Candidate Hard Unique Game*

Subhash Khot † Dana Moshkovitz ‡

November 2, 2015

Abstract

We propose a candidate reduction for ruling out polynomial-time algorithms for unique games, either under plausible complexity assumptions, or unconditionally for Lasserre semidefinite programs with a constant number of rounds. We analyze the completeness and Lasserre solution of our construction, and provide a soundness analysis in a certain setting of interest. Addressing general settings is tightly connected to a question on Gaussian isoperimetry.

Our construction is based on a suggestion in [30] wherein the authors study the complexity of approximately solving a system of linear equations over reals and suggest it as an avenue towards a (positive) resolution of the Unique Games Conjecture. The construction employs a new encoding scheme that we call the *real code*. The real code has two useful properties: like the long code, it has a unique local test, and like the Hadamard code, it has the so-called *sub-code covering* property.

*A preliminary version of this paper without a soundness analysis appeared as ECCC TR14-142.

†khot@cims.nyu.edu. Computer Science Department, Courant Institute of Mathematical Sciences, New York University. Research supported by NSF grants CCF 1422159, 1061938, 0832795 and Simons Collaboration on Algorithms and Geometry grant.

‡dmoshkov@csail.mit.edu. Department of Electrical Engineering and Computer Science, MIT. This material is based upon work supported by the National Science Foundation under grants number 1218547 and 1452302.

1 Introduction

1.1 The Unique Games Conjecture

The Unique Games Conjecture [27] is currently one of the important questions in theoretical computer science. It is a perplexing question in the sense that researchers have no consensus regarding its correctness and a tantalizing question in the sense that its resolution might possibly be on the horizon. As shown in [29], the conjecture can be phrased equivalently in terms of solving a nearly-satisfiable system of *discrete* linear equations, where each equation depends on two variables:

Definition 1. 2Lin($\mathbb{F}$) Problem: *Given N variables $x_1, \ldots, x_N$ taking values over a finite field $\mathbb{F}$ and M equations $C_1, \ldots, C_M$ where each equation C_i is of the form $x_{i_1} - x_{i_2} = b_i$ and $b_i \in \mathbb{F}$. The goal is to find an assignment that maximizes the fraction of equations satisfied.*

Note that in each equation, for any value for either of the two variables, there is a *unique* value for the other variable that satisfies the equation. The Unique Games problem is a bit more general: each constraint is on two variables, the variables take values from an alphabet Σ, and for any value for either of the two variables in a constraint, there is a *unique* value for the other variable that satisfies the constraint (so a constraint corresponds to a permutation on Σ and different constraints may correspond to different permutations). As shown in [29], the essence of the Unique Games problem is captured even when the constraints are linear over a finite field and one may restrict to the 2Lin($\mathbb{F}$) problem. In particular, the Unique Games Conjecture can be stated as:

Definition 2 (The Unique Games Conjecture)**.** *For any constant $\varepsilon > 0$, there is a finite field $\mathbb{F}$ of a constant size, such that given a* 2Lin($\mathbb{F}$) *instance that has an assignment that satisfies $(1 - \varepsilon)$ fraction of all equations, it is NP-hard to find an assignment that satisfies even an ε fraction of all equations.*

If true, the Unique Games Conjecture implies optimal NP-hardness of approximation for a large number of optimization problems and in some cases, for entire classes of problems (see the surveys [28, 47] for more background on the Unique Games Conjecture). For example, Raghavendra [39] shows that, assuming the Unique Games Conjecture (and P$\neq$NP), basic semidefinite programs (SDP) yield optimal approximation algorithms for constraint satisfaction problems and in particular, for the Unique Games problem itself. Indeed, researchers had already designed algorithms for the Unique Games problem based on semidefinite programs [20, 27, 13, 46], as well as constructed matching integrality gaps, showing that these algorithms do not disprove the Unique Games Conjecture [33]. After Raghavendra's work, the integrality gap constructions have been extended to SDPs that are more general, amounting to a combination of a basic SDP and a super-constant number of rounds of the so-called Sherali-Adams linear programming relaxation [32, 40].

More recently, researchers have been looking at semidefinite programs that are even more general. A promising approach is to consider the Lasserre hierarchy of semidefinite programs [6, 22]. In contrast to all the semidefinite programs considered before, current techniques seem inadequate to construct integrality gaps against the Lasserre hierarchy. Some of the limitations of current techniques towards constructing Lasserre integrality gaps have been formalized in [3]. The latter work shows that integrality gaps against weaker SDPs can be solved using only a *constant* number of rounds of the Lasserre hierarchy.

1.2 A Weak Unique Games Conjecture

In this paper we focus on a weak version of the Unique Games Conjecture that rules out polynomial-time algorithms for unique games, either under plausible complexity assumptions, or unconditionally for Lasserre semidefinite programs with a constant number of rounds. The weak UGC focuses on a special case of unique games where the underlying field is boolean.

Definition 3. Boolean 2Lin: *Given N variables $x_1, \ldots, x_N$ taking $\{-1, 1\}$-values and M equations $C_1, \ldots, C_M$ where each equation C_i is of the form $x_{i_1} \cdot x_{i_2} = b_i$ and $b_i \in \{-1, 1\}$. The goal is to find an assignment that maximizes the fraction of equations satisfied.*

Assuming the Unique Games Conjecture, given an instance of Boolean 2Lin where $1 - \varepsilon$ fraction of the equations can be satisfied, it is NP-hard to satisfy $1 - \Omega(\sqrt{\varepsilon})$ fraction of the equations [29, 18]. We refer to this approximation problem as the $(1 - \varepsilon, 1 - \Omega(\sqrt{\varepsilon}))$ gap problem of 2Lin. An NP-hardness result for it could perhaps be equivalent to the Unique Games Conjecture. Even though an equivalence is not known formally (in fact, a promising direction for proving it was ruled out in [42, 5]), researchers tend to agree that the boolean case captures the main difficulty of general Unique Games. If one were able to prove NP-hardness of $(1 - \varepsilon, 1 - \Omega(\sqrt{\varepsilon}))$ gap in the boolean case, the proof might likely extend to NP-hardness of $(1 - \varepsilon, 1 - K(\mathbb{F}) \cdot \sqrt{\varepsilon})$ gap for the general finite field case, where $K(\mathbb{F})$ is a constant with $K(\mathbb{F}) \to \infty$ as $|\mathbb{F}| \to \infty$. The latter result would then be enough, via parallel repetition, to amplify the gap to $(1 - o(1), o(1))$ and prove the Unique Games Conjecture [41]! At present however, we do not even know a $(1 - \varepsilon, 1 - C \cdot \varepsilon)$ gap with $C \to \infty$, even for general Unique Games, and even as Lasserre integrality gap. Hence, the weak UGC focuses on this "modest" goal as opposed to the more ambitious NP-hardness result.

In order to rule out a time n^d algorithm for unique games it suffices to show a size-$2^{\epsilon n/d}$ reduction from a problem with a (conjectured) exponential lower bound $2^{\epsilon n}$ (n is the input size). The problem we focus on is $k\mathrm{CSP}(P_{H\mathrm{Lin}})$. In this problem one is given a set of boolean variables together with arity-k constraints over them. Each constraint restricts its k variables to lie in a certain linear subspace (specifically, the constraint is the hypergraph linearity predicate used in [43]; see Definition 7). The goal is to find an assignment to the variables that satisfies as many constraints as possible. Chan [12] showed that the $(1 - o(1), (1 + o(1))(k + 1)/2^k)$ gap problem of $k\mathrm{CSP}(P_{H\mathrm{Lin}})$ is NP-hard. Tulsiani [48] showed hardness for the natural Lasserre semidefinite program with linear number of rounds, even for random instances of $k\mathrm{CSP}(P_{H\mathrm{Lin}})$.

Definition 4 (Weak Unique Games Conjecture)**.** *For any $C \geq 1$, any $t \geq 1$, any sufficiently small $\varepsilon > 0$, any $\zeta > 0$ and any sufficiently large $n \geq 1$, there is a reduction from size-n instances of random $k\mathrm{CSP}(P_{H\mathrm{Lin}})$ with gap $(1-o(1), (1+o(1))(k+1)/2^k)$ to size-$2^{\zeta n}$ instances of Boolean* 2LIN *with gap $(1-\varepsilon, 1-C\cdot\varepsilon)$. Furthermore, the reduction maps $\Omega(n)$-round Lasserre solutions of $k\mathrm{CSP}(P_{H\mathrm{Lin}})$ to t-round Lasserre solutions of Boolean* 2LIN*.*

The weak Unique Games Conjecture has two implications: (1) It rules out polynomial-time algorithms for the $(1 - \varepsilon, 1 - C \cdot \varepsilon)$ gap version of Boolean 2LIN under the assumption that the appropriately gapped version of random $k\mathrm{CSP}(P_{H\mathrm{Lin}})$ requires exponential time. (2) Via Tulsiani's result [48], it rules out any constant round Lasserre-based algorithm for the $(1-\varepsilon, 1-C\cdot\varepsilon)$ gap version of Boolean 2LIN.

In this paper we construct a *candidate* reduction towards a proof of the weak Unique Games Conjecture. We show the completeness of the reduction, i.e., how assignments satisfying $1-o(1)$ fraction of the constraints of $k\mathrm{CSP}(P_{H\mathrm{Lin}})$ translate to assignments satisfying $1-\varepsilon$ fraction of the

constraints of Boolean 2LIN. More than that, $\Omega(n)$-round Lasserre solutions for $k\mathrm{CSP}(P_{H\mathrm{LIN}})$ translate to t-round Lasserre solutions of Boolean 2LIN. We do not know how to prove the soundness of the reduction, namely that assignments that satisfy at least $1 - C \cdot \varepsilon$ fraction of the equations of Boolean 2LIN translate to assignments that satisfy $(1 + o(1))(k + 1)/2^k$ fraction of the constraints of $k\mathrm{CSP}(P_{H\mathrm{LIN}})$. However, we show soundness for a fairly general family of assignments. Our analysis (or potentially a variant of it) might imply soundness in general, but this depends on a solution to a certain question about Gaussian isoperimetry.

1.3 The Real Code and Gaussian Isoperimetry

Inapproximability reductions meant to prove the hardness of a certain target problem typically follow these steps:

1. *PCP Theorem* [2, 1]: Start with a kCSP instance on variables X and constraints $C_1, \ldots, C_M$. The PCP Theorem establishes the NP-hardness of approximating kCSP to within some constant.

2. *Parallel repetition* [42]: Generate a new instance whose variables are sets of variables in X, where each set induces several constraints from $C_1, \ldots, C_M$. Constraints on the new instance check consistency between sets that intersect in several variables. The parallel repetition theorem shows that the new instance is harder to approximate.

3. *Long Code* [7, 23]: Replace each set with an encoding of its assignment via the *long code*. Replace each constraint with constraints of the target problem over the long code variables.

Inapproximability of unique games does not follow from this framework, since the long code does not have a unique test for checking consistency between sets. In this work we suggest a related framework in order to prove the weak Unique Games Conjecture. In this framework one considers all sets of a certain large size (this is the reason that the reduction is nearly exponential). The long code is replaced with a new code, the *real code*, that does have a unique consistency test (see Section 1.4).

Unlike the long code, the real code no longer uses functions over the boolean hypercube $\{-1, 1\}^n$, but over $\mathbb{R}^n$ with the underlying space equipped with the standard Gaussian measure. The range of the functions is $\{-1, 1\}$ still. Specifically, we use half-spaces (or rather a "periodized" version of half-spaces[1]). For $z \in \mathbb{R}$, define:

$$\mathsf{interval}(z) = \begin{cases} +1 & \text{if} \quad z \in [2k, 2k+1) \text{ for some } k \in \mathbb{Z} \\ \\ -1 & \text{if} \quad z \in [2k-1, 2k) \text{ for some } k \in \mathbb{Z}. \end{cases}$$

The *real encoding* of a string $\sigma = (\sigma_1, \ldots, \sigma_n) \in \{-1, 1\}^n$ is now defined as the function $f_\sigma : \mathbb{R}^n \to \{-1, 1\}$:

$$f_\sigma(x) = \mathsf{interval}\left(\sum_{i=1}^{n} \sigma_i \cdot x_i\right).$$

Borell [9] showed that half-spaces are the solutions to the isoperimetric problem in Gaussian space. That is, among all sets of measure $1/2$, half-spaces have the least surface area. This

[1] In the periodized version, adding 1 to any coordinate flips the sign of the function. This property is useful for arguing that the function strongly depends on all its coordinates.

suggests the following unique "noise test" for checking whether an odd function $f : \mathbb{R}^n \to \{-1, 1\}$, $f(-x) \equiv -f(x)$, is a half-space (or a real code function): pick Gaussian $x \in \mathbb{R}^n$, and pick a small Gaussian shift of x, which we denote $x' \in \mathbb{R}^n$. Check whether $f(x) = f(x')$. The probability that the test rejects is roughly proportional to the surface area of the set $f^{-1}(1)$. The oddness of the function guarantees that the set is of measure 1/2. Borell [10] shows that half-spaces pass the test with maximal probability. Further work [19, 11, 14, 37, 38] establishes that any odd function that passes the test with maximal probability, or even slightly less, must be a half-space, except on a set of points of small measure. There are, however, odd functions that are not even correlated with a half-space, yet pass the test with high probability. Consider, for instance, the XOR of three far apart half-spaces. the probability that the XOR fails the test is bounded by three times the probability that a single half-space fails the test. More generally, we define real code juntas as follows:

Definition 5 (Real code junta)**.** *A* real code junta *is a function of the form*

$$f(x) = G(f_{\sigma_1}(x), \ldots, f_{\sigma_l}(x))$$

where $G : \{-1, 1\}^l \to \{-1, 1\}$ *is a combining function on which each coordinate has constant influence (i.e., for every* $1 \le i \le l$ *we have* $\mathbf{E}_{y^{-i}}[\mathbf{Var}_{y_i}[G_S(y_1, \ldots, y_l)]] \ge \Omega(1)$*). The number* l *is the size of the junta.*

The probability that a real code junta fails the test is at most l times the probability that a real code function fails the test. We show the soundness of our construction assuming that only (approximate) real code juntas can be used in lieu of real code functions.

