Approximating CVP to within Almost-Polynomial Factors is NP-hard

I. Dinur * G. Kindler * R. Raz † S. Safra *

Abstract

This paper shows the problem of finding the closest vector in an n-dimensional lattice to be NP-hard to approximate to within factor $n^{c/\log\log n}$ for some constant $c > 0$.

1 Introduction

An n-dimensional lattice $L = L(v_1, .., v_n)$, for linearly independent vectors $v_1, .., v_n \in R^k$ is the additive group generated by the vectors, i.e. the set $L = \{\sum a_i v_i \mid a_i \in \mathbb{Z}\}$. Given L and an arbitrary vector y, the Closest Vector Problem (CVP) is to find a vector in L closest to y in a certain norm. The Shortest Vector Problem (SVP) is a homogeneous analog of CVP, and is defined to be the problem of finding the shortest non-zero vector in L.

These lattice problems have been introduced in the 19th century, and have been studied since. Minkowsky and Dirichlet tried, with little success, to come up with approximation algorithms for these problems. It was much later that the lattice reduction algorithm was presented by Lenstra, Lenstra and Lovász [LLL82] , achieving a polynomial-time algorithm approximating the Shortest Lattice Vector to within the exponential factor $2^{n/2}$, where n is the dimension of the lattice. Babai [Bab86] applied LLL's methods to present an algorithm that approximates CVP to within a similar factor. Schnorr [Sch85] improved on LLL's technique, reducing the factor of approximation to $(1+\varepsilon)^n$, for both CVP and SVP, where the polynomial running time depends on $\frac{1}{\varepsilon}$ in the exponent. These positive approximation results are quite weak, achieving only extremely large (exponential) factors. The question naturally arises: What are the factors of approximation to within which these problems can be approximated in polynomial time?

Interest in lattice problems has been renewed due to a result of Ajtai [Ajt96], showing a reduction, from a version of SVP, to the *average-case* of the same problem.

CVP was shown to be NP-hard for any l_p norm in [vEB81], where it was also conjectured that SVP is NP-hard. Arora et al. [ABSS93] utilized the PCP characterization of NP to show that CVP is NP-hard to approximate to within any constant, and quasi-NP-hard to approximate to within $2^{(\log n)^{1-\varepsilon}}$ for any constant $\varepsilon > 0$ (i.e. an approximation algorithm for such factors would imply $NP \subseteq \mathrm{DTIME}(2^{polylogn})$.

*Tel-Aviv University

†Weizmann Institute of Science

As for SVP, it is NP-hard to approximate to within [Ajt98, Mic98] some constant factor (see also [CN98]). The proof in [Mic98] relies on the PCP characterization of NP and is carried out via a reduction from gap-CVP (shown NP-hard for any constant gap in [ABSS93]). Using gap-CVP allows, in addition to the significant improvement in the hardness-of-approximation factor, a major simplification of the main technical lemma from [Ajt98]. Better hardness results for gap-CVP may result in improved approximation hardness results for SVP.

So far there is still a huge gap between the positive results, showing approximations for these problems with exponential factors, and the above hardness results. Nevertheless, some other results provide a discouraging indication for improving the hardness result beyond a certain factor. [LLS90] showed that approximating CVP to within $n^{1.5}$ is in co-NP, and later [GG98] showed that approximating both SVP and CVP to within $\sqrt{n}$ is in NP$\cap$ co-AM. Hence it is unlikely for any of these problems to be NP-hard.

The strongest NP-hardness result likely to be true for these problems, hence, is that they are NP-hard to approximate to within a constant power of the dimension.

Our Results. We improve on [ABSS93] in two ways. First, we go beyond the factor of $2^{(\log n)^{1-\varepsilon}}$ for any constant $\varepsilon > 0$, which was the previous hardness-of-approximation factor known for CVP. Instead, we achieve a factor of $2^{\frac{\log n}{\log \log n}} = n^{\frac{1}{\log \log n}}$. Furthermore, we show approximating CVP is *NP-hard* for these large factors, compared to the previously known *quasi* NP-hardness.

The known PCP characterizations of NP seem inadequate in order to show hardness of approximating CVP to within large factors. The proof of [ABSS93] utilizes amplification techniques, in which the dimension of the instance grows faster than the factor for which approximation hardness is obtained. It is therefore unlikely that using this technique, even if allowing a super-polynomial blow-up, one can obtain such strong results. It seems that with this method it will always be the case that the factor for which hardness of approximation is proven never reaches beyond the barrier of $2^{(\log n)^{1-\varepsilon}}$ for any constant $\varepsilon > 0$.

We introduce a new NP-hard gap-problem, Super-SAT ($\mathcal{SSAT}$ for short), that we use to prove our result. The $\mathcal{SSAT}$ problem is a gap version of SAT, minimizing a new, appropriately defined, objective function. Although the $\mathcal{SSAT}$ characterization differs from the PCP characterizations, its proof relies on similar techniques.

Let SAT[F] be the following problem: An instance of SAT[F] is a set of local-constraints (Boolean functions) called *tests*, on variables from a common set, each variable ranging over a finite set F. Each test is represented by a list of assignments for its variables, which are said to satisfy the test. The goal is to attach to each test one of the assignments that satisfies it, such that consistency is maintained among the assignments, that is, each variable is given the same value by the assignments of all tests that depend on it. If this is possible, the instance is accepted, and otherwise it is rejected.

Our gap version of this problem, $\mathcal{SSAT}$, is as follows: $\mathcal{SSAT}$ is the same as SAT[F] except not all non-satisfiable instances must be rejected. We generalize the notion of assignment to that of super-assignment – formal linear combinations of assignments with integer coefficients – and modify the acceptance condition accordingly: Previously accepted instances must still be accepted. An instance must be rejected only if there is no super-assignment to the tests, whose norm (see Definition 2) is smaller than g, and which is "everywhere consistent" (in a sense similar to that described above). If the instance is somewhere in-between (i.e. minimizing the norm of its consistent super-assignments gives a value greater than 1 but less than g), then that instance is not necessarily rejected (any outcome is ok).

We show (Theorem 1) that solving this problem is NP-hard for $g = n^{1/\log\log n}$ (n denotes, as usual, the size of the instance). We then reduce this problem to CVP, preserving the approximation factor. Improving the hardness of approximation factor of $\mathcal{SSAT}$ to a constant power of n, namely where $g = n^{\varepsilon}$ for some constant ε (Conjecture 2), would directly imply CVP to be NP-hard to approximate to within a constant power of the dimension.

For simplicity, our proof works with l_1 norm, however it can be extended to l_p norm for any $1 < p < \infty$ as shown in Subsection 6.3.

Outline. We begin, in Section 2, by presenting the new NP-hard gap-problem, $\mathcal{SSAT}$. We first formally define $\mathcal{SSAT}$ and then state Theorem 1 asserting it is NP-hard to approximate to within large factors of approximation ($n^{1/\log\log n}$). Section 3 gives some definitions and techniques which are the basis of the construction. The NP-hardness of $\mathcal{SSAT}$, that is the most technical part of this work, is established in two parts. In Section 4, we describe the reduction from a low error-probability PCP characterization of NP, to $\mathcal{SSAT}$. We proceed to prove the correctness of the reduction (Theorem 1) in Section 5. In Finally, in Section 6 we show a simple reduction from $\mathcal{SSAT}$ to CVP.

2 Super-SAT - $\mathcal{SSAT}$

In this section we introduce a new NP-hard problem, $\mathcal{SSAT}$. Let us begin by defining SAT[$\mathcal{F}$], which is actually SAT over non-Boolean variables, presented from a different point of view. An instance of SAT[$\mathcal{F}$]

$$I = \langle \Psi = \{\psi_1, .., \psi_n\}, V = \{v_1, .., v_m\}, \{\mathcal{R}_{\psi_1}, .., \mathcal{R}_{\psi_n}\}\rangle$$

is a set Ψ of *tests* (Boolean functions) over a common set V of variables that take values in a field $\mathcal{F}$. In what follows $|\mathcal{F}|$, m, and $|\mathcal{R}_{\psi_i}|$ will always be bounded by a polynomial in $n = |\Psi|$. Each test $\psi \in \Psi$ has associated with it a list $\mathcal{R}_\psi$ of assignments to its variables, called the *satisfying assignments* or the *range* of the test ψ. Having both ψ and $\mathcal{R}_\psi$ is convenient yet somewhat redundant since the list $\mathcal{R}_\psi$ actually specifies all there is to know about the test ψ.

An *assignment* for an instance maps to each *test*, a satisfying assignment from its range. An instance is accepted iff there is an assignment to the tests that is everywhere consistent, that is, each variable is given the same value by the assignments to all tests that depend on it. It is easy to see that SAT[$\mathcal{F}$] is NP-complete.

$\mathcal{SSAT}$ is a gap variant of this problem, obtained by setting a new measure on the non-satisfiability of an instance. While in PCP we measured the fraction of tests, satisfiable by a single assignment, in $\mathcal{SSAT}$ we will define a measure of a different nature - we will introduce a notion of super-assignments to the tests, that is, formal linear combinations of assignments. We will then measure the 'length' of a super-assignment, and ask how 'short' it may get while maintaining 'consistency'.

