (2). The proof of this property is exactly the same as in [GPW15] but we redo it here in the terminology of **AFPC** for the sake of completeness. By definition of thickness, it suffices to show that $\mathsf{MinDeg}_{x'}(A_{-x}) \geq \mathsf{MinDeg}_{x'}(A)$ for all $x' \in I \setminus \{x\}$. Indeed, for every non-zero degree right node $f_{-\{x,x'\}} := \left\{\{f(x'')\}_{x'' \in I \setminus \{x,x'\}}\right\}$ of $\mathsf{Graph}_{x'}(A_{-x})$, there exists a value $f(x')$ such that $f_{-\{x,x'\}} \cup \{f(x')\} \in A_{-x}$. Thus by the definition of A_{-x} there exists a value $f(x) \in L_x$ such that $\left(f_{-\{x,x'\}} \cup \{f(x')\} \cup \{f(x)\}\right) \in A$. Therefore, by the definition of $\mathsf{MinDeg}_{x'}(A)$ applied to the set of non-zero degree right nodes $f_{-x'} := \left\{\{f(x'')\}_{x'' \in I \setminus \{x'\}}\right\}$ of $\mathsf{Graph}_{x'}(A)$, we have that $\left(f_{-x'} \cup \{\hat{f}(x')\}\right) \in A$ holds for at least $\mathsf{MinDeg}_{x'}(A)$ different values $\hat{f}(x')$. All of these values satisfy $\left(f_{-\{x,x'\}} \cup \{\hat{f}(x')\}\right) \in A_{-x}$. Hence the degree of the right node $f_{-\{x,x'\}}$ in $\mathsf{Graph}_{x'}(A_{-x})$ is at least $\mathsf{MinDeg}_{x'}(A)$.

Property (3). Observe that $|A_{-x}| = |A|/\mathsf{AvgDeg}_x(A)$. Since $A_{-x} \subseteq \mathcal{F}_{I \setminus \{x\}}$ and $a(A) = \log(k^{|I|}/|A|)$ by definition, we have

$$a(A_{-x}) = \log(k^{|I|-1}/|A_{-x}|) = \log(k^{|I|-1} \cdot \mathsf{AvgDeg}_x(A)/|A|) =$$

$$= \log(k^{|I|}/|A|) + \log(\mathsf{AvgDeg}_x(A)/k) = a(A) + \log \mathsf{AvgDeg}_x(A) - \log k.$$

Property (4). Recall that the set V is a random variable of U, so we henceforth write $V = V_U$ to avoid confusion. Recall that B_{-x}^{x,V_U} is the set of all possible g-values on $\mathcal{L}_I \setminus \{L_x\}$, for which there is an extension of g to L_x so that $g|_{\mathcal{L}_I} \in B$, and $g(y) \equiv z$ for *every* $y \in U$. We want to show that $|B_{-x}^{x,V_U}|$ shrinks by at most a factor of 2 relative to $|B|$. This is the content of the following claim from [RM99], which we state without a proof:

Claim B.4 ([RM99], Claim 6.5 in Section 6.4)**.** *Suppose that $k \geq r^{1000}$. If $U' \subseteq_R L_x$ is a random subset of size $|U'| = k^{5/20}$, then*

$$\Pr_{U'}\left[|B_{-x}^{x,V_{U'}}| \geq |B|/(2 \cdot r^k)\right] \geq 3/4.$$

We remark that in [RM99], $[r]$ represents a set of absolute “colors” for all of Bob’s inputs (so Bob’s input is in $(r^m)^N$ where the parameter m of [RM99] denotes the length of each “vector”), while, as noted above, in our setting each $x \in G_{\alpha,n}$ has its own set of allowable colors ($|\mathcal{B}_{5\varepsilon}(x)| = r$). We observe that the proof remains exactly the same if we substitute $[m] \leftrightarrow L_x$ (i.e., $m \leftrightarrow k$), and observing that the condition $|\mathcal{B}_{5\varepsilon}(x)| = r \leq k^{1/1000}$ still applies, by (17).

Viewing the random set U (of size $k^{7/20}$) as a collection of $k^{2/20}$ random independent subsets $U'_i \subseteq_R L_x$ each of size $k^{5/20}$, Claim B.4 applied independently to each U'_i ensures that

$$\Pr_U\left[|B_{-x}^{x,V_U}| \geq |B|/(2 \cdot r^k)\right] \geq \left(\Pr_U[|B_{-x}^{x,V_{U'_1}}| \geq |B|/(2 \cdot r^k)]\right)^{k^{2/20}} \geq (3/4)^{k^{2/20}}.$$

This completes the proof of the Projection Lemma.

$\square$

ECCC ISSN 1433-8092

http://eccc.hpi-web.de