Theorem 1.1 (Main)**.** *Our reduction from random* $k\mathrm{CSP}(P_{H\mathrm{LIN}})$ *to Boolean* 2LIN *satisfies the following:*

1. Completeness: *An assignment that satisfies* $1 - o(1)$ *fraction of the constraints of the* $k\mathrm{CSP}(P_{H\mathrm{LIN}})$ *instance can be translated to an assignment that satisfies* $1 - \varepsilon$ *fraction of the equations of the* 2LIN *instance.*

2. Lasserre completeness: *A vector solution to* $\Omega(n)$*-round Lasserre for the* $k\mathrm{CSP}(P_{H\mathrm{LIN}})$ *instance can be translated to a* t*-round Lasserre solution for* 2LIN*.*

3. Restricted soundness: *If there is no assignment to the* $k\mathrm{CSP}(P_{H\mathrm{LIN}})$ *instance that satisfies* $(1 + o(1))(k + 1)/2^k$ *fraction of the constraints, then there is no assignment that satisfies* $1 - C \cdot \varepsilon$ *fraction of the equations of the* 2LIN *instance while only using approximate,* $O(1)$*-sized, real code juntas in lieu of real code functions.*

Whether the weak Unique Games Conjecture could follow (either directly from our analysis or from a strengthening of it) depends on whether the case addressed in our restricted soundness is "essentially" the only case. The underlying mathematical question is of independent interest:

> **Robust Gaussian Isoperimetry:** Which functions[2] fail the noise test with probability only a constant times larger than a (periodized) half-space? Are functions "influenced" by a constant number of (periodized) half-spaces the only such functions?

[2]Among functions that satisfy $f(-x_1, \ldots, -x_n) = -f(x_1, \ldots, x_n)$ and $f(x_1, \ldots, x_i + 1, \ldots, x_n) = -f(x_1, \ldots, x_i, \ldots, x_n)$ for all $x_1, \ldots, x_n \in \mathbb{R}$ and $1 \le i \le n$.

Without periodization, low degree polynomial threshold functions pass the noise test with probability not much larger than half-spaces [25]. We remark that our techniques apply even if the function is correlated with a real code junta only after a random restriction of a constant fraction of the coordinates.

It is interesting to note the analogy between real code testing and long code testing. The long code consists of functions $f_i(x) = x_i$ for $1 \leq i \leq n$. It too has a unique test, obtained by comparing $f(x) = f(x')$ where x is uniform in the boolean hypercube and x' is obtained from x by flipping each coordinate with a small probability. Long code functions depend only on a single coordinate, and therefore pass the test with exceptionally high probability. Juntas, which are functions that depend on a constant number of coordinates, fail the test with probability that is only a constant times larger than that of long code functions. It is known [29, 18] that the only functions that pass the test with high probability are those that have a constant number of influential coordinates for an appropriate definition of influence. (In an amusing turn of events, this is proved by drawing an analogy between functions with no influential coordinates over the boolean hypercube and functions in Gaussian space, and relying on Borell's theorem mentioned above.)

1.4 The Real Code: Combining Advantages of Long Code and Hadamard

Inapproximability reductions typically either use the long code or use the Hadamard code (see, e.g., [26]). The long code encodes an index $i \in [n]$, or equivalently a $\log n$ bit string, as the dictator function $f(x_1, \ldots, x_n) = x_i$ defined over $(x_1, \ldots, x_n) \in \{-1, 1\}^n$. Hadamard code encodes a string $(\sigma_1, \ldots, \sigma_n) \in \{-1, 1\}^n$ as the function $f(x_1, \ldots, x_n) = \prod_{j:\sigma_j=-1} x_j$ defined over $(x_1, \ldots, x_n) \in \{-1, 1\}^n$. Evidently, the long code has a much worse rate than the Hadamard code; that is, one encodes much less information using the long code compared to the Hadamard code.

Unlike the long code, the Hadamard code does not have a unique local test (it is nevertheless very useful in other applications, thanks to tests with three or more queries [8]). The reason is simple: for any two distinct locations $x, x' \in \{-1, 1\}^n$, half of the legitimate Hadamard codes satisfy $f(x) = f(x')$ and the remaining half satisfy $f(x) \neq f(x')$, and thus any unique local test fails with probability $\frac{1}{2}$ on some legitimate Hadamard code.

However, the Hadamard code has the following *sub-code covering* property: the Hadamard code of a string $(\sigma_1, \ldots, \sigma_n)$ is nearly uniformly covered by the Hadamard codes of its proper substrings.[3] This property is potentially useful as follows: an encoding of a set is nearly uniformly covered by the encodings of slightly smaller sets. Since the encodings of the smaller sets are already "contained" in the encodings of the larger sets, only the latter explicitly appear in the "PCP proof", the former being "present implicitly". Now, one may simply run a test (a

[3]Here is a more precise statement. Pick a random subset $S \subseteq [n]$ of size $(1 - \delta)n$ and a random string $x' \in \{-1, 1\}^S$; define a string $x \in \{-1, 1\}^n$ by letting $x_j = x'_j$ if $j \in S$ and $x_j = 1$ otherwise; then the distribution of x is statistically close to uniform over $\{-1, 1\}^n$ (provided $\delta \ll \frac{1}{\sqrt{n}}$, see [31]). Note that x denotes a typical location in the Hadamard code of a string $\sigma \in \{-1, 1\}^n$. The bit of the code at this location is $\prod_{j:\sigma_j=-1} x_j$ and since the coordinates of x outside S are set to 1, this bit depends only on the coordinates in S, i.e. on x'. On the other hand, x' denotes a typical location in the Hadamard code of the substring $\sigma|_S$, i.e. σ restricted to S. In this sense, the Hadamard codes of substrings of σ of length $(1 - \delta)n$ nearly uniformly cover the Hadamard code of σ. The specific manner in which x is chosen can be restated as follows. Pick a random subset $S \subseteq [n]$ of size $(1 - \delta)n$; pick the coordinates in S uniformly at random from $\{-1, 1\}$; pick the coordinates outside S to be "small" in value. In the boolean field $\{-1, 1\}$, "small" amounts to the value 1, i.e. a value that has no effect on the bit of the code at location x.

unique local test if one intends to show hardness of the Unique Games problem) on the encoding of the larger sets. In the soundness analysis, one is able to "list-decode" this encoding. Since the encoding is nearly uniformly covered by the encodings of smaller sets, essentially the same list-decoding also serves as the list-decoding for smaller sets, leading to "consistent decodings" on all sets, completing the soundness analysis. In [31], this recipe is demonstrated using the Hadamard code. Therein the application is different (and not to the Unique Games problem, since the Hadamard code does not have a unique local test).

The real code combines the advantages of both the long code (unique local test) and Hadamard (sub-code covering property). We discussed the local test in the previous section. Moreover, the real code has a property analogous to the sub-code covering property (see footnote for a comparison with the Hadamard code). Pick a random subset $S \subseteq [n]$ of size $(1-\delta)n$ and a random input $x' \in \mathbb{R}^S$ from the standard Gaussian measure; define an input $x \in \mathbb{R}^n$ by letting $x_j = x'_j$ if $j \in S$ and a uniformly random number in $[-\delta, \delta]$ otherwise; then x "looks like" an input chosen from $\mathbb{R}^n$ with the standard Gaussian measure. The reason is that a typical input chosen from $\mathbb{R}^n$ with the standard Gaussian measure does have a fraction δ of the coordinates with magnitude $O(\delta)$ and so "looks like" the input x. Akin to the Hadamard code, the coordinates of x outside S are small in value and do not influence much the bit of the real encoding at location x.

Remark 1.1. *In recent years, researchers suggested the "short code" (aka the "low degree long code") as a more efficient alternative to the long code [4, 16]. The short code has a unique test, but does not have the sub-code covering property.*

1.5 Soundness for Real Code Juntas

In this section we discuss what goes into the proof that our reduction is sound assuming that the only functions used in lieu of the real code are (approximate) real code juntas.

Like most PCP constructions, our construction composes an outer construction with large alphabet and an inner construction with binary alphabet. In our setup the outer construction does not have the uniqueness property, while the inner construction does, therefore resulting in a unique game. The key difficulty is that in our construction – unlike in standard constructions – the inner construction leaks information about the outer construction. The outer construction can be described as a direct product game. A verifier picks at random two sets over a known universe such that the sets intersect. Each set is sent to a different player. Each player responds with a label for each of the elements in its set. The verifier checks that the two players agree on the intersection of the two sets. In the composed construction there is $f_S : \mathbb{R}^n \to \{-1, 1\}$ for every set $S \subseteq [N]$, $|S| = n$. An outer verifier performs a direct product game with sets S, R and the inner verifier queries different parts of f_S, f_R depending on the intersection $S \cap R$. List decodings of f_S, f_R partition $\mathbb{R}^n$ in adversarial ways, and may therefore give information about the intersection.

We present a general technique for handling direct product games with leakage, and show that a direct product game with limited leakage (approximately) contains a slightly smaller standard direct product game. In our analysis we build on information-theoretic ideas that originated in the study of parallel repetition [42, 24]. Parallel repetition too can be viewed in terms of games with leakage. There the leakage takes the form of the event that the two players succeed in certain rounds, and the goal is to fix $n-1$ elements such that the remaining element is approximately unaffected by leakage. In our setup the leakage is more general and the goal is to fix only a small fraction of elements.

1.6 Approximate Real Linear Equations

An important intermediate point of our construction is the approximate real linear equations problem. In this problem, one is given a system of linear equations over reals and each equation is of the form:

$$\sum_{j=1}^{k} b_j y_j = 0,$$

where $b_j \in \{-1, 1\}$ and y_j are real variables. One wishes to satisfy the equations approximately, and not necessarily exactly. Also, one wishes to assign at least a constant fraction of the variables values that are at least a constant in magnitude (and so, one cannot "cheat" by assigning the zero value to all the variables). Given a real valued assignment to the variables, the margin (or the error) on a typical equation as above is $\left|\sum_{j=1}^{k} b_j y_j\right|$ and the goal is to find an assignment that (approximately) minimizes the average margin over all the equations. Note that, assigning the variables y_j random $\{-1, 1\}$ values, one can always achieve a margin of $O(\sqrt{k})$ on average, so the question is whether one can do better (and the answer is negative as explained next).

In analogy to the discrete case, we refer to the approximate real linear equations problem as kLin($\mathbb{R}$). In the paper [30], the authors prove an optimal NP-hardness result for the problem 3Lin($\mathbb{R}$), i.e. even when each equation has only three variables. The authors show that when an assignment with average margin ε exists, it is NP-hard to find an assignment with average margin $O(\sqrt{\varepsilon})$, i.e. there is a quadratic gap. This is shown to be optimal in the sense that there is a matching SDP algorithm (effectively a *least square fit* algorithm). In [30], the coefficients b_j in the equations are allowed to be in a bounded interval, as opposed to being $\{-1, 1\}$, but this is a minor point. In the current paper, we show a reduction from random kCSP($P_{H\text{Lin}}$) to kLin($\mathbb{R}$), and reduce from kLin($\mathbb{R}$) to Boolean 2Lin.