Definition 1 (Super-Assignment to Tests) *A super-assignment is a function S mapping each $\psi \in \Psi$ to a value from $\mathbb{Z}^{\mathcal{R}_\psi}$. $S(\psi)$ is a vector of integer coefficients, one for each value $r \in \mathcal{R}_\psi$. Denote by $S(\psi)[r]$ the r^{th} coordinate of $S(\psi)$.*

If $S(\psi)[r] \neq 0$, we say that the value r *appears* in $S(\psi)$. A *natural super-assignment* assigns each $\psi \in \Psi$ a unit vector $e_i \in \mathbb{Z}^{\mathcal{R}_\psi}$ with a 1 in a single coordinate i corresponding to an assignment for that test in the usual sense (i.e. an assignment which maps $r \in \mathcal{R}_\psi$ to ψ corresponds to the natural super-assignment $S(\psi)$ such that $S(\psi)[r] = 1$ and $S(\psi)[r'] = 0$ for all $r' \neq r$). We use the average over the l_1 norms of the vectors $S(\psi)$, $\|S(\psi)\|$, to measure the closeness of S to a natural super-assignment,

Definition 2 (Norm of a Super-Assignment) *The norm of a super-assignment S is the average norm of its individual assignments $\|S\| = \frac{1}{|\Psi|}\sum_{\psi\in\Psi}\|S(\psi)\|$, where $\|S(\psi)\|$ denotes the l_1 norm of the vector $S(\psi)$.*

The norm of a natural super-assignment is 1. The gap of $\mathcal{SSAT}$ will be formulated in terms of the norm of the minimal super-assignment that maintains consistency. A natural assignment $r \in \mathcal{R}_\psi$ to a test ψ induces an assignment to each variable x, denoted $r|_x$. In the SAT[$\mathcal{F}$] problem an assignment is called consistent if for every pair of tests with a common variable, the assignments to the tests, restricted to the variable, are equal. We extend this notion of consistency to super-assignments by defining the projection of a super-assignment $S(\psi)$ onto each of ψ's variables. Consistency between tests will amount to equality of projections on common variables.

Definition 3 (Projection) *Let $S : \Psi \to \bigcup_\psi \mathbb{Z}^{\mathcal{R}_\psi}$ be a super-assignment to the tests. We define the projection of $S(\psi)$ on a variable x of ψ, $\pi_x(S(\psi)) \in \mathbb{Z}^{\mathcal{F}}$, as follows:*

$$\forall a \in \mathcal{F} : \qquad \pi_x(S(\psi))[a] \stackrel{def}{=} \sum_{r\in\mathcal{R}_\psi,\, r|_x = a} S(\psi)[r]$$

Namely, we partition the assignments in $\mathcal{R}_\psi$ according to their value $a \in \mathcal{F}$ on the variable x (we associate with $a \in \mathcal{F}$ all assignments $r \in \mathcal{R}_\psi$ for which $r|_x = a$). For each value $a \in \mathcal{F}$, we then add the coefficients $S(\psi)[r]$ of the assignments associated with it, and this is the value of the coefficient $\pi_x(S(\psi))[a]$.

We shall now proceed to define the notion of consistency between tests. If the projections of two tests on each common variable x are equal (in other words, they both give x the same super-assignment), we say that the super-assignments of the tests are consistent.

Definition 4 (Consistency) *Let S be a super-assignment to the tests in Ψ. S is consistent if for every pair of tests ψ_i and ψ_j with a common variable x,*

$$\pi_x(S(\psi_i)) = \pi_x(S(\psi_j))$$

S is said to be *non-trivial* if every variable $x \in V$ there is at least one test $\psi \in \Psi$ that isn't 'cancelled' on x: $\pi_x(S(\psi)) \neq \vec{0}$. For a variable x we think of all the values $a \in \mathcal{F}$ receiving non-zero coefficients in $\pi_x(S(\psi))$ (i.e. values for which $\pi_x(S(\psi))[a] \neq 0$)) as being simultaneously 'assigned' to x by ψ. The non-triviality requirement means that each variable must be assigned at least one value.
We can now define the $\mathcal{SSAT}$ problem.

Definition 5 **(g-$\mathcal{SSAT}$)** *An instance of $\mathcal{SSAT}$ with parameter g*

$$I = \langle \Psi = \{\psi_1, .., \psi_n\}, V = \{v_1, .., v_m\}, \{\mathcal{R}_{\psi_1}, .., \mathcal{R}_{\psi_n}\} \rangle$$

consists of a set Ψ of tests *over a common set V of variables that take values in a field $\mathcal{F}$. The parameters m and $|\mathcal{F}|$ and $|\mathcal{R}_\psi|$ are always bounded by some polynomial in n. Each test $\psi \in \Psi$ has associated with it a list $\mathcal{R}_\psi$ of assignments to its variables, called the* satisfying assignments *or the* range *of the test ψ. The problem is to distinguish between the following two cases,*

Yes: There is a consistent natural super-assignment for Ψ.

No: Every non-trivial consistent super-assignment for Ψ has norm $> g$.

Theorem 1 **($\mathcal{SSAT}$ Theorem)** *There is some constant $c > 0$, such that $\mathcal{SSAT}$ is NP-hard for $g = n^{c/\log\log n}$.*

The $\mathcal{SSAT}$ theorem (Theorem 1) can be viewed as an extension of Cook's theorem [Coo71, Lev73] in the following way. An algorithm solving $\mathcal{SSAT}$ is required to accept if the test system is satisfiable. However, the algorithm is allowed to accept non-satisfiable instances that have a consistent super-assignment of norm $\leq g$. It must only reject when every consistent super-assignment for Ψ has norm $> g$. We are, in fact, adding slackness between the acceptance and rejection cases.

We suggest a stronger conjecture which, if true, would imply that CVP is NP-hard to approximate to within a *constant power* of the lattice-dimension.

Conjecture 2 *$\mathcal{SSAT}$ is NP-hard for $g = n^c$ for some constant $c > 0$.*

It is unlikely that the conjecture remain true for $c \geq \frac{1}{2}$ due to the result of [GG98] showing that approximating CVP to within $\sqrt{n}$ is in NP$\cap$ co-AM. Our reduction from $\mathcal{SSAT}$ to CVP is linear, and hence it follows that approximating $\mathcal{SSAT}$ to within $\sqrt{n}$ is in NP$\cap$ co-AM as well.

3 Tools and Definitions

3.1 Preliminaries

Let $\mathcal{F}$ denote a finite field $\mathcal{F} = \mathbb{Z}_p$ for some prime number $p > 1$.

Definition 6 (Low Degree Function) *A function $f : \mathcal{F}^d \to \mathcal{F}$ is said to have degree r if its values are the point evaluation of a polynomial on $\mathcal{F}^d$ with degree $\leq r$ in each variable. In this case we say that f is an $[r, d]$-LDF, or $f \in LDF_{r,d}$.*

Sometimes we omit the parameters and refer simply to an LDF. The *total degree* of a function is the total degree of the corresponding polynomial, i.e. the maximum over its monomials, of the sum of degrees of each variable in the monomial. Every $[r, d]$-LDF has total degree at most rd.

For an LDF $P : \mathcal{F}^d \to \mathcal{F}$, we define its restriction and re-parameterization $P|_{\mathcal{C}} : \mathcal{F}^D \to \mathcal{F}$ to the D dimensional cube (affine subspace) $\mathcal{C} = \bar{x}_0 + \mathrm{span}(\bar{x}_1, .., \bar{x}_D)$ (where $\bar{x}_0, .., \bar{x}_D \in \mathcal{F}^d$), in the natural way. Namely,

$$\forall (t_1, .., t_D) \in \mathcal{F}^D, \qquad P_{\mathcal{C}}(t_1, .., t_D) = P(\bar{x}_0 + \sum_{i=1}^{D} t_i \bar{x}_i)$$

Observe that the total degree of $P|_{\mathcal{C}}$ is at most that of P, namely $\leq rd$.

Definition 7 (Low Degree Extension) *Let $m, d > 0$ be natural numbers, and let $\mathcal{H} \subset \mathcal{F}$ such that $|\mathcal{H}|^d = m$. A vector $(a_0, .., a_{m-1}) \in \mathcal{F}^m$ can be naturally identified with a function $A\colon \mathcal{H}^d \to \mathcal{F}$ by looking at points in $\mathcal{H}^d$ as representing numbers in base $|\mathcal{H}|$.*

Let $\hat{A} : \mathcal{F}^d \to \mathcal{F}$ be defined by

$$\hat{A}(x_1, .., x_d) = \sum_{(h_1, .., h_d) \in \mathcal{H}^d} \prod_{\substack{i \in \mathcal{H} \\ i \neq h_1}} \frac{(x_1 - i)}{(h_1 - i)} \cdot \prod_{\substack{i \in \mathcal{H} \\ i \neq h_2}} \frac{(x_2 - i)}{(h_2 - i)} \cdots \prod_{\substack{i \in \mathcal{H} \\ i \neq h_d}} \frac{(x_d - i)}{(h_d - i)} \cdot A(h_1, .., h_d)$$

$\hat{A}$ is a $(|\mathcal{H}| - 1, d)$-LDF called the $|\mathcal{H}| - 1$ degree extension of A in $\mathcal{F}$.

Let $V = \{v_0, .., v_{m-1}\}$ be a set of variables, and identify every assignment $A : V \to \mathcal{F}$ with the vector $(a_0, .., a_{m-1}) \in \mathcal{F}^m$ where $a_i = A(v_i)$. One can extend A to a larger set of variables $\hat{V} \supset V$ via the low-degree-extension of $(a_0, .., a_{m-1})$. Namely, we identify the variables V with the points in $\mathcal{H}^d$, and add new variables for the rest of the points in $\mathcal{F}^d$. The new set of variables $\hat{V}$ correspond each to a point in $\mathcal{F}^d$. $\hat{A}$ is thus viewed as an assignment to $\hat{V} \supset V$ that (1) extends A, and (2) is a point-evaluation of an $[|\mathcal{H}| - 1, d]$-LDF.

Similar to the definition of super-assignments, we define a *super-LDF* $\mathcal{G} : \mathrm{LDF} \to \mathbb{Z}$ to be a formal integer linear combination of LDFs, and denote by $\mathcal{G}[P]$, the integer coefficient assigned to the LDF P. We say that the LDF P *appears* in $\mathcal{G}$ iff $\mathcal{G}[P] \neq 0$. This definition arises naturally from the fact that the tests in our final construction will range over LDFs. We further define the *norm* of a super-LDF to be the norm of the corresponding coefficient vector (same as with super-assignments). We say that a super-LDF has total degree r if every LDF appearing in it has total degree $\leq r$.