1.7 Organization

We discuss constraint satisfaction problems, their Lasserre semidefinite programs, and Tulsiani's result in Section 2. We obtain the integrality gap for approximate real linear equations problem in Section 3. We discuss the real code in Section 4, and how to incorporate constraint test in the real code in Section 5. We show how to check consistency between real codes in Section 6. The overall candidate integrality gap instance for the Boolean 2Lin problem is in Section 7, and in Sections 8, 9 and 10 we prove soundness for real code juntas.

Appendix

2 Constraint Satisfaction Problems and their Lasserre Semidefinite Programs

A predicate $P : \{-1,1\}^k \mapsto \{0,1\}$ leads to a constraint satisfaction problem (CSP) as follows. There are N variables taking values in $\{-1,1\}$ and M constraints, each defined on some (ordered) tuple of k variables. In a constraint, each variable first gets a $\{-1,1\}$ sign ("polarity"), and then the predicate P is applied on the tuple of polarized variables:

Definition 6. *For a predicate $P : \{-1,1\}^k \mapsto \{0,1\}$, an instance of CSP(P) consists of N variables $x_1, \ldots, x_N$ and M constraints $C_1, \ldots, C_M$, where each constraint C is over a k-tuple of variables $\{x_{i_1}, \ldots, x_{i_k}\}$ and is of the form $P(b_1 x_{i_1}, \ldots, b_k x_{i_k})$ where $b_1, \ldots, b_k \in \{-1,1\}$.*

We overload notation by using C to denote a typical constraint, as well as the tuple of variables appearing in it. For $j \in [k]$, we let $C[j] \in [N]$ denote the index of the j^{th} variable in C and $b_C \in \{-1,1\}^k$ be a vector such that its j^{th} coordinate $b_C[j]$ indicates the polarity of the j^{th} variable in C. For $u, v \in \{-1,1\}^k$, let $u \circ v \in \{-1,1\}^k$ denote the coordinate-wise product of u, v. A random instance of CSP(P) is one where the constraints are on randomly chosen k-tuples of variables, and the polarities of variables are randomly chosen as well (independently for occurrences in different constraints).

The optimization problem associated with CSP(P) is to find an assignment to the variables that satisfies the largest number of constraints. If Φ is an instance of CSP(P), then we denote the maximum number of satisfied constraints by $OPT(\Phi)$.

The t-round Lasserre semidefinite program for the CSP(P) problem has a vector variable $V_{S,\alpha}$ for every set of variables $S \subseteq [N]$, $|S| \leq t$, and an assignment $\alpha \in \{-1,1\}^{|S|}$ to the variables in S. In the intended solution, $V_{S,\alpha}$ is some (globally fixed) unit vector if S is assigned α, and is the zero vector otherwise. If $\alpha_1 \in \{-1,1\}^{t_1}$ is an assignment to a set $S_1 \subseteq [N]$ of variables, and $\alpha_2 \in \{-1,1\}^{t_2}$ is an assignment to a set $S_2 \subseteq [N]$ of variables, then we say that α_1 and α_2 agree if they assign the same values to variables in $S_1 \cap S_2$; otherwise, we say that they disagree. The Lasserre program attempts to maximize the number of satisfied constraints as reflected by the vector variables, subject to consistency constraints on the variables:

Lasserre semidefinite program

$$\begin{aligned}
\max \quad & \sum_{i=1}^{M} \sum_{\alpha \in \{-1,1\}^k} P(\alpha \circ b_C) \|V_{C_i,\alpha}\|^2 & & \\
\text{s.t.} & & & \\
\textbf{Orthogonality}: \quad & \langle V_{S,\alpha}, V_{S,\beta} \rangle = 0 & \forall S, \alpha \neq \beta & (1) \\
\textbf{Consistency}: \quad & V_{S,\alpha} = V_{S \cup \{x\}, \alpha \cup \{+1\}} + V_{S \cup \{x\}, \alpha \cup \{-1\}} & \forall S, \alpha, x \notin S & (2) \\
\textbf{Non-negativity}: \quad & \langle V_{S,\alpha}, V_{T,\beta} \rangle \geq 0 & \forall S, T, \alpha, \beta & (3) \\
\textbf{Normalization}: \quad & \|V_{\phi,\phi}\|^2 = 1. & & (4)
\end{aligned}$$

Local distributions: In any feasible solution to the semidefinite program, for every S, we have

$$\sum_{\beta \in \{-1,1\}^{|S|}} \|V_{S,\beta}\|^2 = 1.$$

Thinking of $\|V_{S,\beta}\|^2$ as probabilities, a vector solution to the Lasserre semidefinite program induces a distribution over assignments to S, referred to as the local distribution on S. In particular, for each constraint C, one has a local distribution on assignments to that constraint. Any set $S \supseteq C$ induces a distribution over assignments to C by picking $\beta \in \{-1,1\}^{|S|}$ with probability $\|V_{S,\beta}\|^2$ and restricting β to the variables in C. The feasibility of the solution ensures that this induced distribution coincides with the local distribution on C.

Note that the objective function of the program, measuring the quality of the solution, depends only on the local distributions. If one only has local distributions and consistency among them, rather than vector solutions and consistency among them, one gets the so-called Sherali-Adams linear programming relaxation. In the first reading, the reader might want to focus only on the local distributions. We rely on a result of Tulsiani concerning CSPs with the following linear predicate:

Definition 7. The Hypergraph Linearity Test Predicate: *For $k = 2^s - 1$, the hypergraph linearity test predicate $P_{H\text{LIN}} : \{-1,1\}^k \mapsto \{0,1\}$ is defined as follows. Index the k coordinates by non-empty subsets $A \subseteq [s]$ and assume w.l.o.g. that the first s coordinates correspond to the singleton sets. Then $x \in \{-1,1\}^k$ is a satisfying assignment of the predicate $P_{H\text{LIN}}$ (i.e., $P_{H\text{LIN}}(x) = 1$) if and only if*

$$x_A = \prod_{i \in A} x_{\{i\}} \quad \forall\ 2 \leq |A| \leq s.$$

In other words, the satisfying assignments of the predicate are precisely the Hadamard codewords and $2^s = k + 1$ in number.

Samorodnitsky and Trevisan [44] constructed a UGC-based PCP using the Hypergraph Linearity Test Predicate, and Chan [12], in a recent remarkable result, constructed a similar PCP without relying on the UGC. Regarding the Lasserre integrality gap, Tulsiani, building on the works of Grigoriev and Schoenebeck [21, 45], shows the following (the statement is tailored to our needs):

Theorem 8 (Tulsiani [48])**.** *Let T be an arbitrarily large constant. Let Φ be a randomly chosen instance of $\mathsf{CSP}(P_{H\text{LIN}})$ with N variables and $M = TN$ constraints for large enough (growing) N. Let $n_0 = \lfloor \eta N \rfloor$ where $\eta = \frac{1}{T^{25}}$. Then with high probability over the choice of Φ:*

1. Completeness: *Φ has an n_0-round Lasserre SDP solution with objective value M, where for every constraint C, the local distribution on C is uniform over the satisfying assignments to C (i.e. uniform over $P^{-1}_{H\text{LIN}}(1) \circ b_C$).*
2. Soundness: *$OPT(\Phi) \leq (1 + o(1))\frac{k+1}{2^k} \cdot M$.*

In the soundness case, note that since the hypergraph linearity test predicate has $k+1$ satisfying assignments, the expected number of constraints satisfied by a random $\{-1,1\}$ assignment to the variables is $((k+1)/2^k) \cdot M$. A standard argument shows that, with high probability over the choice of the instance, no assignment satisfies a slightly larger fraction of constraints (here the $o(1)$ term becomes arbitrarily small as T increases). What is remarkable is that in the completeness case, there is a SDP solution, up to a linear number of rounds of Lasserre, which "pretends" that there is an assignment satisfying all the constraints.

3 Lasserre Integrality Gap for Approximate Real Linear Equations

In this section, we construct a Lasserre integrality gap for approximate real linear equations problem. The construction is essentially a re-interpretation of Tulsiani's Lasserre integrality gap for the hypergraph linearity test predicate, where we re-interpret a predicate over boolean domain as an equation over reals with carefully chosen coefficients. As Tulsiani's instance is a random instance of $\mathrm{CSP}(P_{\mathrm{HLIN}})$, our integrality gap can be thought of as a random, or average-case, analog of our NP-hardness result in [30].

Recall that the predicate $P_{H\mathrm{LIN}} : \{-1,1\}^k \mapsto \{0,1\}$ has exactly $k+1$ satisfying assignments and $k = 2^s - 1$. It is easily verified that for any two distinct satisfying assignments $a, b \in P^{-1}_{H\mathrm{LIN}}(1)$, we have $\sum_{j=1}^{k} a_j b_j = -1$. Indeed, since the k co-ordinates are indexed by non-empty subsets $A \subseteq [s]$, for some distinct $x, y \in \{-1,1\}^s$, we have

$$\sum_{j=1}^{k} a_j b_j = \sum_{A \subseteq [s],\ A \neq \phi} \prod_{i \in A} x_i \cdot \prod_{i \in A} y_i = -1.$$

Let Φ be an instance of $\mathrm{CSP}(P_{H\mathrm{LIN}})$ with N variables and M constraints. A typical constraint is denoted as C along with the vector b_C of polarities. We construct a system of linear equations over reals by replacing each constraint C by a set of $k+1$ linear equations over reals, one equation, as below, for each *sign vector* $\epsilon \in P^{-1}_{H\mathrm{LIN}}(1)$:

$$\sum_{j=1}^{k} \epsilon_j \cdot b_C[j] \cdot y_{C[j]} = 0,$$

where $y_1, \ldots, y_N$ are real-valued variables. In the following, a constraint C will refer to the constraint as in the $\mathrm{CSP}(P_{H\mathrm{LIN}})$ instance, and also to any of the $k+1$ real linear equations constructed from it. It should be clear from the context which is being referred to. We observe that a uniformly random satisfying assignment to the constraint C (i.e. uniform in $P^{-1}_{H\mathrm{LIN}}(1) \circ b_C$) is, on average, a good assignment to each of the linear equations constructed from it, in terms of the average ℓ_1 error (i.e. margin).

Fact 3.1. *Let $\epsilon \in P^{-1}_{H\mathrm{LIN}}(1)$ be any fixed sign vector. Then*

$$\mathop{\mathbf{E}}_{\alpha \in P^{-1}_{H\mathrm{LIN}}(1) \circ b_C} \left[\left| \sum_{j=1}^{k} \epsilon_j \cdot b_C[j] \cdot \alpha_j \right| \right] = \frac{2k}{k+1}.$$

Proof. Substituting $\alpha = \beta \circ b_C$ above and canceling out the polarities, the expectation is

$$\mathop{\mathbf{E}}_{\beta\ \in P^{-1}_{H\mathrm{LIN}}(1)} \left[\left| \sum_{j=1}^{k} \epsilon_j \cdot \beta_j \right| \right].$$

Note that $P_{H\mathrm{LIN}}$ has $k+1$ satisfying assignments. Out of these, there is one assignment that equals ϵ and the inner sum equals k. For the remaining k assignments $\beta \neq \epsilon$ and the inner sum equals -1 as observed before. □

Motivated by the observation that $\{-1,1\}$-valued assignments to the variables suffice for approximate satisfaction of the real linear equations we defined, we continue to refer to the Lasserre semidefinite program we described before, which has a variable $V_{S,\alpha}$ per set S of at most t variables and per $\{-1,1\}$ assignment α to the variables in S. We keep the feasibility conditions of this program, but drop the objective function (which talks about satisfying the predicate $P_{H\text{LIN}}$).

Theorem 8 now directly implies our Lasserre integrality gap for approximate real linear equations problem, stated as Theorem 9 below. In the completeness part, we have a feasible vector solution inducing local distributions that on average approximate each equation up to a margin $O(1)$. In the soundness part, we have that every boolean assignment to the variables has average margin $\Omega(\sqrt{k})$, averaged over all the equations.