Given a super-$[r, d]$-LDF $\mathcal{G}$, we define its restriction $\pi_{\mathcal{C}}(\mathcal{G})$ to a D-dimensional cube $\mathcal{C}$, (which is a super-LDF of dimension D and degree rd) in the natural way. Namely,

$$\forall P \in \mathrm{LDF}_{rd,D} \qquad \pi_{\mathcal{C}}(\mathcal{G})[P] \stackrel{def}{=} \sum_{Q \in \mathrm{LDF}_{r,d},\ Q|_{\mathcal{C}} = P} \mathcal{G}[Q]$$

We say that a point x is ambiguous for a super-LDF $\mathcal{G}$ if there are two LDFs appearing in $\mathcal{G}$, that agree on x. The following (simple) property of super-LDFs will be very important.

Proposition 1 (Low Ambiguity) *Let $\mathcal{G}$ be an $[r, d]$-super-LDF of norm $\leq g$. The fraction of ambiguous points for $\mathcal{G}$ is $\leq \mathrm{amb}(r, d, g) \stackrel{def}{=} \binom{g}{2} \frac{rd}{|\mathcal{F}|}$.*

Proof: Two distinct $[r,d]$-LDFs agree on at most $\frac{rd}{|\mathcal{F}|}$ of their points. At most g LDFs appear in any super-LDF of norm $\leq g$, and so there are no more than $\binom{g}{2}$ pairs. ■

Two LDFs can coincide on only a small fraction of cubes,

Proposition 2 *Let P, Q be two $[r,d]$-LDFs. The fraction of cubes $\mathcal{C}$ (affine subspaces of dimension $D < d$) on which $P|_{\mathcal{C}} = Q|_{\mathcal{C}}$ is $\leq \frac{rd}{|\mathcal{F}|}$.*

This follows from the fact that two distinct $[r,d]$-LDFs agree on at most $\frac{rd}{|\mathcal{F}|}$ of their domain, and by the fact that selecting a random point in a random cube gives a uniform distribution on the entire domain, which implies that the restriction of an LDF to a random cube, is even less likely to avoid all points for which $P(x) \neq Q(x)$.

3.2 Embedding Extension

An important technique utilized herein is adapted from [DFK+99], and shows how to represent an LDF over a low-dimensional domain $\mathcal{C} = \mathcal{F}^t$ by a lower-degree LDF over a domain of higher dimension $\mathcal{D} = \mathcal{F}^{kt}$. The points in the domain $\mathcal{C}$ are embedded in the domain $\mathcal{D}$ by taking each 'axis' in $\mathcal{F}^t$ and replacing it by k new ones (thus the extended domain $\mathcal{F}^{kt}$ has dimension $k \cdot t$) so that an LDF of degree r (in each variable) on the original domain $\mathcal{F}^t$ is transformed to an LDF of degree $\sqrt[k]{r}$ (in each variable) on the extended domain $\mathcal{F}^{kt}$.

Definition 8 (embedding extension) *Let $b \geq 2$, $k > 1$ and t be natural numbers. We define the* embedding extension mapping *$E_b : \mathcal{F}^t \to \mathcal{F}^{t \cdot k}$ as follows. E_b maps any point $x = (\xi_1, .., \xi_t) \in \mathcal{F}^t$ to $y \in \mathcal{F}^{t \cdot k}$, $y = E_b(x) = (\eta_1, .., \eta_{t \cdot k})$ by*

$$E_b(\xi_1, .., \xi_t) \stackrel{def}{=} \left(\xi_1, (\xi_1)^b, (\xi_1)^{b^2}, .., (\xi_1)^{b^{k-1}}, \ldots, \xi_t, (\xi_t)^b, (\xi_t)^{b^2}, .., (\xi_t)^{b^{k-1}}\right)$$

Hence $E_b(\mathcal{F}^t) \subset \mathcal{F}^{kt}$ is a manifold (multi-dimensional curve) in $\mathcal{F}^{kt}$. Each of $\mathcal{F}^{kt}$'s axes corresponds to some preset power of an axis of $\mathcal{F}^t$, and $E_b(\mathcal{F}^t)$ consists of exactly the points in which those axes indeed match.

The following proposition shows that any LDF on $\mathcal{F}^t$ can be represented by an LDF on $\mathcal{F}^{t \cdot k}$ with significantly lower degree:

Proposition 3 *Let $f : \mathcal{F}^t \to \mathcal{F}$ be a $[b^k - 1, t]$-LDF, for integers $t > 0, b > 1, k > 1$. There is a $[b-1, t \cdot k]$-LDF $f_{\text{ext}} : \mathcal{F}^{t \cdot k} \to \mathcal{F}$ such that*

$$\forall x \in \mathcal{F}^t : \quad f(x) = f_{\text{ext}}(E_b(x))$$

Proof: We rewrite f as an LDF $f_{\text{ext}} : \mathcal{F}^{t \cdot k} \to \mathcal{F}$ by replacing each power $(\xi_i)^p$ of ξ_i,

$$0 < i \leq t \quad 0 < p < b^k \qquad (\xi_i)^p \longrightarrow (\eta_{i,0})^{\beta_0} \cdot (\eta_{i,1})^{\beta_1} \cdots (\eta_{i,k-1})^{\beta_{k-1}}$$

where $\langle \beta_0 \beta_1 ... \beta_{k-1} \rangle$ is the base b representation of p, and we 're-index' $\eta_{i,j} \stackrel{def}{=} \eta_{(i-1)k+j+1}$. The degree in each variable of f_{ext} is $b-1$, and the dimension is $t \log_b b^k = t \cdot k$. The restriction of f_{ext}

to the manifold $E_b(\mathcal{F}^t)$, will give f, as seen from substituting the manifold equations $\eta_{i,j} = (\eta_{i,0})^{b^j}$ into each of the monomials). ■

Note that an arbitrary $[r, tk]$-LDF f on the larger domain $\mathcal{F}^{t \cdot k}$ can be viewed, when restricted to the manifold, as a $[\tilde{r}, t]$-LDF $\tilde{f}$ with $\tilde{r} = r \cdot (1 + b + b^2 + \ldots + b^{k-1}) \leq r \cdot (b^k - 1)$. This LDF is the re-parameterization of the LDF obtained by substituting in the manifold equations. Note that if the total degree of f is s, then the total degree of $\tilde{f}$ is $\leq s \cdot b^{k-1}$.

4 Reducing PCP to $\mathcal{SSAT}$

In this section, we present a reduction from a low error-probability PCP characterization of NP, to $\mathcal{SSAT}$. Starting with a PCP instance, we show how to construct an instance of $\mathcal{SSAT}$. The correctness of the reduction is proven in the next section.

Let $\Phi = \{\varphi_1, .., \varphi_n\}$ be a system of *tests* over Boolean variables $V_\Phi = \{v_1, .., v_m\}$, (assume $m = n^c$ for some constant $c > 0$) such that each test depends on $D = O(1)$ variables. The following theorem is a direct corollary of [AS98, ALM+98]:

Theorem 3 *It is NP-hard to distinguish between the following two cases:*

Yes: There is an assignment to V_Φ such that all $\varphi_1, ..., \varphi_n$ are satisfied.

No: No assignment can satisfy more than $1/2$ of the tests in Φ.

Starting from Φ, we will construct an $\mathcal{SSAT}$ test-system Ψ over variables $V_\Psi \supset V_\Phi$. Our new variables V_Ψ will range over a larger, non-Boolean, range, namely a field $\mathcal{F}$. An assignment to V_Ψ can be interpreted as an assignment to V_Φ by identifying the value $0 \in \mathcal{F}$ with the Boolean value `true` and every non-zero value $a \in \mathcal{F}$ with the Boolean value `false`.

4.1 Constructing the CR-Forest

We construct Ψ from Φ by replacing each $\varphi \in \Phi$ with a set of new tests ψ. These tests essentially test that φ is satisfied, and that some set of variables (that encode φ's variables) are an LDF. The construction relies on strong 'error-correcting' properties of LDFs (in a similar manner to proofs of PCP theorems) to eventually 'decode' any consistent low-norm super-assignment for Ψ into a satisfying assignment for the original test-system Φ. The idea is to embed Φ's variables into a geometric domain and then recursively encode this domain by multiple new domains, adding new variables along the way.

We describe the construction via an underlying tree structure, one tree per test $\varphi \in \Phi$. Each node in the tree is associated with a set of variables such that the variables of all of the offspring of a node encode that node's variables. For each leaf of the tree, Ψ will have one test that depends on the variables associated with that leaf.

The key to the construction lies in understanding how the variables associated with different nodes relate to each other This is described in Subsection 4.2. The variables of the root node contain φ's variables, plus some additional ones that together represent the points of a domain $\mathcal{F}^{d_0}$. In fact, every node in the tree will associated with a domain $\mathcal{F}^d$, and each offspring of that node will be associated with a cube $\mathcal{C} \subset \mathcal{F}^d$ in that domain. This is roughly how the points of the parent domain are distributed among its offspring. The variables of each offspring will consist of some of the parent's variables but also some new "extension" variables, together corresponding to points in a new domain $\mathcal{F}^d$ where the parent "cube" variables are mapped via the embedding extension mapping into the new domain.

The idea is that a consistent super-assignment to the tests of Ψ, essentially assigning a super-LDF to each leaf node, can be inductively decoded into super-LDFs on domains of nodes residing higher up in the tree, reaching all the way up to the root. For this decoding to work, certain points in a domain, containing more 'information' than others, need to have a larger proportion of offspring representing them. This is established (in Subsection 4.3) by defining for each domain a set of 'distinguished points'. Then, a mechanism of labels serves to obtain the correct proportion of offspring encoding the distinguished and the non-distinguished points.