Theorem 9. *Let T be an arbitrarily large constant. Let Φ be a randomly chosen instance of kLIN$(\mathbb{R})$ with N variables and $M = TN$ constraints for large enough (growing) N, as described above. Note that every equation is of the form $\sum_{j=1}^{k} \epsilon_j \cdot b_C[j] \cdot y_{C[j]} = 0$. Let $n_0 = \lfloor \eta N \rfloor$ where $\eta = \frac{1}{T^{25}}$. Then w.h.p. over the choice of Φ we have:*

1. **Completeness:** *Φ has an n_0-round Lasserre SDP feasible solution, where for every constraint C, the local distribution is uniform on its satisfying assignments, i.e. $P^{-1}_{H\text{LIN}}(1) \circ b_C$. In particular, for every sign vector $\epsilon \in P^{-1}_{H\text{LIN}}(1)$, when taking expectation over the local distribution on C,*

$$\mathop{\mathbf{E}}_{\sigma \in P^{-1}_{H\text{LIN}}(1) \circ b_C} \left[\left| \sum_{j=1}^{k} \epsilon_j \cdot b_C[j] \cdot \sigma(C[j]) \right| \right] = \frac{2k}{k+1} \;\leq\; 2. \tag{5}$$

2. **Soundness:** *For some absolute constant $c > 0$, the following holds: for any global assignment $\sigma : [N] \mapsto \{-1,1\}$, for a c fraction of the equations (where an equation is specified by constraint C and a sign vector ϵ), the margin is at least $c \cdot \sqrt{k}$, i.e.*

$$\left| \sum_{j=1}^{k} \epsilon_j \cdot b_C[j] \cdot \sigma(C[j]) \right| \;\geq\; c \cdot \sqrt{k}. \tag{6}$$

Proof. The completeness part follows from Theorem 8 and Fact 3.1. The soundness part is a standard probabilistic argument over the choice of the instance Φ: fix a global assignment σ, fix the tuples of variables that appear in all the constraints, and consider the choice of the polarities for all the constraints. Given a constraint C and a sign vector ϵ, over the choice of random polarities $b_C[j]$, the sum $\sum_{j=1}^{k} \epsilon_j \cdot b_C[j] \cdot \sigma(C[j])$ is at least $c \cdot \sqrt{k}$ in magnitude with probability c for some absolute constant $c > 0$. Since the polarities are chosen independently for different constraints and the number of constraints TN is large relative to the number of variables N, one can apply Chernoff bound and a union bound. □

Remark 3.1. *(1) In the soundness case above, it is easy to extend the conclusion to all assignments $\sigma : [N] \mapsto \mathbb{R}$ (as opposed to only boolean assignments $\sigma : [N] \mapsto \{-1,1\}$) as long as σ assigns, to a constant fraction of the variables, values that are at least a constant in magnitude. (2) The gap $(O(1),\ \Omega(\sqrt{k}))$ in the completeness versus the soundness case in Theorem 9 is our starting gap. We re-emphasize some of the points mentioned before. Dividing by a normalization factor of k, the gap here is $(O(\frac{1}{k}), \Omega(\frac{1}{\sqrt{k}}))$, i.e., a quadratic gap. In [30], the authors indeed*

prove that it is NP-hard to distinguish such a quadratic gap even when equations involve three variables. Thus Theorem 9 is the integrality gap analogue of NP-hardness result in [30]. The authors therein propose that there may be a further reduction from the system of equations with three variables (3Lin($\mathbb{R}$)) to a system of equations with two variables (2Lin($\mathbb{R}$)) and/or the closely related Boolean 2Lin problem. This note may substantiate their proposal, albeit in the context of Lasserre integrality gaps. We are at present unable to show soundness of our construction. If the construction is sound and the techniques to analyze soundness are developed, it may be possible to extend the integrality gap construction to a NP-hardness reduction.

4 The Real Code and the Gaussian Noise Test

The encoding scheme in our construction is the real code as explained in the introduction. Recall that for some assignment $\sigma : [n] \mapsto \{-1, 1\}$, its real code encoding $f_\sigma : \mathbb{R}^n \mapsto \{-1, 1\}$ is supposed to be the interval function of σ. Specifically, defining for $z \in \mathbb{R}$,

$$\mathsf{interval}(z) = \begin{cases} +1 & \text{if} \quad z \in [2k, 2k+1) \text{ for some } k \in \mathbb{Z} \\ \\ -1 & \text{if} \quad z \in [2k-1, 2k) \text{ for some } k \in \mathbb{Z}, \end{cases} \tag{7}$$

we let, with underlying standard Gaussian measure on $\mathbb{R}^n$ (denoted $\mathcal{N}^n$),

$$f_\sigma(x) = \mathsf{interval}\left(\sum_{i=1}^{n} \sigma(i) \cdot x_i\right). \tag{8}$$

Let Φ be an instance of kLin($\mathbb{R}$) as in Theorem 9 along with the Lasserre SDP solution. For every set $S \subseteq [N]$, $|S| = n$, our final construction has a copy of the "gadget" $f_S : \mathbb{R}^n \mapsto \{-1, 1\}$ which is supposed to be the real encoding of the assignment to S. We use a standard PCP trick called *folding* to enforce certain basic properties of f_S:

Folding: The encoding $f = f_\sigma$ in Equation (8) satisfies

$$f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) = -f(x_1, \ldots, x_{i-1}, x_i + 1, x_{i+1}, \ldots, x_n),$$

for any $i \in [n]$ and $x \in \mathbb{R}^n$. Moreover f is odd, i.e., $f(-x) = -f(x)$ for any $x \in \mathbb{R}^n$. By using these identities for evaluating f, we can assume that the functions f in our Boolean 2Lin instance always satisfy these identities.

Now we describe a test that checks that a given function $f : \mathbb{R}^n \mapsto \{-1, 1\}$, at least loosely speaking, resembles an encoding of some assignment $\sigma : [n] \mapsto \{-1, 1\}$, or more generally of some *reasonable* assignment $\sigma : [n] \mapsto \mathbb{R}$. We call it the *low boundary test*. It fails with only a *small* probability whenever f is indeed a correct encoding of a boolean assignment (there will be more tests, namely a *constraints test* and a *consistency test* that will be added later). The test below is applied on a given function $f : \mathbb{R}^n \mapsto \{-1, 1\}$. We think of the parameter α as infinitesimally small.

Low Boundary Test with Parameter α

- Pick $x, w \in \mathcal{N}^n$ independently and let $y = (1-\alpha)x + \sqrt{2\alpha - \alpha^2}w$ (thus y is a α-perturbation of x).

- Reject if and only if $f(x) \neq f(y)$.

Lemma 4.1. *If $f : \mathbb{R}^n \mapsto \{-1, 1\}$ is an encoding of a boolean assignment $\sigma : [n] \mapsto \{-1, 1\}$ as in Equation (8), then f rejects the low boundary test with probability $O(\sqrt{\alpha n})$.*

Proof. This is because f, viewed as a partition of $\mathbb{R}^n$, has a Gaussian boundary/surface-area $O(\sqrt{n})$ and then one uses Corollary 14 in [35]. Alternately, it is easily seen that the two sums $\sum_{i=1}^n \sigma(i) \cdot x_i$ and $\sum_{i=1}^n \sigma(i) \cdot y_i$ are typically spread over a band of width $\Theta(\sqrt{n})$ around the origin and typically differ by $O(\sqrt{\alpha n})$. Thus the probability that the two sums lie in adjacent odd/even intervals is $O(\sqrt{\alpha n})$. □

5 The Real Code Augmented with Constraint Test

We now augment the basic gadget with an additional test that allows for checking kLin($\mathbb{R}$) constraints. Suppose that there is a constraint $C \subseteq [n]$ of the kLin($\mathbb{R}$) instance of the form:

$$\sum_{j=1}^{k} \epsilon_j \cdot b_C[j] \cdot y_{C[j]} = 0.$$

For any such constraint C, let $\mathbf{v}_C$ denote the unit vector in $\mathbb{R}^n$ that has $\frac{\epsilon_j \cdot b_C[j]}{\sqrt{k}}$ in the position $C[j]$ and zero elsewhere. Let β be a parameter thought of as infinitesimally small.

The Constraint Test for a Given Constraint C and Parameter β

- Pick $x, y \in \mathbb{R}^n$ such that both x, y are distributed as $\mathcal{N}^n$ and $y = x + \beta\ell\mathbf{v}_C$ and $\ell \sim \mathcal{N}$. Specifically, x, y are picked by first selecting their common component orthogonal to $\mathbf{v}_C$ and then selecting their components along $\mathbf{v}_C$ in a $(1 - \frac{\beta^2}{2})$-correlated manner.[4]
- Reject if and only if
$$f(x) \neq f(y). \tag{9}$$

Note that x, y differ only on coordinates in C. Next we analyze the behavior of the test on local distributions induced by the Lasserre SDP solution:

Lemma 5.1. *Suppose an assignment $\sigma : [n] \mapsto \{-1, 1\}$ is sampled from the local distribution on set S as given by the Lasserre SDP solution. Let $C \subseteq [n]$ be a constraint with a corresponding linear equation of the form $\sum_{j=1}^{k} \epsilon_j \cdot b_C[j] \cdot y_{C[j]} = 0$. Then the average rejection probability of the Constraint Test (w.r.t. constraint C) over the choice of σ is at most $O(\frac{\beta}{\sqrt{k}})$.*

Proof. Note that in the Lasserre solution, the restriction of σ to the constraint C is uniformly distributed over the satisfying assignments to C. Hence, over the choice of σ,

$$\sum_{j=1}^{k} \epsilon_j \cdot b_C[j] \cdot \sigma(C[j]) = \begin{cases} -1 & \text{with probability } \frac{k}{k+1} \\ k & \text{with probability } \frac{1}{k+1}. \end{cases} \tag{10}$$

[4] If y^* and x^* denote the components along $\mathbf{v}_C$, then this amounts to saying $y^* = (1 - \frac{\beta^2}{2})x^* + \sqrt{\beta^2 - \frac{\beta^4}{4}}w^*$ for $x^*, w^* \sim \mathcal{N}$. Thus $y^* - x^* \sim \beta\mathcal{N}$.

In the first case, $\langle \sigma, \mathbf{v}_C \rangle = -\frac{1}{\sqrt{k}}$, hence

$$\langle \sigma, y - x \rangle \sim \frac{\beta}{\sqrt{k}} \mathcal{N},$$

and the Constraint Test (9) rejects with probability $O(\frac{\beta}{\sqrt{k}})$. This is because the sums $\sum_{i=1}^{n} \sigma(i) \cdot x_i$ and $\sum_{i=1}^{n} \sigma(i) \cdot y_i$ are spread over a band of width $\Theta(\sqrt{n})$ around the origin, and their difference is distributed as $\frac{\beta}{\sqrt{k}} \mathcal{N}$ as shown. Thus the probability that the two sums lie in adjacent odd/even intervals is $O(\frac{\beta}{\sqrt{k}})$.

Similarly, in the second case, $\langle \sigma, \mathbf{v}_C \rangle = \sqrt{k}$, hence

$$\langle \sigma, y - x \rangle \sim \beta\sqrt{k}\, \mathcal{N},$$

and the Constraint Test (9) rejects with probability $O(\beta\sqrt{k})$. Overall, the Constraint Test rejects with probability $O(\frac{\beta}{\sqrt{k}})$. □

The constraint test examines the behavior of the function along a small number k of coordinates among the n coordinates, whereas the low boundary test might be insensitive to changes in such a small number of coordinates. This motivates a generalization of the low boundary test which focuses on any given subset K of the coordinates (this generalization will also be a part of our final construction). For $x \in \mathbb{R}^n$, let x_K denote the restriction of x to the coordinates in $K \subseteq [n]$, and let $x_{\overline{K}}$ denote the restriction of x to coordinates in $\{1, \ldots, n\} \setminus K$. As before, we think of the parameter α as infinitesimally small.

General Low Boundary Test with Parameter α on Subset $K \subseteq \{1, \ldots, n\}$

- Pick $x, w \in \mathcal{N}^n$ independently. Let $y_{\overline{K}} = x_{\overline{K}}$ and $y_K = (1 - \alpha)x_K + \sqrt{2\alpha - \alpha^2} w_K$.
- Reject if and only if $f(x) \neq f(y)$.

When $K = \{1, \ldots, n\}$, this test is same as the low boundary test we defined before. Lemma 4.1, analyzing the boundary test on the real code, continues to hold for any $K \subseteq \{1, \ldots, n\}$ with the appropriate scaling:

Lemma 5.2. *If $f : \mathbb{R}^n \mapsto \{-1, 1\}$ is an encoding of a boolean assignment $\sigma : [n] \mapsto \{-1, 1\}$ as in Equation (8), then f rejects the general low boundary test on subset K with probability $O(\sqrt{\alpha\,|K|})$.*

6 The Consistency Test

Our proposed integrality gap instance for the Boolean 2Lin problem (in Section 7) has a block of variables for every subset $S \subseteq [N], |S| = n$, for an appropriate setting of the parameter n. The variables correspond to points in $\mathbb{R}^n$ (discretized appropriately and weighed according to the standard Gaussian measure). The variables are boolean and an assignment to the variables in a block corresponds to a function $f_S : \mathbb{R}^n \mapsto \{-1, 1\}$. For some global assignment $\tau : [N] \mapsto \{-1, 1\}$, the function f_S is intended to be the encoding of $\sigma = \tau|_S$ as in Equation (8). On each block, we are going to perform the two tests described so far. In addition, we need a test to check consistency between different blocks (i.e., to check, in at least some loose sense, that the

functions f_S for different blocks are encodings of block assignments $\sigma(S)$ that are consistent across blocks, giving rise to consistent global assignment). We describe the consistency test next.

Roughly speaking if there are two blocks S and R such that $|S \cap R| \approx (1-\delta)n$, then the (intended) linear interval functions f_S and f_R are nearly the same and thus we may test that this is indeed the case. We describe the test formally below. An absolutely crucial aspect of the test is that the coordinates in $S \setminus R$ and $R \setminus S$ are *very small* compared to the coordinates in $S \cap R$.

The test has a parameter δ satisfying $\frac{1}{n^{1/2}} \ll \delta \ll \frac{1}{n^{1/3}}$. It is not clear what the correct setting should be. For now think of $\delta = \frac{1}{\sqrt{n}}$. Let $I = [-s, s]$ be an interval whose measure w.r.t. the standard Gaussian is δ (and thus $s \approx \sqrt{2\pi}\ \frac{\delta}{2}$). Let $\mathcal{D}_I$ and $\mathcal{D}_{\bar{I}}$ denote the distribution of $x \sim \mathcal{N}$ conditional on being $x \in I$ and $x \notin I$ respectively.

Consistency Test with Parameter δ

Given functions $\{f_S : \mathbb{R}^n \mapsto \{-1,1\} \mid S \subseteq [N], |S| = n\}$.

- Pick a set $S \subseteq [N], |S| = n$ at random.
- Pick $U \subseteq S$ by including each element of S with probability $1-\delta$. Pick $U' \subseteq S, |U'| = n/2$ at random. Pick $R \subseteq [N], |R| = n$ such that $S \cap R = U \cup (\overline{U} \cap U')$.
- Pick $x \sim \mathcal{D}_{\bar{I}}^{|U|}$. Pick $x' \sim \mathcal{D}_I^{|\overline{U} \cap U'|}$, $y^S, y^R \sim \mathcal{D}_I^{n-|U \cup (\overline{U} \cap U')|}$ independently. We think of the coordinates of x, x', y^S, y^R as indexed by the elements of U, $\overline{U} \cap U'$, $S \setminus (S \cap R)$ and $R \setminus (S \cap R)$ respectively.
- Reject if and only if
$$f_S(x, x', y^S) \neq f_R(x, x', y^R).$$

Note that the distribution of both the queries (x, x', y^S) and (x, x', y^R) is precisely $\mathcal{N}^n$. Moreover, note that (x, x', y^S) (similarly, (x, x', y^R)) does not give away $S \cap R$ ($S \cap R$ is known to contain the large coordinates, but might contain other coordinates as well). We have the following lemma regarding the rejection probability of the test when the functions f_S and f_R are indeed encodings of consistent assignments.

Lemma 6.1. *Suppose the functions f_S and f_R are encodings of assignments $\sigma(S) : S \mapsto \{-1,1\}$ and $\sigma(R) : R \mapsto \{-1,1\}$ respectively such that $\sigma(S)|_{S\cap R} = \sigma(R)|_{S\cap R} =: \pi$. Then the failure probability of the Consistency Test above is at most $O(\delta\sqrt{\delta n})$ (which is $\ll 1$ by our choice of $\delta \ll \frac{1}{n^{1/3}}$).*

Proof. We have:

$$f_S(x, x', y^S) = \mathsf{interval}\left(\sum_{i \in U} \pi(i) \cdot x_i + \sum_{i \in \overline{U} \cap U'} \pi(i) \cdot x'_i + \sum_{\ell \in S \setminus (S \cap R)} \sigma(S)(\ell) \cdot y^S_\ell\right).$$

$$f_R(x, x', y^R) = \mathsf{interval}\left(\sum_{i \in U} \pi(i) \cdot x_i + \sum_{i \in \overline{U} \cap U'} \pi(i) \cdot x'_i + \sum_{\ell \in R \setminus (S \cap R)} \sigma(R)(\ell) \cdot y^R_\ell\right).$$

Note that the sums are spread over a band of width $\Theta(\sqrt{n})$ around the origin whereas the difference between the two sums is attributed to y^S and y^R and is typically $O(\delta\sqrt{n-m}) = O(\delta\sqrt{\delta n})$ in magnitude. Therefore the two interval functions differ with probability $O(\delta\sqrt{\delta n})$. □

7 The Overall Construction

We are now ready to describe our proposed construction of the Boolean 2Lin integrality gap instance. Let Φ be the instance of $k\text{Lin}(\mathbb{R})$ with N variables and $M = TN$ constraints as in Theorem 9. We think of k as a large constant, T as a constant large enough after choosing k and N as a growing parameter. As we noted, every constraint C of Φ is a homogeneous linear equation over reals:

$$\sum_{j=1}^{k} \epsilon_j \cdot b_C[j] \cdot y_{C[j]} = 0.$$

Let $n = \zeta n_0/(t \log N)$ for n_0 as in Theorem 9, a parameter ζ that dictates the size of the construction and a parameter t which is the number of Lasserre rounds for the instance we construct. As mentioned before, in our Boolean 2Lin instance, there is a block of variables for every subset $S \subseteq [N], |S| = n$. The variables correspond to points in the space $\mathbb{R}^n$ (discretized appropriately) and an assignment to this block corresponds to a function $f_S : \mathbb{R}^n \mapsto \{-1, 1\}$. Note that the size of the construction is bounded by $2^{\zeta M}$.

Choice of Parameters: Let α, β be infinitesimally small, $\frac{1}{n^{1/2}} \ll \delta \ll \frac{1}{n^{1/3}}$.

Test: Run the following three tests with appropriate probabilities:

(1a) *Low Boundary Test with parameter* α is carried out with probability proportional to $\frac{1}{\sqrt{\alpha n}}$.

Pick a set $S \subseteq [N], |S| = n$ at random. Run the Low Boundary Test with parameter α on f_S.

(1b) *General Low Boundary Test with parameter* α is carried out with probability proportional to $\frac{1}{\sqrt{\alpha k}}$.

Pick a set $S \subseteq [N], |S| = n$ at random. Pick a constraint $C \subseteq S$ at random (note that $|C| = k$). Run the General Low Boundary Test with parameter α on subset C and f_S.

(2) *Constraint Test with Parameter* β is carried out with probability proportional to $\frac{1}{\beta/\sqrt{k}}$.

Pick a set $S \subseteq [N], |S| = n$ at random. Pick a constraint $C \subseteq S$ at random. Run the Constraint Test for constraint C and parameter β on f_S.

(3) *Consistency Test with Parameter* δ is carried out with probability proportional to $\frac{1}{\delta\sqrt{\delta n}}$.

Remark 7.1. *Note that the probability with which each test is performed is inversely proportional to the rejection probability of the test in Lemmas [4.1, 5.2], 5.1 and 6.1 respectively. These are the rejection probabilities in the "completeness case" (see below). Thus, in the completeness case, the different tests contribute equally towards the overall rejection probability. In the soundness case, it is enough to show that for any integral solution, for at least one of the tests, the rejection probability is significantly larger than that in the completeness case.*

Next we describe how the n_0-round Lasserre integrality gap for approximate real linear equations instance in Section 3 can be transformed into a feasible solution for t-rounds of Lasserre SDP for our Boolean 2Lin instance. We also analyze the objective value achieved by the Lasserre solution.

Let $\{V_{S,\alpha}\}$ be the vector solution for approximate real linear equations instance. Let us denote the vector solution for our Boolean 2Lin instance by $\{U_{T,\beta}\}$. Let T be a set of at most t variables from the Boolean 2Lin instance. Let the i^{th} variable in T correspond to a pair (S_i, x_i) where $S_i \subseteq [N], |S| = n$ and $x_i \in \mathbb{R}^n$, i.e. the variable appears in the block S_i and corresponds to the point $x_i \in \mathbb{R}^n$ in that block. Let $S = \bigcup_{i=1}^{|T|} S_i$, so that $|S| \leq t \cdot n = n_0$. The set S will be referred to as the super-block corresponding to set T. The main observation is that every assignment $\alpha \in \{-1,1\}^{|S|}$ to the variables in S induces an assignment $\alpha(T) \in \{-1,1\}^{|T|}$ for the variables in T as follows: an assignment $\alpha \in \{-1,1\}^{|S|}$ induces, by restriction, assignments $\sigma_i : S_i \mapsto \{-1,1\}$ to the blocks, which in turn induce assignments $f_{\sigma_i}(x_i)$ to the points $x_i \in \mathbb{R}^n$ via the encodings $f_{\sigma_i} : \mathbb{R}^n \mapsto \{-1,1\}$. This yields the induced assignment $\alpha(T)$ is claimed. With this observation in mind, for every $\beta \in \{-1,1\}^{|T|}$, let

$$U_{T,\beta} = \sum_{\alpha:\alpha(T)=\beta} V_{S,\alpha}.$$

For every T, by orthogonality,

$$\sum_{\beta\in\{-1,1\}^{|T|}} \|U_{T,\beta}\|_2^2 = \sum_{\beta\in\{-1,1\}^{|T|}} \sum_{\alpha:\alpha(T)=\beta} \|V_{S,\alpha}\|_2^2 = \sum_{\alpha} \|V_{S,\alpha}\|_2^2 = 1.$$

The local distributions associated with $\{U_{T,\beta}\}$ assign β to T with probability $\|U_{T,\beta}\|_2^2$. As we show below, the consistency conditions of the solution $\{U_{T,\beta}\}$ follow from the consistency conditions of the solution $\{V_{S,\alpha}\}$.

Orthogonality: Let T be a set of at most t variables from our Boolean 2Lin instance. Let $\beta_1 \in \{-1,1\}^{|T|}$, $\beta_2 \in \{-1,1\}^{|T|}$ be distinct assignments to the variables of T. Let $S \subseteq [N]$ be the super-block associated with T. For every assignments α_1, α_2 to S such that $\alpha_1(T_1) = \beta_1$ and $\alpha_2(T_2) = \beta_2$, it holds that α_1, α_2 disagree. Hence, from the feasibility of the original solution, $\langle V_{S,\alpha_1}, V_{S,\alpha_2}\rangle = 0$. Therefore,

$$\begin{aligned}
\langle U_{T,\beta_1}, U_{T,\beta_2}\rangle &= \langle \sum_{\alpha_1:\alpha_1(T)=\beta_1} V_{S,\alpha_1}, \sum_{\alpha_2:\alpha_2(T)=\beta_2} V_{S,\alpha_2}\rangle \\
&= \sum_{\alpha_1:\alpha_1(T)=\beta_1} \sum_{\alpha_2:\alpha_2(T)=\beta_2} \langle V_{S,\alpha_1}, V_{S,\alpha_2}\rangle \\
&= 0.
\end{aligned}$$

Consistency: Let T be a set of at most $t-1$ variables from the Boolean 2Lin instance. Let $p \notin T$ be an additional variable. Let $S, S^+ \subseteq [N]$ be the super-blocks associated with T and $T \cup \{p\}$ respectively. Note that either $S^+ = S$ or $S^+ = S \cup S'$ for some $S' \subseteq [N]$, $|S'| = n$. If

$S^+ = S$, then

$$\begin{aligned} U_{T,\beta} &= \sum_{\alpha:\alpha(T)=\beta} V_{S,\alpha} \\ &= \sum_{\alpha:\alpha(T\cup\{p\})=\beta\cup\{+1\}} V_{S,\alpha} + \sum_{\alpha:\alpha(T\cup\{p\})=\beta\cup\{-1\}} V_{S,\alpha} \\ &= U_{T\cup\{p\},\beta\cup\{+1\}} + U_{T\cup\{p\},\beta\cup\{-1\}}. \end{aligned}$$

If $S^+ = S \cup S'$, then

$$\begin{aligned} U_{T,\beta} &= \sum_{\alpha:\alpha(T)=\beta} V_{S,\alpha} \\ &= \sum_{\alpha:\alpha(T)=\beta} \sum_{\alpha'} V_{S\cup S',\alpha\cup\alpha'} \\ &= \sum_{\alpha:\alpha(T)=\beta} \sum_{\alpha':\alpha'(p)=+1} V_{S\cup S',\alpha\cup\alpha'} + \sum_{\alpha:\alpha(T)=\beta} \sum_{\alpha':\alpha'(p)=-1} V_{S\cup S',\alpha\cup\alpha'} \\ &= U_{T\cup\{p\},\beta\cup\{+1\}} + U_{T\cup\{p\},\beta\cup\{-1\}}. \end{aligned}$$

Non-negativity: For any sets T_1, T_2 of up to t variables and assignments β_1, β_2 to them, letting S_1, S_2 be the super-blocks associated with them, we have,

$$\begin{aligned} \langle U_{T_1,\beta_1}, U_{T_2,\beta_2} \rangle &= \langle \sum_{\alpha_1:\alpha_1(T_1)=\beta_1} V_{S_1,\alpha_1}, \sum_{\alpha_2:\alpha_2(T_2)=\beta_2} V_{S_2,\alpha_2} \rangle \\ &= \sum_{\alpha_1:\alpha_1(T_1)=\beta_1} \sum_{\alpha_2:\alpha_2(T_2)=\beta_2} \langle V_{S_1,\alpha_1}, V_{S_2,\alpha_2} \rangle \geq 0. \end{aligned}$$

Completeness: The completeness of our construction follows from the completeness of the approximate real linear equations instance in Theorem 9, and Lemmas [4.1, 5.2], 5.1 and 6.1 analyzing the boundary, constraint and consistency tests, respectively. We elaborate a bit more.