Let us begin by defining the composition-recursion forest (CR-forest), which holds the underlying structure of Ψ.

Let $\mathcal{F}$ be a field of size $|\mathcal{F}| = |V_\Phi|^{\Theta(1)/\log\log n} = n^{c_1/\log\log n}$ for some constant $c_1 > c$ (recall we denoted $|V_\Phi| = n^c$). Let $d_0 = \lceil \log\log n \rceil$, recall that D denotes the number of variables each test in Φ depends on, and set $d = 4D + 8$. Let $L = \lceil c_2 \log\log n \rceil$, (the constant $c_2 > 0$ will be specified later).

Let $B(\mathcal{F}^d, t_1, t_2)$ denote the number of different affine-subspaces of dimension t_1 (in a domain $\mathcal{F}^d$) that contain a certain affine subspace of dimension $0 \le t_2 \le t_1$. It is easy to see that $B(\mathcal{F}^d, t_1, t_2) \le |\mathcal{F}|^{d(t_1 - t_2)}$.

Definition 9 ($\mathbf{F}_n(\Phi)$) *The composition-recursion forest (CR-forest) $\mathbf{F}_n(\Phi) = \{\mathbf{T}_\varphi \mid \varphi \in \Phi\}$ is a set containing one depth-L tree $\mathbf{T}_\varphi$ for every test $\varphi \in \Phi$. The root node (level-0) of $\mathbf{T}_\varphi$ has $B(\mathcal{F}^{d_0}, D+2, D-1) = n^{O(1)}$ offspring, and all nodes in levels $i = 1, \ldots, L-1$ have $2\,|\mathcal{F}|^{D+2} \cdot B(\mathcal{F}^{4D+8}, D+2, 0) = |\mathcal{F}|^{O(1)}$ offspring. Note that although the forest $\mathbf{F}_n(\Phi)$ depends on many parameters (L, D, d_0) which can all be derived from Φ, we single out the parameter n according to which the size of the generated instance is measured.*

The forest $\mathbf{F}_n(\Phi)$ will be the base upon which Ψ's variables and tests will be defined as follows. With each node $v \in \mathbf{T}_\varphi$ $(\varphi \in \Phi)$, we associate a distinct geometric domain, denoted $\mathbf{dom}_v$. For the root $root_\varphi$ of every tree, $\mathbf{dom}_{root_\varphi} \overset{def}{=} \mathcal{F}^{d_0}$, while for non-root nodes v, $\mathbf{dom}_v \overset{def}{=} \mathcal{F}^d$. For a node v, we associate with each point in $\mathbf{dom}_v$ a distinct variable from V_Ψ, by defining an injection $\mathbf{var}_v : \mathbf{dom}_v \to V_\Psi$. Points from domains of distinct nodes may be mapped to the same variable. In particular, the variables that φ depends on will belong to $\mathbf{var}_v(\mathbf{dom}_v)$ for many of the leaves in the tree $\mathbf{T}_\varphi$.

We can already at this point define the tests of Ψ,

Definition 10 (tests) *Ψ will have one test ψ_v for each leaf v in the forest. ψ_v will depend on the variables in $\mathbf{var}_v(\mathbf{dom}_v)$. An assignment A for ψ_v's variables is considered satisfying if and only if the following two conditions hold:*

1. *A is an $[r_L, d]$-LDF on $\mathbf{var}_v(\mathbf{dom}_v)$ (where $r_L \le 2(D+2) = O(1)$ will be defined below).*
2. *If $v \in \mathbf{T}_\varphi$ for $\varphi \in \Phi$ and all of φ's variables appear in $\mathbf{var}_v(\mathbf{dom}_v)$, then A must satisfy φ.*

The instance of $\mathcal{SSAT}$ that we construct, must have a list of satisfying assignments for each test. Note that the size of this list is bounded by the number of $[r_L, d]$-LDFs which is $|\mathcal{F}|^{O(1)}$, i.e. polynomial in n. Having defined the tests in Ψ and the satisfying assignments for each test, it now only remains to specify the variables that each test accesses, i.e. define for each node v, the mapping $\mathbf{var}_v : \mathbf{dom}_v \to V_\Psi$.

4.2 Variables

We begin by defining the variable mappings for the root nodes of the trees in the forest. Recall that for the root node $root_\varphi$ of each tree $\mathbf{T}_\varphi$, we set $\mathbf{dom}_{root_\varphi} = \mathcal{F}^{d_0}$. Let $\hat{V}_\Phi \supset V_\Phi$ be the variables representing the low-degree-extension (Definition 7, with parameters $m = |V_\Phi|$, $d_0 = \lceil \log\log n \rceil$, and $\mathcal{H} \subset \mathcal{F}$ such that[1] $|\mathcal{H}|^{d_0} = |V_\Phi|$) of V_Φ, i.e. $\hat{V}_\Phi$ is a set of $|\mathcal{F}|^{d_0}$ variables each representing a distinct point in $\mathcal{F}^{d_0}$. We define the mapping $\mathbf{var}_{root_\varphi}$ as follows,

Definition 11 ($\mathbf{var}_{root_\varphi}$) *The bijection* $\mathbf{var}_{root_\varphi} : \mathbf{dom}_{root_\varphi} \to \hat{V}_\Phi$ *maps the points of* $\mathbf{dom}_{root_\varphi} = \mathcal{F}^{d_0}$ *to* $\hat{V}_\Phi$ *in the following manner. Take* $\mathcal{H} \stackrel{def}{=} \{0, .., h-1\} \subset \mathcal{F}$ *such that* $|\mathcal{H}|^{d_0} = h^{d_0} = |V_\Phi|$ *(i.e.* $|\mathcal{H}| = |V_\Phi|^{\frac{1}{d_0}} = n^{c/d_0} = n^{c/\log\log n}$ *and since* $|\mathcal{F}| = n^{c_1/\log\log n}$ *we have* $|\mathcal{H}|^{c_1/c} = |\mathcal{F}|$*). We define* $\mathbf{var}_{root_\varphi}$ *to be a bijection* independent of φ, *taking the points of* $\mathcal{H}^{d_0} \subset \mathcal{F}^{d_0}$ *to* V_Φ, *and the remaining points* $\mathcal{F}^{d_0} \setminus \mathcal{H}^{d_0}$ *to* $\hat{V}_\Phi \setminus V_\Phi$.

Note that for every $\varphi \in \Phi$ the points of $\mathbf{dom}_{root_\varphi}$ were mapped to *the same* variables, hence each of the $|\mathcal{F}|^{d_0}$ variables in $\hat{V}_\Phi$ has $|\Phi|$ pre-images (so far).

For simplicity we assume that for each $\varphi \in \Phi$, the points mapped to φ's variables are in general position (i.e. they span a $(D-1)$-dimensional affine-subspace of $\mathcal{F}^{d_0}$), otherwise, we choose an arbitrary $(D-1)$-dimensional affine subspace containing these points.

Before we continue to define the mappings $\mathbf{var}_v$ for non-root nodes, let us examine the purpose of these mappings. Picture a super-assignment to the tests of Ψ, as a labeling of each leaf in the forest by a super-LDF. We will prove (see Lemma 6) that such an assignment, if consistent and of low-norm, 'induces' a low-norm super-LDF for the domain of each *internal* node, and in particular – a low-norm super-LDF $\mathcal{G}$ for the 'root-domain', $\mathcal{F}^{d_0}$. We now use the fact that the variables representing this root-domain are common to the roots of all $\mathbf{T}_\varphi$'s, to interpret $\mathcal{G}$ as a global assignment for the variables in V_Φ. Namely, we will show that any LDF that appears in $\mathcal{G}$ with a non-zero coefficient assigns V_Φ values that satisfy most of the tests in Φ.

The idea behind the CR-forest is that the domain $\mathbf{dom}_u$ of a node u is 'represented' by its offspring' domains. u's domain's points are distributed among the domains of each of u's offspring. The aforementioned Lemma 6 will show how to join the LDFs of u's offspring into one LDF for u. The advantage we gain by representing one LDF over u's domain by many LDFs over u's offspring' domains is that we can enforce the degree of the LDFs in the leaves to be very low, compared to the degree of the LDF on the root that they represent (the dimension of the LDFs is maintained low as well). Therefore the list of satisfying assignments for the tests in Ψ (corresponding to LDFs on the leaves' domains) is not too long. We can afford to list all LDFs (i.e. satisfying assignments) only when the degree (and dimension) of the LDFs is small enough, because for a higher degree the length of the list would not be polynomial in n.

[1]If $\sqrt[d]{m}$ is not an integer, we add dummy variables to V_Φ.

The key to understanding the construction is to see how a node u is 'represented' by its offspring. Pictorially, u's domain's points are distributed among the domains of u's offspring, each offspring v receives a slice of u's domain. Some of v's points correspond to v's slice of u's variables. The rest of v's points are some (low-degree) encoding or extension of these points.

Consider a non-root node v, and denote its parent by u. Assuming $\mathbf{var}_u$ is already defined, we now specify the mapping $\mathbf{var}_v : \mathbf{dom}_v \to V_\Psi$. Some (exactly $|\mathcal{F}|^{D+2}$) of the points in $\mathbf{dom}_v$ 'represent' points from $\mathbf{dom}_u$, and will thus be mapped to u's variables ($\mathbf{var}_u(\mathbf{dom}_u)$). The rest of the points in $\mathbf{dom}_v$ will be mapped to fresh new variables $V_v \subset V_\Psi$ ($|V_v| = |\mathcal{F}|^d - |\mathcal{F}|^{D+2}$) associated with the node v. Only points in domains of nodes in v's sub-tree may be mapped to V_v. For uniformity of notation, we define $V_{root_\varphi} \stackrel{def}{=} \hat{V}_\Phi$, for every root $root_\varphi$, again stressing the fact that the roots of all of the trees share the same variables. Altogether

$$V \stackrel{def}{=} V_\Psi = \bigcup_{\substack{v \in \mathbf{T}_\varphi \\ \varphi \in \Phi}} V_v \ .$$

u's variables are distributed among its offspring by letting each offspring v of u 'represent' an affine sub-space $\mathcal{C}_v \subset \mathbf{dom}_u$ of dimension $D+2$ (a $(D+2)$-cube). More formally, we label (as specified later in Subsection 4.3) each offspring v of u by a $(D+2)$-cube $\mathcal{C}_v \subset \mathbf{dom}_u$. We represent a cube $\mathcal{C}_v$ by $D+3$ points $x_0, .., x_{D+2}$ such that $\mathcal{C}_v = x_0 + \text{span}(x_1, .., x_{D+2})$ (this yields a natural way of viewing $\mathcal{C}_v$ as $\mathcal{F}^{D+2}$).