Consider a hypothetical scenario that the instance in Theorem 8 and Theorem 9 has a perfectly satisfying (global) assignment. Considering its restrictions to blocks, we have: (1) for each block S, an assignment $\sigma(S)$ such that (2) $\sigma(S)$ satisfies all the constraints C that appear inside S and (3) the assignments $\sigma(S)$ and $\sigma(R)$ for any two blocks are consistent, i.e. they agree on $S \cap R$. In this scenario, letting each function f_S to be the correct encoding $f_{\sigma(S)}$, the failure probability of all the tests is bounded as in Lemmas [4.1, 5.2], 5.1 and 6.1.

Of course, the instance in Theorem 8 and Theorem 9 is highly unsatisfiable and the scenario is impossible. Still, the main point is that the Lasserre SDP solution to the instance effectively pretends that the hypothetical scenario holds. Namely, we have (1) for each block S, a set of assignments $\sigma(S)$ (the "local" distribution is uniform on this set) such that (2) every assignment $\sigma(S)$ satisfies all the constraints C that appear inside S and (3) sampling a random assignment τ for the block $S \cup T$ and letting $\sigma(S) = \tau|_S$ and $\sigma(R) = \tau|_R$ yields assignments to blocks S and R that are consistent.

Moreover, there are vectors $V_{S,\sigma(S)}$ that satisfy all the Lasserre feasibility conditions. Now, to each block S, instead of assigning an encoding $f_{\sigma(S)} : \mathbb{R}^n \mapsto \{-1, 1\}$, we effectively assign a "vector super-position" of such encodings, informally written as

$$\sum_{\sigma(S)} f_{\sigma(S)} \cdot V_{S,\sigma(S)}.$$

To be precise, if a typical variable in the Boolean 2Lin instance is denoted by a pair (S, x), then

$$U_{(S,x),\{+1\}} = \sum_{\sigma(S):f_{\sigma(S)}(x)=+1} V_{S,\sigma(S)},$$

$$U_{(S,x),\{-1\}} = \sum_{\sigma(S):f_{\sigma(S)}(x)=-1} V_{S,\sigma(S)}.$$

We need to show that the SDP solution achieves an objective value that is same as the failure probability of the tests in Lemmas [4.1, 5.2], 5.1 and 6.1. We demonstrate this for the low boundary test (i.e. Lemmas [4.1) and the others are treated similarly.

The low boundary test (for some fixed block S) picks two points $x, y \in \mathbb{R}^n$ and rejects if $f_S(x) \neq f_S(y)$. The analogue of the rejection probability from the viewpoint of the SDP objective is

$$\|U_{\{(S,x),(S,y)\},\{+1,-1\}}\|^2 + \|U_{\{(S,x),(S,y)\},\{-1,+1\}}\|^2,$$

or more precisely, the expectation of this expression over the choice of x and y. Using the feasibility conditions, this is same as

$$\langle U_{(S,x),\{+1\}}, U_{(S,y),\{-1\}}\rangle + \langle U_{(S,x),\{-1\}}, U_{(S,y),\{+1\}}\rangle.$$

Using the expression for the vector $U_{(S,x),\{+1\}}$ and others as observed, this is same as

$$\sum_{\sigma(S):f_{\sigma(S)}(x)\neq f_{\sigma(S)}(y)} \|V_{S,\sigma(S)}\|^2.$$

Taking the expectation over the choice of x and y, the SDP objective is

$$\sum_{\sigma(S)} \|V_{S,\sigma(S)}\|^2 \cdot \Pr_{x,y}\left[f_{\sigma(S)}(x) \neq f_{\sigma(S)}(y)\right].$$

Now we observe that $\sum_{\sigma(S)} \|V_{S,\sigma(S)}\|^2 = 1$ and $\Pr_{x,y}\left[f_{\sigma(S)}(x) \neq f_{\sigma(S)}(y)\right]$ is the rejection probability of the test as in Lemma 4.1.

Soundness road map: In the next few sections we analyze the soundness of the construction for (approximate) real code juntas. We start by proving several information theoretic lemmas in Section 8. Then we analyze direct product testing in the presence of leakage in Section 9. Finally, we derive the soundness proof for (approximate) real code juntas in Section 10.

8 Information Theoretic Lemmas

8.1 Extractors

Definition 10 (Min entropy)**.** *A distribution P over a space U has min-entropy k if the maximum probability $P(u)$ over $u \in U$ is 2^{-k}.*

Definition 11 (Extractor)**.** *A bipartite graph $G = (A, B, E)$ is a (δ, ε)-extractor if for every distribution P_A over A with min-entropy at least $\log(\delta\,|A|)$ the distribution induced on Y by picking $a \sim P_A$ and a uniform neighbor of a is ε-close to uniform over B.*

Lemma 8.1. *Let $0 < \varepsilon < 1$. Consider the bipartite graph that has on one side all sets $S \subseteq [N]$, $|S| = n$, and on the other side all M constraints. A set is connected to all the constraints it contains. Then, the graph is an $((1/\varepsilon)\cdot 2^{-\varepsilon^2 n/k}, O(\varepsilon))$-extractor ($k$ is the arity of the constraints).*

8.2 Basic Facts From Information Theory

The *informational divergence* $D(Y||X)$ of two random variables Y, X over a space U is $D(Y||X) \doteq \sum_{u \in U} \Pr[Y = u] \log \frac{\Pr[Y=u]}{\Pr[X=u]}$. The convention is that $\log(0/0) = 0$, whereas if there exists $u \in U$ such that $\Pr[X = u] = 0$ and $\Pr[Y = u] > 0$, then $D(Y||X) = \infty$.

Lemma 8.2. *If $X_1, \ldots, X_n$ are independent random variables, and $Y_1, \ldots, Y_n$ are any random variables, then*

$$\sum_i D(Y_i||X_i) \leq D(Y_1, \ldots, Y_n||X_1, \ldots, X_n).$$

Lemma 8.3. *If E is an event with probability at least 2^{-d}, X is a random variable, and $X' = X|E$, then*

$$D(X'||X) \leq d.$$

Lemma 8.4.

$$|Y - X|_1 \leq \sqrt{2 \ln 2 \cdot D(Y||X)}.$$

8.3 A Couple of Lemmas

The setup for the next couple of lemmas is as follows. Let $n \geq k \geq 1$ be natural numbers. Let $d \geq 1$. Let $b_1, \ldots, b_n$ be binary i.i.d random variables. Let E be an event that depends only on $b_1, \ldots, b_n$ and such that $\Pr[E] \geq 2^{-d}$. For $S \subseteq [n]$ let $B_S = \{b_i \mid i \in S\}$ and $B'_S = \{b_i \mid i \in S\} | E$.

Lemma 8.5. *Pick $S \subseteq [n]$, $|S| = k$, uniformly at random. Then,*

$$\mathop{\mathbf{E}}_S \left[\left| B'_S - B_S \right|_1 \right] \leq \sqrt{\frac{2 \ln 2 \cdot dk}{n}}.$$

Proof. Applying Lemma 8.4 and convexity,

$$\mathop{\mathbf{E}}_S \left[\left| B'_S - B_S \right|_1 \right] \leq \mathop{\mathbf{E}}_S \left[\sqrt{2 \ln 2 \cdot D(B'_S||B_S)} \right] \leq \sqrt{2 \ln 2 \cdot \mathop{\mathbf{E}}_S \left[D(B'_S||B_S) \right]}.$$

We can pick S by first picking a partition $(S_1, \ldots, S_{n/k})$ of $[n]$ into parts of size k, and then picking one of the part of S. Hence,

$$\mathop{\mathbf{E}}_S \left[D(B'_S||B_S) \right] = \mathop{\mathbf{E}}_{S_1, \ldots, S_{n/k}} \left[\frac{k}{n} \sum_i D(B'_{S_i}||B_{S_i}) \right].$$

By linearity of expectation and Lemma 8.2,

$$\mathop{\mathbf{E}}_{S_1, \ldots, S_{n/k}} \left[\frac{k}{n} \sum_i D(B'_{S_i}||B_{S_i}) \right] \leq \frac{k}{n} \cdot \mathop{\mathbf{E}}_{S_1, \ldots, S_{n/k}} \left[D(B'_{S_1} \cdots B'_{S_{n/k}}||B_{S_1} \cdots B_{S_{n/k}}) \right].$$

By Lemma 8.3,

$$\frac{k}{n} \cdot \mathop{\mathbf{E}}_{S_1, \ldots, S_{n/k}} \left[D(B'_{S_1} \cdots B'_{S_{n/k}}||B_{S_1} \cdots B_{S_{n/k}}) \right] \leq \frac{dk}{n}.$$

The lemma follows by combining all of the above. □

Lemma 8.6. *Let V be a random variable such that $b_1, \ldots, b_n$ are independent conditioned on $V = v$ for every v. Pick $S \subseteq [n]$, $|S| = k$, uniformly at random. Then,*

$$\mathop{\mathbf{E}}_{S}\left[\mathop{\mathbf{E}}_{v|E}\left[|B_S|(V=v,E) - B_S|(V=v)|_1\right]\right] \leq \sqrt{\frac{2\ln 2 \cdot dk}{n}}.$$

Proof. By linearity of expectation, Lemma 8.5 and convexity arguments,

$$\begin{aligned}
\mathop{\mathbf{E}}_{S}\left[\mathop{\mathbf{E}}_{v|E}\left[|B_S|(V=v,E) - B_S|(V=v)|_1\right]\right] &= \mathop{\mathbf{E}}_{v|E}\left[\mathop{\mathbf{E}}_{S}\left[|B_S|(V=v,E) - B_S|(V=v)|_1\right]\right] \\
&\leq \mathop{\mathbf{E}}_{v|E}\left[\sqrt{\frac{2\ln 2 \cdot k}{n} \cdot \log\left(\frac{1}{\Pr\left[E|V=v\right]}\right)}\right] \\
&\leq \sqrt{\frac{2\ln 2 \cdot k}{n} \cdot \log\left(\mathop{\mathbf{E}}_{v|E}\left[\frac{\Pr\left[V=v\right]}{\Pr\left[E\right]\cdot\Pr\left[V=v|E\right]}\right]\right)} \\
&= \sqrt{\frac{2\ln 2 \cdot k}{n} \cdot \log\left(\frac{1}{\Pr\left[E\right]}\right)} \\
&\leq \sqrt{\frac{2\ln 2 \cdot dk}{n}}
\end{aligned}$$

□

8.4 Correlated Sampling

Correlated sampling [34, 24] is a protocol for two non-communicating players with access to shared randomness to pick a common element from a space U with high probability, where the first player has a distribution D_1 over U, the second players has a distribution D_2 over U, and $|D_1 - D_2|_1 \leq \varepsilon$. This is done by using the shared randomness to specify a sequence of pairs $(u, p) \in U \times [0, 1]$. Each player scans the sequence and outputs the first u such that p is bounded by u's probability according to its distribution. The probability that the two players output different u's is most ε.

9 Direct Product Game With Leakage

9.1 Direct Product With Nearly Identical Sets

Let U be a set of elements, n be a natural number, and $0 < \delta, q < 1$. We define the n-direct product game as follows. A verifier picks a set $S \subseteq U$, $|S| = n$, uniformly at random, and a sequence of n bits, $b_1, \ldots, b_n \in \{0, 1\}$. We have $b_i = 1$ with probability δ independently. The verifier orders the n elements in S in some arbitrary fashion. The verifier constructs a set $R \subseteq U$, $|R| = n$, by picking for every $1 \leq i \leq n$ such that $b_i = 1$ a uniform $u_i \in U$, and picking the i'th element of S for every $1 \leq i \leq n$ such that $b_i = 0$. One player receives S, and the other player receives R. Each player responds with n bits one for each element in its set. The verifier picks each element in $S \cap R$ with probability q. The players are said to q-win if the bits they assign to the elements picked by the verifier are the same. The strategy of the players is given by functions $A, B : \binom{U}{n} \to \{-1, 1\}^n$.

Observe that the players can agree on an element in their intersection with probability $1-\delta$, and can devise their assignments depending on the said element. Direct product testing theorems show that for a small random set $S_0 \subseteq U$ of elements and an assignment to them $s_0 : S_0 \to \{-1,1\}$, the players' strategy on sets that contain S_0 and on which the players agree on assignment s_0 to S_0 is mostly consistent with a single function $F_{S_0 \leftarrow s_0} : U \to \{-1,1\}$.

In this work δ is very small, and hence the two sets S, R are nearly identical. Previous work on direct products analyzed the case of δ which is a large constant, but it can be extended to the nearly identical case.

Theorem 12 (Direct Product Testing with Nearly Identical Sets [17, 36, 15])**.** *For every $0 < \eta < 1$ there exists $0 < \eta' < 1$, such that the following holds. Consider a strategy $A, B : \binom{U}{n} \to \{-1,1\}^n$ of the players that $1/2$-wins with probability at least $2^{-\delta\eta' n}$. Let $S_0 \subseteq U$, $|S_0| = \eta n$, be a random set and let $s_0 : S_0 \to \{-1,1\}$.*

There exists $F_{S_0 \leftarrow s_0} : U \to \{-1,1\}$ such that when picking uniformly $S \supseteq S_0$, $|S| = n$, such that $A(S)_{|S_0} = s_0$, the probability that for a random element $x \in S - S_0$ it holds that $A(s)(x) = F_{S_0 \leftarrow s_0}(x)$ is at least $1 - \eta$.

9.2 Direct Product With Leakage

The n-direct product game with d-leakage extends the n-direct product game. A verifier picks a set $S \subseteq U$, $|S| = n$, uniformly at random, and a sequence of n bits, $b_1, \ldots, b_n \in \{0,1\}$. We have $b_i = 1$ with probability δ independently. The verifier also picks uniformly a set $T \subseteq [n]$, $|T| = n/2$. The verifier picks a random order on the elements and sorts the n elements in S. The verifier constructs $R \subseteq U$, $|R| = n$, by picking for every $i \in T$ such that $b_i = 1$ a uniform $u_i \in U$, and picking the i'th element of S for every $1 \leq i \leq n$ such that $b_i = 0$ or $i \notin T$. The first player receives S, and the second player receives R. In addition to S, R, the players are provided with additional information in the form of leakage y that depends on $b_1, \ldots, b_n$ but does not depend on T. The min-entropy of y is at most d. It is guaranteed that the distribution of $b_1, \ldots, b_n$ conditioned on y and S is the same as the distribution of $b_1, \ldots, b_n$ conditioned on y and R. Each player responds with n bits, one for each element in its set. The verifier picks each elements in $S \cap R$ with probability q. The players are said to q-win if the bits they assign to the elements picked by the verifier are the same.

Even with leakage, no player knows for sure which of the n elements in its set also appears in the other player's set. Each player may have information on which elements are more likely to appear in the other player's set, as well as on which elements definitely do not appear in the other player's set.

We'll show that players that win the n-direct product game with d-leakage also win a certain $\Omega(n)$-direct product game.

Lemma 9.1. *A strategy that q-wins an n-direct product game with d-leakage with probability at least p has a sub-strategy that q-wins the k-direct product game with probability at least $p - O(\varepsilon)$, where $\varepsilon = \sqrt{\frac{2 \ln 2 \cdot dk}{n/2-k}}$.*

Proof. We obtain a strategy for the k-direct product game from a strategy for the n-direct product game with d-leakage by presenting an embedding of the k-direct product game inside the n-direct product game with d-leakage. In the sequel S and R are the sets in the n-direct product game with d-leakage and y is the leakage.

- The players use shared randomness to fix a value to y. Let E be the event that y is as fixed. Note that $\Pr[E] \geq 2^{-d}$.
- The players pick uniformly at random $I \subseteq \{1, \ldots, n\}$, $|I| = k$. Let S_I, R_I be the elements of S, R in positions I, respectively. By Lemma 8.5, the expected distance between $S_I, R_I|E$ and the elements picked in the k-direct product game is bounded by ε.
- By Markov inequality, for a large fraction of the I's, we have that $S_I, R_I|E$ are $O(\varepsilon)$-close to the distribution in a k-direct product. We will focus on such I's. Each player embeds the k elements from the k-direct product game in positions I.

In the sequel we will further restrict our choice of I, each time maintaining the probability that an appropriate I exists positive. It remains to show that the players of a k-direct product game can jointly sample the rest of the sets for the n-direct product game with leakage. Crucially, the elements in positions $\{1, \ldots, n\} - I$ may be dependent on the elements in positions I given E. In particular, $\{b_i \mid i \in I\}$, which is determined by the elements that the players so far picked, is unknown to any of the players, yet might be dependent on $\{b_i \mid i \notin I\}$ given E. This may result in correlations within $\{b_i \mid i \notin I\}$ given what was so far decided that are unknown to any of the players. To the rescue comes T that picks at random half of the positions as definitely common to the two players and "breaks" the unknown correlations.

- The players use shared randomness to agree on $T - I$. The set $T - I$ contains all the indices outside I where S and R do not share common variables (and other indices as well).
- Recall that we are guaranteed that $\{b_i \mid i \in T - I\}|(S_I, E)$ is distributed the same as $\{b_i \mid i \in T - I\}|(S_I, R_I, E)$ and as $\{b_i \mid i \in T - I\}|(R_I, E)$.
- By Lemma 8.6, we have that $\{b_i \mid i \in T - I\}|(S_I, R_I, E)$ is ε-close in expectation to $\{b_i \mid i \in T - I\}$. We focus on I such that there is $O(\varepsilon)$-closeness. The players can use correlated sampling to agree on $\{b_i \mid i \in T - I\}$ conditioned on all their choices so far.

Next the players agree on the elements in positions $\{1, \ldots, n\} - I$. As we mentioned above, the elements in positions $\{1, \ldots, n\} - I$ may depend on the elements in positions I given E. The key observation is that each of the players has essentially all the information about the elements in their set in positions $\{1, \ldots, n\} - I$.

- Let $R_{\overline{I}}, S_{\overline{I}}$ be the elements in positions $\{1, \ldots, n\} - I$ in R, S, respectively. By Lemma 8.6, $R_I|(S_I, S_{\overline{I}}, E)$ is ε-close in expectation to $R_I|S_I$. We focus on I's in which $O(\varepsilon)$-closeness holds.
- Hence, $S_{\overline{I}}|(S_I, R_I, E)$ is $O(\varepsilon)$-close in expectation to $S_{\overline{I}}|(S_I, E)$:

$$\mathop{\mathbf{E}}_{s_{\overline{I}}, s_I, r_I} \left[\Pr\left[S_{\overline{I}} = s_{\overline{I}} | S_I = s_I \wedge R_I = r_I, E \right] \right]$$

$$
\begin{aligned}
&= \mathop{\mathbf{E}}_{s_{\overline{I}},s_I,r_I}\left[\frac{\Pr\left[S_{\overline{I}}=s_{\overline{I}}\wedge S_I=s_I\wedge R_I=r_I|E\right]}{\Pr\left[S_I=s_I\wedge R_I=r_I|E\right]}\right]\\
&= \mathop{\mathbf{E}}_{s_{\overline{I}},s_I,r_I}\left[\frac{\Pr\left[R_I=r_I|S_{\overline{I}}=s_{\overline{I}}\wedge S_I=s_I,E\right]\cdot\Pr\left[S_{\overline{I}}=s_{\overline{I}}\wedge S_I=s_I|E\right]}{\Pr\left[S_I=s_I\wedge R_I=r_I|E\right]}\right]\\
&\leq \mathop{\mathbf{E}}_{s_{\overline{I}},s_I,r_I}\left[\frac{\Pr\left[R_I=r_I|S_I=s_I\right]}{\Pr\left[S_I=s_I\wedge R_I=r_I|E\right]}\cdot\Pr\left[S_{\overline{I}}=s_{\overline{I}}\wedge S_I=s_I|E\right]\right]+O(\varepsilon)\\
&= \mathop{\mathbf{E}}_{s_{\overline{I}},s_I,r_I}\left[\frac{\Pr\left[R_I=r_I\wedge S_I=s_I\right]}{\Pr\left[S_I=s_I\wedge R_I=r_I|E\right]}\cdot\frac{\Pr\left[S_{\overline{I}}=s_{\overline{I}}\wedge S_I=s_I|E\right]}{\Pr\left[S_I=s_I\right]}\right]+O(\varepsilon)\\
&\leq \mathop{\mathbf{E}}_{s_{\overline{I}},s_I,r_I}\left[\Pr\left[S_{\overline{I}}=s_{\overline{I}}|S_I=s_I,E\right]\right]+O(\varepsilon)
\end{aligned}
$$

The first inequality follows since $R_I|S_I, S_{\overline{I}}, E$ is $O(\varepsilon)$-close to $R_I|S_I$. The second inequality follows since $R_I, S_I|E$ is $O(\varepsilon)$-close to R_I, S_I.

- This means that the first player, who knows S_I and E, but does not know R_I, can approximately sample $S_{\overline{I}}$. Similarly, $R_{\overline{I}}|(R_I, S_I, E)$ is $O(\varepsilon)$-close to $R_{\overline{I}}|(R_I, E)$, so the second player, who knows R_I and E, but does not know S_I, can approximately sample $R_{\overline{I}}$. Moreover, the players can jointly sample $S_{\overline{I}}, R_{\overline{I}}$ (recall that $S_{\overline{I}}, R_{\overline{I}}$ are correlated). The players already agreed which elements they have in common. Therefore, they can use correlated sampling to pick $R_{\overline{I}}$ and $S_{\overline{I}}$.

□

10 Soundness for Approximate Real Code Juntas

We consider strategies that correspond to a list decoding of the real code. Recall that the real code consists of the functions

$$f_\sigma(x) = \mathsf{interval}\left(\sum_{i=1}^{n}\sigma(i)\cdot x_i\right).$$

The f_σ's all satisfy the folding constraints and pass the low boundary test with high probability $1-O(\sqrt{\alpha n})$. However, these are not the only functions with this property. For three different $\sigma_1,\sigma_2,\sigma_3\in\{-1,1\}^n$, take $f_{\sigma_1}\oplus f_{\sigma_2}\oplus f_{\sigma_3}$. This function too satisfies folding, and the probability that the low boundary test rejects it is at most three times the probability that it rejects a single linear interval function. Note that unlike in the standard coding theoretic sense, $f_{\sigma_1}\oplus f_{\sigma_2}\oplus f_{\sigma_3}$ has no correlation with any single linear interval function. In this section we handle functions like $f_{\sigma_1}\oplus f_{\sigma_2}\oplus f_{\sigma_3}$ that correspond to a "list decoding" of the real code.

Definition 13. *An (l,γ)-list decoding strategy is as follows. For at least $1-\gamma$ fraction of the sets $S\subseteq[N]$, $|S|=n$, there are is a real code junta $J_S:\mathbb{R}^n\to\{-1,1\}$ depending on at most l real code functions, such that $\Pr_x\left[f_S(x)=J_S(x))\right]\geq 1-\gamma$.*

Remark 10.1. *Without loss of generality, we assume that for every S and $1\leq i\neq j\leq l$, if $f_{\sigma_{S,i}}$ and $f_{\sigma_{S,j}}$ are the i'th and the j'th real code functions in J_S, then the vectors $\sigma_{S,i}$ and $\sigma_{S,j}$ disagree on at least $\Omega(\gamma/l^2)$ fraction of the coordinates (if this is not the case, one can remove one of $\sigma_{S,i},\sigma_{S,j}$, adapt G_S to use the other instead, and introduce only $O(\gamma)$ approximation error overall).*

The outcome of $f_{\sigma_{S,1}}(x), \ldots, f_{\sigma_{S,l}}(x)$ can give information on which are the small coordinates of x (e.g., if $x_1 \geq x_2 \geq x_3 \geq \cdots \geq 0$, then it's quite likely that x_1 is large). The identity of the small coordinates leaks information about which R's could have been picked in the consistency test together with S. We use our analysis of the direct product game with leakage to show that for a random Φ the rejection probability for any (l, γ)-list decoding strategy is much higher than the rejection probability in the completeness case.