We embed all points of the cube $\mathcal{C}_v \subset \mathbf{dom}_u$ into the domain $\mathbf{dom}_v$ by the embedding extension mapping, defined above in Subsection 3.2, $E_{b_i} : \mathcal{C}_v \to \mathbf{dom}_v$ (the parameter b_i depends on the level $i \geq 1$ of the node v, and is specified shortly below). Via this mapping, we can transform LDFs on the cube $\mathcal{C}_v$ to *lower-degree* LDFs on the domain $\mathbf{dom}_v$. This will allow us to represent a satisfying assignment to Φ by $[r_i, d]$-LDFs on the domains $\mathbf{dom}_v$ of level-i nodes (the degree r_i will be defined below). The construction is aimed to *lower* the degree r_i of the LDFs, from $r_0 \stackrel{def}{=} |\mathcal{H}| \approx n^{1/\log\log n}$ to $r_L = O(1)$.

We think of the point $y = E_{b_i}(x) \in \mathbf{dom}_v$ as 'representing' the point $x \in \mathcal{C}_v \subset \mathbf{dom}_u$, and define $\mathbf{var}_v : \mathbf{dom}_v \to V_\Psi$ as follows,

Definition 12 ($\mathbf{var}_v$, for a non-root node v) *Let v be a non-root node, let u be v's parent, and let $\mathcal{C}_v \subset \mathbf{dom}_u$ be the label attached to v (the* label *of a node is defined below, Definitions 13,14). For each point $y \in E_{b_i}(\mathcal{C}_v) \subset \mathbf{dom}_v$ define $\mathbf{var}_v(y) \stackrel{def}{=} \mathbf{var}_u(E_{b_i}^{-1}(y))$, i.e. points that 'originated' from $\mathcal{C}_v$ are mapped to the previous-level variables, that their pre-images under E_{b_i} in $\mathcal{C}_v$ were mapped to. For each 'new' point $y \in \mathbf{dom}_v \setminus E_{b_i}(\mathcal{C}_v)$ we define $\mathbf{var}_v(y)$ to be a distinct variable from V_v.*

The parameters used for the embedding extension mappings E_{b_i} are $t = D+2$, $k = d/t$. We set $r_0 = |\mathcal{H}| = |\mathcal{F}|^{c/c_1}$ and r_{i+1} and b_{i+1} ($i \geq 0$) are defined by the following recursive formulas:

$$\begin{aligned} b_{i+1} &= \lceil \sqrt[4]{r_i(D+2)+1} \rceil \\ r_{i+1} &= b_{i+1} - 1 \end{aligned}$$

(we will show in Subsection 5.1 that b_i, r_i decrease until for some $L < \log\log n$, $r_L \leq 2(D+2) = O(1)$).

In order to complete the description of the test-system, we now only need to describe the cube-labeling of all of the offspring of each node. This will describe how the representation of a node u is distributed among its offspring.

4.3 Labeling Nodes

We define the *offspring-labels* of a node u, thereby completing the description of the construction. As described above, each offspring of the node u 'represents' an affine subspace in the domain $\mathbf{dom}_u$, i.e. the variables of u's offspring represent an encoding of u's variables. This representation has some error. To control this error, we proportion the offspring so that more important variables are represented by more offspring. Roughly speaking, the 'importance' of a variable $x \in V$ is determined by how high up (towards the root) in the tree this variable appears. The closer the variable is to the root, the more information it represents about Φ's original variables, V_Φ.

We will use a mechanism of 'distinguished-points' to promote the importance of certain points more than others. This mechanism works by having a higher proportion of descendants of v 'represent' the distinguished points of v.

Let us begin by defining the labels of the offspring of a root node $root_\varphi$. The tests at the leaves of the tree $\mathbf{T}_\varphi$ represent the test $\varphi \in \Phi$. Therefore, the variables that φ depends on are 'very important' to represent and their corresponding points are the distinguished points of each root node,

Definition 13 (offspring-labels for a root node) *Let*

$$\mathbf{dst}(root_\varphi) \stackrel{def}{=} \left\{ x \in \mathbf{dom}_{root_\varphi} \,\middle|\, \varphi \text{ depends on } \mathbf{var}_{root_\varphi}(x) \right\}$$

be a set of distinguished points for $root_\varphi$ (recall our assumption that $\mathbf{dst}(root_\varphi)$ is a set of exactly D points in general position). We label each offspring of $root_\varphi$ by a distinct cube from the following set:

$$\mathbf{labels}(root_\varphi) \stackrel{def}{=} \left\{ \mathcal{C} \text{ is a } (D+2)\text{-cube in } \mathcal{F}^{d_0} \mid \mathcal{C} \supset \mathbf{dst}(root_\varphi) \right\}$$

The number of labels $|\mathbf{labels}(root_\varphi)| = B(\mathcal{F}^{d_0}, D+2, D-1)$ is the number of $(D+2)$-cubes containing the $(D-1)$-cube spanned by the points mapped to φ's variables (assuming, as mentioned above, that these points are in general position).

For a general non-root node $v \in \mathbf{T}_\varphi$, we consider two levels of 'important' variables: (1) variables that belong to some ancestor (direct and indirect) of v (there are $|\mathcal{F}|^{D+2}$ such variables, all mapped from v's parent) and (2) variables mapped from the distinguished points of v (there will always be exactly one or D such variables). The node v will correspondingly have two equal-weight sets of offspring,

Definition 14 (offspring-labels for non-root nodes) *Let v be a non-root node. We define two multi-sets of offspring-labels for v. For each variable $\mathrm{x} \in \mathbf{var}_v(\mathbf{dom}_v) \setminus V_v$, i.e. x that belongs to some ancestor of v, we define*

$$\mathbf{labels}_{\mathrm{x}}(v) \stackrel{def}{=} \{ \mathcal{C} \subset \mathbf{dom}_v \text{ is a } (D+2)\text{-cube} \mid \mathrm{x} \in \mathbf{var}_v(\mathcal{C}) \}$$

we then take $\mathbf{labels}_1(v)$ *to be the multi-set*

$$\mathbf{labels}_1(v) \stackrel{def}{=} \bigcup_{\mathrm{x} \in \mathbf{var}_v(\mathbf{dom}_v) \setminus V_v} \mathbf{labels}_{\mathrm{x}}(v)$$

For every offspring w of v, labelled by a cube from $\mathbf{labels}_{\mathrm{x}}(v)$*, we define* $\mathbf{dst}(w)$ *to be the singleton set consisting of the point in* $\mathbf{dom}_w$ *that is mapped to* x*, i.e.* $\mathbf{dst}(w) \stackrel{def}{=} \{\mathbf{var}_w{}^{-1}(\mathrm{x})\} \subset \mathbf{dom}_w$*.*

The second multi-set (actually set) of offspring-labels is devoted to representing the distinguished points of v. We simply take

$$\mathbf{labels}_2(v) \stackrel{def}{=} \{\mathcal{C} \subset \mathbf{dom}_v \textit{ is a } (D+2)\textit{-cube} \mid \mathcal{C} \supset \mathbf{dst}(v)\}$$

For each offspring w of v labelled by a cube from $\mathbf{labels}_2(v)$*, we set* $\mathbf{dst}(w) \stackrel{def}{=} E_{b_i}(\mathbf{dst}(v))$ *(where i is w's level in the tree), i.e. w distinguishes the same set of points as v after embedding via* E_{b_i}*.*

The final multi-set $\mathbf{labels}(v)$ *is the union of* $\mathbf{labels}_1(v)$ *and* $\lfloor M \rfloor$ *copies of* $\mathbf{labels}_2(v)$*, where the number* $M = |\mathcal{F}|^{D+2} B(\mathcal{F}^d, D+2, 0)/B(\mathcal{F}^d, D+2, |\mathbf{dst}(v)|-1)$ *is chosen so that at least half of the labels are from* $\mathbf{labels}_1$*, and at least a third of the labels are from* $\mathbf{labels}_2$*.*

4.4 Construction Size

Recall that we defined $d_0 \stackrel{def}{=} \lceil \log\log n \rceil$ and $d \stackrel{def}{=} 4D+8$. We also set $r_0 = |\mathcal{H}| = |\mathcal{F}|^{c/c_1} = n^{c/\log\log n}$, and defined $b_{i+1} = \lceil \sqrt[4]{(D+2)r_i + 1} \rceil$ and $r_{i+1} = b_{i+1} - 1$ for every $i \geq 0$.

We claim that indeed $r_L = O(1)$ for some $L \leq \log\log n$. For this purpose we prove by simple induction that $r_i \leq \max(\lceil r_0^{1/2^i} \rceil, 2(D+2))$. For r_0 this indeed holds, and assuming it true for r_i we have that if $r_i > 2(D+2)$ and $b_{i+1} > 2$, then

$$r_{i+1} < b_{i+1} = \lceil \sqrt[4]{(D+2)r_i + 1} \rceil < \lceil \sqrt[4]{2r_i(D+2)} \rceil \leq \lceil (r_i)^{2/4} \rceil \leq \lceil \sqrt{r_i} \rceil \leq \lceil (r_0)^{1/2^{i+1}} \rceil.$$

We set L to be the first index for which $r_L \leq 2(D+2) = O(1)$. Obviously, until that point r_i, b_i decrease monotonically, and since $r_0 = 2^{c\log n/\log\log n}$, $L \leq \lfloor \log(c\log n/\log\log n) \rfloor + 1 < \log\log n$ (assuming n is large enough). Clearly $b_i > 2$ for all $0 \leq i \leq L$, and the induction is complete.

The size of the $\mathcal{SSAT}$ instance (Recall Definition 5) also depends on the range of the tests, which is the parameter we are about to bound.

The Range of the Tests. The tests of the test-system range over $[r_L, d]$-LDFs. The number of monomials of degree $r_L \leq 2(D+2) = O(1)$, and dimension $d = 4D+8 = O(1)$ is bounded by $(r_L+1)^d = O(1)$. The number of $[r_L, d]$-LDFs is hence bounded by $|\mathcal{F}|^{O(1)} < O(n)$ and therefore the range of the tests is polynomial in n.

The Number of Tests and Variables. It is only left to verify that the size of the forest is polynomial. We have $|\Phi| = n$ trees, so let's verify that the number of nodes in each tree is polynomially-bounded.

Consider a tree $\mathbf{T} = \mathbf{T}_\varphi \in \mathbf{F}_n(\Phi)$. $root_\varphi$ has $B(\mathcal{F}^{d_0}, D+2, D-1) \le |\mathcal{F}|^{3d_0} = n^{O(1)}$ offspring and each node in level i $(0 < i < L)$ has $2\,|\mathcal{F}|^{D+2} \cdot B(\mathcal{F}^d, D+2, 0) = |\mathcal{F}|^{O(1)}$ offspring. Altogether the number of nodes in $\mathbf{T}$ is bounded by

$$n^{O(1)} \cdot \prod_{i=1}^{L} |\mathcal{F}|^{O(1)} = n^{O(1)} \cdot |\mathcal{F}|^{O(L)} = n^{O(1)} \cdot (2^{\log n/\log\log n})^{O(\log\log n)} = n^{O(1)}$$

Hence the number of tests in Ψ is polynomial, and the number of variables is $\le |\mathcal{F}|^d \cdot |\Psi| = n^{O(1)}$.

5 Correctness of the Reduction

In this section we prove the completeness and soundness of the reduction presented in the previous section.

5.1 Completeness

Lemma 4 (Completeness) *If there is an assignment $\mathcal{A} : V_\Phi \to \{\mathtt{true}, \mathtt{false}\}$ satisfying all of the tests in Φ, then there is a natural assignment $\mathcal{A}_\Psi : V_\Psi \to \mathcal{F}$ satisfying all of the tests in Ψ.*

Of course, this assignment $\mathcal{A}_\Psi$ is equivalent to a consistent natural super-assignment. We extend $\mathcal{A}$ following the rationale of the construction, by taking its low-degree-extension to the variables $\hat{V}_\Phi$, and then repeatedly taking the embedding extension of the previous-level variables, until we've assigned all of the variables in the system. More formally,

Proof: We construct an assignment $\mathcal{A}_\Psi : V_\Psi \to \mathcal{F}$. We first set (for every $\varphi \in \Phi$) $P_{root_\varphi} : \mathbf{dom}_{root_\varphi} \to \mathcal{F}$ to be the $[r_0, d_0]$-LDF that is the low degree extension (see Definition 7) of $\mathcal{A}$ (we think of $\mathcal{A}$ as assigning each variable a value in $\{0,1\} \subset \mathcal{F}$ rather than $\{\mathtt{true}, \mathtt{false}\}$, see discussion in the beginning of Section 4). We proceed to inductively obtain $[r_i, d]$-LDFs $P_v : \mathbf{dom}_v \to \mathcal{F}$ for every level-i $(i > 0)$ node v of every tree in the CR-forest, as follows. Assume we've defined an $[r_i, d]$-LDF (an $[r_i, d_0]$-LDF in case $i = 0$) P_u consistently for all level-i nodes, and let v be an offspring of u, labelled by $\mathcal{C}_v$. The restriction $f = P_u|_{\mathcal{C}_v}$ of P_u to the cube $\mathcal{C}_v$ is a $[r_i(D+2), D+2]$-LDF. f can be written as a $[\lceil \sqrt[4]{r_i(D+2)+1}\rceil - 1,\ 4D+8]$-LDF f_{ext} over the larger domain $\mathcal{F}^d$, as promised by Proposition 3 taking $k = 4$. We define $P_v = f_{\text{ext}}$ to be that $[r_{i+1}, d]$-LDF (recall that $d = 4D+8$ and $r_{i+1} = b_{i+1} - 1 = \lceil \sqrt[4]{r_i(D+2)+1}\rceil - 1$).

Finally, for a variable $\mathrm{x} \in \mathbf{var}_v$, $\mathrm{x} = \mathbf{var}_v(x)$, we set $\mathcal{A}_\Psi(\mathrm{x}) \stackrel{def}{=} P_v(x)$. The construction implies that there are no collisions, i.e. $\mathrm{x}' = \mathbf{var}_{v'}(x') = \mathbf{var}_v(x) = \mathrm{x}$ implies $P_v(x) = P_{v'}(x')$. ∎

5.2 Soundness

In this subsection we show that a 'no' instance of PCP is always mapped to a 'no' instance of $\mathcal{SSAT}$. We assume that the constructed $\mathcal{SSAT}$ instance has a consistent super-assignment of norm $\le g$, and show that Φ – the PCP test system we started with – is satisfiable.

Lemma 5 (Soundness) *Let $g \stackrel{def}{=} |\mathcal{F}|^{c_g}$ where $c_g > 0$ is some small enough constant, say $c_g = 1/1000$. If there exists a non-trivial consistent super-assignment of norm $\leq g$ for Ψ, then Φ is satisfiable.*

Let us first sketch a brief outline of the proof. The proof follows the structure of the trees underlying the construction. Since the tree structure is different for the first level nodes and for all other levels, we divide the proof accordingly.

We begin with a few definitions preparing for the proof itself. We then state Lemma 6 that encapsulates the inductive part, handling all internal nodes in levels ≥ 1 of the tree, and proving that a non-trivial consistent super-assignment at the leaves can be decoded into "consistent" super-LDFs on "most" internal nodes. Relying on this lemma, we proceed to prove the soundness lemma (Lemma 5). The heart of the proof is a consistency lemma (Lemma 7) that allows us to combine "consistent" super-LDFs on domains of offspring of a node into one super-LDF on the domain of that node. We use this lemma to combine the super-LDFs on the root's offspring (i.e. level-1 nodes) into one global super-LDF on the common domain $\mathcal{F}^{d_0}$, and from it deduce an assignment satisfying the original PCP test-system Φ.

We then return to the inductive proof of Lemma 6 again relying on the same consistency lemma (Lemma 7) for the inductive step.

The proof of the consistency lemma (Lemma 7) itself follows in Subsection 5.3.

Proof: Let $\mathcal{SA}$ be a non-trivial consistent super-assignment for Ψ, of norm $\|\mathcal{SA}\| \leq g$. It induces (by projection) a super-assignment to the variables

$$m : V_\Psi \longrightarrow \mathbb{Z}^{|\mathcal{F}|}$$

i.e. for every variable $\mathrm{x} \in V_\Psi$, m assigns a vector $\pi_{\mathrm{x}}(\mathcal{SA}(\psi))$ of integer coefficients, one per value in $\mathcal{F}$ where ψ is some test depending on x. Since $\mathcal{SA}$ is consistent, m is well defined (independent of the choice of test ψ). Alternatively, we view m as a labeling of the points $\bigcup_{v \in \mathbf{T}_\varphi \in \mathbf{F}_n(\Phi)} \mathbf{dom}_v$ by a 'super-value' – a formal linear combination of values from $\mathcal{F}$. The label of the point $x \in \mathbf{dom}_v$ for some $v \in \mathbf{T}_\varphi \in \mathbf{F}_n(\Phi)$, is simply $m(\mathbf{var}_v(x))$, and with a slight abuse of notation, is sometimes denoted $m(x)$. m is used as the "underlying point super-assignment" for the rest of the proof, and will serve as an anchor by which we test consistency. Since $\mathcal{SA}$ is non-trivial, $m(x) \neq \vec{0}$ for every x.

For a node u, we denote by $\mathrm{Avg}(u)$ the average of $\|\mathcal{SA}(\psi_v)\|$ over the leaves v in u's sub-tree. We will show that whenever $\mathrm{Avg}(u)$ is not too high for a node u, then u's subtree is, in a sense, consistent. We thus define a 'good' node as one having a low average norm on its subtree's leaves:

Definition 15 (Good Nodes) *Fix $C > 0$ large enough, e.g. $C = 301$. A node u in level i in the CR-forest is said to be* good *if*

$$\mathrm{Avg}(u) \leq g_i \stackrel{def}{=} g \cdot C^{i+1}.$$

We denote by $\mathbf{nodes}_i^*$ the set of good nodes in level i.

For any node v, denote by $\mathbf{ofsp}(v)$ the set of v's direct offspring. It is easy to see that most offspring of a good node are themselves good:

Proposition 4 *If $u \in \mathbf{nodes}_i^*$, then*

$$\Pr_{v \in \mathbf{ofsp}(u)}(v \in \mathbf{nodes}_{i+1}^*) \geq 1 - 1/C$$

Proof: All subtrees rooted at offspring of u have the same number of offspring. u is good, thus by definition $\mathrm{Avg}(u) \leq g_i$. Had u more than $1/C$ bad offspring, then the total average of $\mathcal{SA}(\psi_v)$ on its sub-tree would be $> g_{i+1} \cdot \frac{1}{C} = g_i$. ∎

The central task of our proof is to show that a consistent low-norm super-assignment to the tests at the leaves induces a low-norm super-LDF on the root domain. The key step in this proof is the inductive step showing that if, for a node v, almost all of its offspring have a low-norm super-LDF that is consistent with m, then we can deduce such a super-LDF $\mathcal{G}_v$ over $\mathbf{dom}_v$.