Theorem 14 (Soundness for list decoding strategies)**.** *Let $0 < \gamma < 1/4$ be a constant and let $l \geq 1$ be a constant. Assume $\beta \ll 1$. Let $\delta\sqrt{\delta} \ll \epsilon \ll \delta$. With high probability over Φ, for any (l, γ)-list decoding strategy, at least one of the two holds:*

1. *The consistency test rejects with probability at least $\Omega(\sqrt{\epsilon n})$ (where the rejection probability in the completeness case is $O(\delta\sqrt{\delta n})$).*
2. *The constraint test rejects with probability at least $\Omega(\beta)$ (where the rejection probability in the completeness case is $O(\beta/\sqrt{k})$).*

Proof. Fix an assignment to our gap instance such that the probability that the consistency test rejects is $o(\sqrt{\epsilon n})$. One can view the consistency test as comprising an outer verifier that picks sets S, R as in the n-direct product game, and an inner verifier that picks queries (x, x', y^S) and (x, x', y^R) to f_S, f_R, respectively. Except with probability 2γ there are list decodings $\sigma_{S,1}, \ldots, \sigma_{S,l}$ and $\sigma_{R,1}, \ldots, \sigma_{R,l}$. Let the leakage specify the information that $f_{\sigma_{S,1}}, \ldots, f_{\sigma_{S,l}}$ and $f_{\sigma_{R,1}}, \ldots, f_{\sigma_{R,l}}$ convey on the identity of the small coordinates. The min-entropy of the leakage is bounded by a constant since l and γ are constants (see Remark 10.1). Let the answers $A(S), B(R)$ of the players be random $\sigma_{S,i}, \sigma_{R,j}$ respectively from each of their lists.

Suppose that for every $1 \leq i, j \leq l$ it holds that $\sigma_{S,i}, \sigma_{R,j}$ disagree on at least ϵn elements in the intersection $S \cap R$. In this case, since $\delta\sqrt{\delta} \ll \epsilon$, the consistency test rejects with probability at least $\Omega(\sqrt{\epsilon n})$. By our assumption, this happens with probability at most $o(1)$.

In contrast, if there exists $1 \leq i, j \leq l$ such that $\sigma_{S,i}, \sigma_{R,j}$ disagree on at most ϵn elements in the intersection $S \cap R$, then the probability of 1/2-win in the direct product game is at least $(1/l^2) \cdot 2^{-\epsilon n} \geq 2^{-o(\delta n)}$. Overall, the probability of 1/2-win is at least $2^{-o(\delta n)}$. By Lemma 9.1 and Theorem 12, for a random set $S_0 \subseteq [N]$ of fraction $|S_0|/n$ that is sufficiently small with respect to γ, $1/l$ and k the arity of the constraints, and for $s_0 : S_0 \to \{-1, 1\}$ there exists $F_{S_0 \leftarrow s_0} : [N] \to \{-1, 1\}$ such that when picking uniformly $S \supseteq S' \supseteq S_0$, $|S| = n$, $|S'| = \Omega(n)$ (where the constant in the $\Omega(\cdot)$ depends on l and γ), with $A(S)_{|S_0} = s_0$, we have $A(s)(x) = F_{S_0 \leftarrow s_0}(x)$ for at least $1 - c/100k$ fraction of $x \in S' - S_0$ (where c is a small constant; see below). We focus on S, S', S_0 and s_0 such that the probability that the constraint test fails conditioned on $S \supseteq S' \supseteq S_0$, $A(S)_{|S_0} = s_0$ and the constraint being contained in $S' - S_0$, is at most a constant times larger than the probability the constraint test fails in general. Let $UNSAT \subseteq [M]$ specify those real linear equations whose margin under the assignment $F_{S_0 \leftarrow s_0}$ is at least $c \cdot \sqrt{k}$. We know that $|UNSAT| \geq c \cdot M$ with high probability. Let us focus on this event. By Lemma 8.1, when picking uniformly $S \supseteq S' \supseteq S_0$ as above, and a constraint in $S' - S_0$, there is a constant probability that all of the following occur:

- The constraint is in $UNSAT$.
- S has a list decoding $f_{\sigma_{S,1}}, \ldots, f_{\sigma_{S,l}}$, and $A(S) = \sigma_{S,i}$ for a uniform $i \in \{1, \ldots, l\}$.
- The $A(S)$ assignment to all the variables in the constraint agrees with $F_{S_0 \leftarrow s_0}$.

When the above items hold, the constraint test rejects $f_{\sigma_{S,i}}$ with probability at least $\Omega(\beta)$. Since $\beta \ll 1$, and $\sigma_{S,1}, \ldots, \sigma_{S,l}$ are sufficiently far apart, it is likely that only one of $f_{\sigma_{S,1}}, \ldots, f_{\sigma_{S,l}}$ changes as a result of the shift of the constraint test. Hence, the constraint test rejects f_S with probability at least $\Omega(\beta)$. Overall, the probability that the constraint test rejects is at least $\Omega(\beta)$. □

Acknowledgements

Many thanks to Madhur Tulsiani and Pratik Worah for discussions that led to the Lasserre integrality gap for approximate real linear equations problem. Thanks to everyone with whom we discussed related issues over the years. Thanks to Boaz Barak for suggestions that improved the presentation.

References

[1] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM*, 45(3):501–555, 1998.

[2] S. Arora and S. Safra. Probabilistic checking of proofs: a new characterization of NP. *Journal of the ACM*, 45(1):70–122, 1998.

[3] B. Barak, F. G. S. L. Brandão, A. W. Harrow, J. A. Kelner, D. Steurer, and Y. Zhou. Hypercontractivity, sum-of-squares proofs, and their applications. In *Proc. 44th ACM Symp. on Theory of Computing*, pages 307–326, 2012.

[4] B. Barak, P. Gopalan, J. Håstad, R. Meka, P. Raghavendra, and D. Steurer. Making the long code shorter. In *Proc. 53rd IEEE Symp. on Foundations of Computer Science*, pages 370–379, 2012.

[5] B. Barak, M. Hardt, I. Haviv, A. Rao, O. Regev, and D. Steurer. Rounding parallel repetitions of unique games. In *Proc. 49th IEEE Symp. on Foundations of Computer Science*, pages 374–383, 2008.

[6] B. Barak, P. Raghavendra, and D. Steurer. Rounding semidefinite programming hierarchies via global correlation. In *Proc. 52nd IEEE Symp. on Foundations of Computer Science*, pages 472–481, 2011.

[7] M. Bellare, O. Goldreich, and M. Sudan. Free bits, PCPs, and nonapproximability—towards tight results. *SIAM Journal on Computing*, 27(3):804–915, 1998.

[8] M. Blum, M. Luby, and R. Rubinfeld. Self-testing/correcting with applications to numerical problems. *Journal of Computer and System Sciences*, 47(3):549–595, 1993.

[9] C. Borell. The brunn-minkowski inequality in gauss space. *Inventiones mathematicae*, 30(2):207–216, 1975.

[10] C. Borell. Geometric bounds on the ornstein-uhlenbeck velocity process. *Z. Wahrsch. Verw. Gebiete*, 70(1):1–13, 1985.

[11] E.A. Carlen and C. Kerce. On the cases of equality in bobkov's inequality and gaussian rearrangement. *Calculus of Variations and Partial Differential Equations*, 13(1):1–18, 2001.

[12] S. O. Chan. Approximation resistance from pairwise independent subgroups. In *Proc. 45th ACM Symp. on Theory of Computing*, pages 447–456, 2013.

[13] M. Charikar, K. Makarychev, and Y. Makarychev. Near-optimal algorithms for maximum constraint satisfaction problems. *ACM Transactions on Algorithms*, 5(3), 2009.

[14] A. Cianchi, N. Fusco, F. Maggi, and A. Pratelli. On the isoperimetric deficit in gauss space. *American Journal of Mathematics*, 133(1):131–186, 2011.

[15] I. Dinur. Generalizing Dinur-Steurer to nearly identical sets. private communication, 2015.

[16] I. Dinur and V. Guruswami. PCPs via low-degree long code and hardness for constrained hypergraph coloring. In *Proc. 54th IEEE Symp. on Foundations of Computer Science*, pages 340–349, 2013.

[17] I. Dinur and D. Steurer. Direct product testing. In *Computational Complexity Conference*, 2014.

[18] R. O'Donnell E. Mossel and K. Oleszkiewicz. Noise stability of functions with low influences invariance and optimality. In *Annals of Mathematics*, volume 171, pages 295–341, 2010.

[19] A. Erhard. Elements extremaux pour les inegalites de brunn-minkowski gaussiennes. *Ann. Inst. H. Poincare*, (22):149–168, 1986.

[20] U. Feige and L. Lovász. Two-prover one-round proof systems: Their power and their problems (extended abstract). In *Proc. 24th ACM Symp. on Theory of Computing*, pages 733–744, 1992.

[21] D. Grigoriev. Linear lower bound on degrees of Positivstellensatz calculus proofs for the parity. *Theoret. Comput. Sci.*, 259(1-2):613–622, 2001.

[22] V. Guruswami and A.K. Sinop. Lasserre hierarchy, higher eigenvalues, and approximation schemes for quadratic integer programming with PSD objectives. In *Proc. 52nd IEEE Symp. on Foundations of Computer Science*, pages 482–491, 2011.

[23] J. Håstad. Some optimal inapproximability results. *Journal of the ACM*, 48(4):798–859, 2001.

[24] T. Holenstein. Parallel repetition: Simplification and the no-signaling case. *Theory of Computing*, 5(1):141–172, 2009.

[25] D. M. Kane. The gaussian surface area and noise sensitivity of degree-d polynomial threshold functions. *computational complexity*, 20(2):389–412, 2011.

[26] S. Khot. Improved inapproximability results for maxclique, chromatic number and approximate graph coloring. In *Proc. 42nd IEEE Symp. on Foundations of Computer Science*, pages 600–609, 2001.

[27] S. Khot. On the power of unique 2-prover 1-round games. In *Proc. 34th ACM Symp. on Theory of Computing*, pages 767–775, 2002.

[28] S. Khot. On the unique games conjecture (invited survey). In *IEEE Conference on Computational Complexity*, pages 99–121, 2010.

[29] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell. Optimal inapproximability results for MAX-CUT and other two-variable CSPs? *SIAM Journal on Computing*, 37(1):319–357, 2007.

[30] S. Khot and D. Moshkovitz. NP-hardness of approximately solving linear equations over reals. In *STOC*, pages 413–420, 2011.

[31] S. Khot and S. Safra. A two prover one round game with strong soundness. In *Proc. 52nd IEEE Symp. on Foundations of Computer Science*, pages 648–657, 2011.

[32] S. Khot and R. Saket. SDP integrality gaps with local ℓ_1-embeddability. In *Proc. 50th IEEE Symp. on Foundations of Computer Science*, pages 565–574, 2009.

[33] S. Khot and N. K. Vishnoi. The unique games conjecture, integrality gap for cut problems and embeddability of negative type metrics into l_1. In *Proc. 46th IEEE Symp. on Foundations of Computer Science*, pages 53–62, 2005.

[34] J. Kleinberg and E. Tardos. Approximation algorithms for classification problems with pairwise relationships: metric labeling and markov random fields. *Journal of the ACM*, 49(5):616–639, 2002.

[35] A. Klivans, R. O'Donnell, and R. Servedio. Learning geometric concepts via gaussian surface area. In *Proc. 49th IEEE Symp. on Foundations of Computer Science*, pages 541–550, 2008.

[36] D. Moshkovitz. Direct product testing with nearly identical sets. Technical Report TR14-182, ECCC, 2014.

[37] E. Mossel and J. Neeman. Robust dimension free isoperimetry in gaussian space. *Annals of Probability*, 43(3):971–991, 2015.

[38] E. Mossel and J. Neeman. Robust optimality of gaussian noise stability. *Journal of the European Math Society (JEMS)*, 17(2):433–482, 2015.

[39] P. Raghavendra. Optimal algorithms and inapproximability results for every csp? In *Proc. 40th ACM Symp. on Theory of Computing*, pages 245–254, 2008.

[40] P. Raghavendra and D. Steurer. Integrality gaps for strong SDP relaxations of unique games. In *Proc. 50th IEEE Symp. on Foundations of Computer Science*, pages 575–585, 2009.

[41] A. Rao. Parallel repetition in projection games and a concentration bound. *SIAM Journal on Computing*, 40(6):1871–1891, 2011.

[42] R. Raz. A parallel repetition theorem. In *SIAM Journal on Computing*, volume 27, pages 763–803, 1998.

[43] A. Samorodnitsky and L. Trevisan. A PCP characterization of NP with optimal amortized query complexity. In *Proc. 32nd ACM Symp. on Theory of Computing*, pages 191–199, 2000.

[44] A. Samorodnitsky and L. Trevisan. Gowers uniformity, influence of variables, and PCPs. *SIAM Journal on Computing*, 39(1):323–360, 2009.

[45] G. Schoenebeck. Linear level lasserre lower bounds for certain k-CSPs. In *Proc. 49th IEEE Symp. on Foundations of Computer Science*, pages 593–602, 2008.

[46] L. Trevisan. Approximation algorithms for unique games. *Theory of Computing*, 4(1):111–128, 2008.

[47] L. Trevisan. On Khot's unique games conjecture. *Bull. Amer. Math. Soc. (N.S.)*, 49(1):91–111, 2012.

[48] M. Tulsiani. CSP gaps and reductions in the lasserre hierarchy. In *Proc. 41st ACM Symp. on Theory of Computing*, pages 303–312, 2009.