It turns out that for a general node u we cannot always deduce a super-LDF agreeing with m on every point in $\mathbf{dom}_u$ (a counter-example can be constructed). Instead, for good nodes u we show that there exists a super-LDF $\mathcal{G}_u$ over $\mathbf{dom}_u$ that agrees with *almost* all of the super-LDFs on u's offspring. By 'agrees' we mean that if v is an offspring of u that is labelled by $\mathcal{C}_v$, then the parent super-LDF $\mathcal{G}_u$ projected on the points of $\mathcal{C}_v$ equals $\mathcal{G}_v$ projected on the manifold points $E_{b_i}(\mathcal{C}_v)$ (see Definition 16 below).

The consistency with m will then follow inductively from the fact that the offspring' super-LDFs were consistent with m. The importance of consistency with m is not the same for all points. For certain points (e.g. those mapped to variables from V_Φ) we cannot allow any inconsistency, while for others we can allow some small error. For every node v we consider, as mentioned before in the construction, two types of special points: The distinguished points $\mathbf{dst}(v)$, and the manifold points,

Definition 16 (Manifold Points) *For a non-root node v labelled by $\mathcal{C}_v$, we define the manifold points $\mathbf{manf}(v) \stackrel{def}{=} E_{b_i}(\mathcal{C}_v)$, where i is v's level in the tree.*

These are the points that originate from all of v's ancestors.

We rely on the manner in which the offspring were proportioned to deduce a high level of consistency for these special points. We can now state the key inductive lemma (Lemma 6) showing how consistency follows from having a low-norm super-assignment. This lemma relies heavily on the precise structure of the forest, and shows that for every good node u, there is a super-LDF on u's domain that is 'almost' consistent with "the anchor" m. The lemma is proved inductively, constructing u's super-LDF from the super-LDFs of u's good offspring. We will later want to construct from these super-LDFs an assignment that satisfies more than half of the tests in Φ. For this purpose, we need the super-LDFs along the way to be legal,

Definition 17 (Legal) *An LDF P is called* legal for a node $v \in \mathbf{T}_\varphi$ *(for some $\varphi \in \Phi$), if it satisfies φ in the sense that if φ's variables have pre-images under $\mathbf{var}_v$ $x_1, .., x_D \in \mathbf{dom}_v$, then $P(x_1), .., P(x_D)$ satisfy φ. A super-LDF $\mathcal{G}$ is called* legal for $v \in \mathbf{T}_\varphi$ *if for every LDF P appearing in $\mathcal{G}$, P is legal for $v \in \mathbf{T}_\varphi$.*

Lemma 6 *Let $u \in \mathbf{nodes}_i^*$ for some $i \geq 1$, and set $\alpha = 3/C$ and $A = 4 \cdot (D+2)^3 = O(1)$. There exists a legal super-LDF $\mathcal{G}_u$ over $\mathbf{dom}_u$ of degree at most $\tilde{r}_i \stackrel{def}{=} A^{L-i} \cdot (r_i + 1)$ and of norm $\leq 2^{L-i} \cdot \mathrm{Avg}(u)$ that agrees with m on $\mathbf{dst}(u)$ and on $1-\alpha$ fraction of the points in $\mathbf{manf}(u)$, i.e.*

$$\forall x \in \mathbf{dst}(u) \quad \pi_x(\mathcal{G}_u) = m(x)$$

and

$$\Pr_{x \in \mathbf{manf}(u)} (\pi_x(\mathcal{G}_u) = m(x)) \geq 1 - \alpha$$

This lemma is the key to our construction. Its proof shows how consistent LDFs on offspring nodes induce an LDF on the parent. Before we prove this lemma, which is somewhat technical, let us first use it to complete the proof of Lemma 5.

Applying Lemma 6 for level-1 nodes, we deduce that $\mathcal{SA}$ induces a legal super-LDF $\mathcal{G}_v$ of degree $\tilde{r}_1$ and with $\|\mathcal{G}_v\| \leq 2^{L-1}\mathrm{Avg}(v)$, for every node v in $\mathbf{nodes}_1^*$. We now join these super-LDFs into one legal super-LDF over the root domain $\mathcal{F}^{d_0}$, and then deduce a satisfying assignment to the tests in Φ from this super-LDF. Let $v \in \mathbf{nodes}_1^*$ be an offspring of $root_\varphi$ for some $\varphi \in \Phi$, and let $\mathcal{C}_v \subset \mathbf{dom}_{root_\varphi}$ be the cube labeling v. We would like to view $\mathcal{G}_v$ as a super-LDF over $\mathcal{C}_v$, by restricting the LDFs in $\mathcal{G}_v$ to the manifold $\mathbf{manf}(v) \subset \mathbf{dom}_v$ that represents $\mathcal{C}_v$. For every $[\tilde{r}_1, d]$-LDF $P : \mathbf{dom}_v \to \mathcal{F}$, define $\tilde{P}$ as the $[d\tilde{r}_1 \cdot (b_1)^3, D+2]$-LDF which is defined as P's restriction to the manifold $\mathbf{manf}(v)$:

$$\forall x \in \mathcal{C}_v \subset \mathbf{dom}_{root_\varphi} : \quad \tilde{P}(x) \stackrel{def}{=} P(E_{b_1}(x))$$

(Note that since P's total degree is $\leq \tilde{r}_1 \cdot d$, the total degree of $\tilde{P}$ is $\leq \tilde{r}_1 \cdot d \cdot (b_1)^3$ because the degree of E_{b_1} is $(b_1)^{4-1} = (b_1)^3$). For every LDF $P : \mathbf{dom}_v \to \mathcal{F}$, $\mathcal{G}_v$ assigns an integer value $\mathcal{G}_v[P]$. We define the super-LDF $\tilde{\mathcal{G}}_v$ to be the same formal linear combination as $\mathcal{G}_v$, replacing each LDF P with $\tilde{P}$:

$$\tilde{\mathcal{G}}_v[\tilde{Q}] = \sum_{P:\, \tilde{P} = \tilde{Q}} \mathcal{G}_v[P]$$

In other words, the super-LDF $\tilde{\mathcal{G}}_v$ is simply the restriction (and re-parameterization) of $\mathcal{G}_v$ to the manifold $\mathbf{manf}(v)$, as discussed in Subsection 3.2. The total degree of $\tilde{\mathcal{G}}_v$ is $\tilde{r}_1 d \cdot (b_1)^3$.

Let $root_\varphi$ be a good root node. Since the average norm of all the tests is $\leq g$, and a good root node is by definition one with $\mathrm{Avg}(root_\varphi) \leq g \cdot C$, there are at least $1 - 1/C = 1 - \alpha/3$ such nodes. For every good offspring v of $root_\varphi$, Lemma 6 guarantees that $\Pr_{x \in \mathcal{C}_v}\left(\pi_x(\tilde{\mathcal{G}}_v) = m(x)\right) \geq 1 - \alpha$, and that for every $x \in \mathbf{dst}(v)$, $\pi_x(\tilde{\mathcal{G}}_v) = m(x)$. Given this assignment of super-LDF $\tilde{\mathcal{G}}_v$ per label $\mathcal{C}_v \subset \mathbf{dom}_{root_\varphi} = \mathcal{F}^{d_0}$, we would like to use the fact that these super-LDFs are consistent with m to deduce the existence of some global super-LDF on $\mathcal{F}^{d_0}$ (that is also consistent with m). The following consistency lemma, when applied for $u = root_\varphi$, will imply just that.

Lemma 7 (Consistency Lemma) *Let $u \in \mathbf{nodes}_i^*$ for some $0 \leq i < L$. Define $\mathcal{S}^*$ to be the multi-set of 'good' cubes: i.e. cubes that label good offspring of u:*

$$\mathcal{S}^* \stackrel{def}{=} \left\{ \mathcal{C}_v \in \mathbf{labels}(u) \,\middle|\, v \in \mathbf{nodes}_{i+1}^* \right\}.$$

If for every good offspring v of u there is a super-LDF $\tilde{\mathcal{G}}_v$ over $\mathcal{C}_v$, of total degree $\leq r = \tilde{r}_i/(D+2)$ and norm $\|\tilde{\mathcal{G}}_v\| \leq 2^{L-i-1} \cdot \mathrm{Avg}(v)$, such that

$$\Pr_{x \in \mathcal{C}_v} (\pi_x(\tilde{\mathcal{G}}_v) = m(x)) \geq 1 - \alpha$$

then there is a super-LDF $\mathcal{G}_u$ over $\mathbf{dom}_u$ of total degree $\tilde{r}_i = r(D+2)$ and norm $\|\mathcal{G}_u\| \leq 2^{L-i} \cdot \mathrm{Avg}(u)$ that obeys

$$\Pr_{\mathcal{C}_v \in \mathcal{S}^*} \left(\pi_{\mathcal{C}_v}(\mathcal{G}_u) = \tilde{\mathcal{G}}_v\right) \geq 1 - \alpha/6$$

We defer the proof of this lemma to the next subsection, and continue with the proof of Lemma 5. As previously mentioned, the super-LDFs $\tilde{\mathcal{G}}_v$ obtained for the $\mathcal{C}_v$s were of total degree $\tilde{r}_1 \cdot d(b_1)^3$, hence the degree is:

$$\begin{aligned}
\tilde{r}_1 \cdot d(b_1)^3 &= A^{L-1}(r_1+1) \cdot d(b_1)^3 = A^{L-1} \cdot d \cdot (b_1)^4 = \\
&= A^{L-1} \cdot 4(D+2) \cdot (r_0(D+2)+1) \\
&< A^{L-1} \cdot 4(D+2)^2(r_0+1) = A^L(r_0+1)/(D+2) = \tilde{r}_0/(D+2)
\end{aligned}$$

using the definitions $A = 4(D+2)^3$, $\tilde{r}_i = A^{L-i}(r_i+1)$, $b_{i+1} = \sqrt[4]{r_i(D+2)+1}$ and $r_{i+1} = b_{i+1} - 1$. Hence we obtain from the consistency lemma a global super-LDF $\mathcal{G}_\varphi$ of degree $\tilde{r}_0$ over $\mathcal{F}^{d_0}$ that agrees with $\tilde{\mathcal{G}}_v$ for $1 - \alpha/6$ of the good offspring v of $u = root_\varphi$.

We next show that $\mathcal{G}_\varphi = \mathcal{G}_{\varphi'}$ for every $\varphi \neq \varphi'$ whose corresponding nodes $root_\varphi$ and $root_{\varphi'}$ are both good. Choose a random offspring v of $root_\varphi$, (by choosing a random label $\mathcal{C}_v \in_R$ **labels**$(root_\varphi)$), and a random point $x \in \mathcal{C}_v \subset \mathcal{F}^{d_0}$. We claim that $\Pr_{x,\mathcal{C}_v}(\pi_x(\tilde{\mathcal{G}}_v) = m(x)) \geq 1 - 2\alpha$. By Proposition 4 the probability that v is not good is $\leq \alpha/3$. If v is good, the above Lemma 7 tells us that with probability at most $\alpha/6$, $\pi_{\mathcal{C}_v}(\mathcal{G}_\varphi) \neq \tilde{\mathcal{G}}_v$ (altogether we have that with probability $\geq 1 - \alpha/3 - \alpha/6 \geq 1 - \alpha$ over the cubes in **labels**$(root_\varphi)$, $\pi_{\mathcal{C}_v}(\mathcal{G}_\varphi) = \tilde{\mathcal{G}}_v$). Now, by Lemma 6, for any good v, $\Pr_{x \in \mathcal{C}_v}(\pi_x(\tilde{\mathcal{G}}_v) \neq m(x)) \leq \alpha$. For all otherwise chosen points, we have $\pi_x(\mathcal{G}_\varphi) = m(x)$, and the claim is proven.

These points constitute roughly [2] a $1 - 2\alpha$ fraction of $\mathcal{F}^{d_0}$. Hence $\mathcal{G}_\varphi$ and $\mathcal{G}_{\varphi'}$ agree with m on the same $\geq 1 - 4\alpha > 1/2$ fraction of the points. Thus, the super-LDF $\mathcal{G}_\varphi - \mathcal{G}_{\varphi'}$ (subtraction is defined as subtraction of the coefficient vectors) is zero when projected on more than half of the points.

Now, utilizing the fact that $\|\mathcal{G}_\varphi - \mathcal{G}_{\varphi'}\| \leq \|\mathcal{G}_\varphi\| + \|\mathcal{G}_{\varphi'}\| \leq 2^{L+1}g$, and by the low-ambiguity property (see Proposition 1) the fraction of ambiguous points (the only candidates on which the projection can be zero) is bounded by

$$\text{amb}(\tilde{r}_0, d_0, 2^{L+1}g) < 2^{2(L+1)}g^2\frac{\tilde{r}_0 d_0}{|\mathcal{F}|} \leq 2^{2(L+1)} \cdot A^L \cdot d_0 \cdot |\mathcal{F}|^{2c_g + c/c_1 - 1} \ll 1/2$$

Thus, we deduce that $\mathcal{G}_\varphi = \mathcal{G}_{\varphi'}$. In addition, $\mathcal{G}_\varphi$ must be non-trivial since $m(x) \neq \vec{0}$ for every x.

We choose an arbitrary LDF P that appears in $\mathcal{G} \stackrel{def}{=} \mathcal{G}_{\varphi'} \neq \vec{0}$ for some good $root_{\varphi'}$, and define an assignment $\mathcal{A}_P : V_\Phi \to \{\texttt{true}, \texttt{false}\}$ for the variables of Φ as follows. For each $v \in V_\Phi$, we define $\mathcal{A}_P(v) \stackrel{def}{=} \texttt{true}$ iff $P(x) = 0$ on the corresponding point $x = \mathbf{var}_{root_\varphi}{}^{-1}(v)$ (see Definition 11), and $\mathcal{A}_P(v) \stackrel{def}{=} \texttt{false}$ otherwise.

The fraction of tests $\varphi \in \Phi$ for which $root_\varphi$ is good is at least $1 - 1/C > 1/2$ (because the total average of $\text{Avg}(root_\varphi)$ over all $\varphi \in \Phi$ is g, and a good root node $root_\varphi$ is defined as a node with $\text{Avg}(root_\varphi) \leq C \cdot g$).

We will show that $\mathcal{A}_P$ satisfies φ for every good node $root_\varphi$, and thus Φ is totally satisfiable. Let $\varphi \in \Phi$ be such that $root_\varphi$ is a good node. By the above consistency lemma, we know that for $1 - \alpha/6$ of the good offspring v of $root_\varphi$, $\pi_{\mathcal{C}_v}(\mathcal{G}) = \tilde{\mathcal{G}}_v$, and unless P is cancelled on $\mathcal{C}_v$, $P|_{\mathcal{C}_v}$

[2] This procedure is *almost* equivalent to choosing a point uniformly at random, however there is a small (negligible) bias in favor of points in the span of **dst**$(root_\varphi)$

appears in $\tilde{\mathcal{G}}_v$. P will be cancelled on $\mathcal{C}_v$ only if there is another LDF Q appearing in $\mathcal{G}$ whose restriction to $\mathcal{C}_v$ equals P's restriction. For each Q the probability for this is bounded by (see Proposition 2) $\frac{\tilde{r}_0 d_0}{|\mathcal{F}|}$. Since there are no more than $2^K g$ possible LDFs Q that appear in $\mathcal{G}$, P is cancelled with probability $\leq \frac{\tilde{r}_0 d_0}{|\mathcal{F}|} \cdot 2^K g \ll 1/2$.

Thus there exists at least one good offspring v of $root_\varphi$ for which $\pi_{\mathcal{C}_v}(\mathcal{G}) = \tilde{\mathcal{G}}_v$ and $P|_{\mathcal{C}_v}$ appears in $\tilde{\mathcal{G}}_v$. Recall that the distinguished points of each offspring v of a root node $root_\varphi$, $\mathbf{dst}(v)$, are mapped to φ's variables. In addition, since v is good, Lemma 6 ensures that $\tilde{\mathcal{G}}_v$ is legal, i.e. for every Q appearing in $\tilde{\mathcal{G}}_v$, Q's restriction to φ's variables satisfies φ. It follows that φ is satisfied by $\mathcal{A}_P$. ∎

This completes the proof of soundness, Lemma 5.

We now fill in the proof of Lemma 6.

Proof: (of Lemma 6) We prove this statement by induction on $L - i$. We ascend from the leaves to the top level, obtaining a super-LDF for each good node from the super-LDFs of its good offspring.

For obtaining the base of the induction ($i = L$), recall that for every leaf $u \in \mathbf{nodes}^*_L$, the test ψ_u is assigned a super-LDF $\mathcal{SA}(\psi_u)$. The definition of $\mathbf{nodes}^*_L$ implies $\|\mathcal{SA}(\psi_u)\| \leq g_L$. Since $\mathcal{SA}$ is a consistent super-assignment, $\mathcal{SA}(\psi_u)$ agrees with m on all of $\mathbf{dom}_u$ (in particular with all of $\mathbf{manf}(u)$ and $\mathbf{dst}(u)$), and thus the base of the induction is established.

To see the inductive step ($1 \leq i < L$), let $u \in \mathbf{nodes}^*_i$ be a good level-i node. By the inductive hypothesis for $L - i - 1$, every good offspring v of u has a legal super-LDF $\mathcal{G}_v$ (of degree $\tilde{r}_{i+1}$) with norm $\leq 2^{L-i-1} \cdot \mathrm{Avg}(v)$ such that $\mathcal{G}_v$ agrees with m on $\mathbf{dst}(v)$ and on $1 - \alpha$ of the points in $\mathbf{manf}(v)$.

Let v be a good offspring of u. We define $\tilde{\mathcal{G}}_v$, as in the proof of Lemma 5, to be the same linear combination as $\mathcal{G}_v$, taking LDFs $\tilde{P}$ instead of P, where $\tilde{P} : \mathcal{C}_v \to \mathcal{F}$ is defined by $\forall x \in \mathcal{C}_v\ \tilde{P}(x) \stackrel{def}{=} P(E_{b_{i+1}}(x))$. It follows from definition 8 and from the inductive hypothesis that $\tilde{\mathcal{G}}_v$ is a super-LDF of total degree $d\tilde{r}_{i+1} \cdot (b_{i+1})^3$ as before. As before, this is bounded by $\leq \tilde{r}_i/(D+2)$.

For any good node v labelled by $\mathcal{C}_v$, let $x \in \mathcal{C}_v$, and let $y = E_{b_{i+1}}(x) \in \mathbf{manf}(v) \subset \mathbf{dom}_v$. Recall that we abbreviated $m(x)$ to mean $m(\mathbf{var}_v(x))$. Furthermore, we defined $\mathbf{var}_v(y) = \mathbf{var}_v(E_{b_{i+1}}(x)) = \mathbf{var}_u(x)$, thus $m(y) = m(x)$. Note also that by definition of $\tilde{\mathcal{G}}_v$, $\pi_x(\tilde{\mathcal{G}}_v) = \pi_y(\mathcal{G}_v)$.

By the inductive hypothesis, and since $E_{b_{i+1}}$ bijects $\mathcal{C}_v$ to $\mathbf{manf}(v)$, we have that the equality:

$$\pi_x(\tilde{\mathcal{G}}_v) = \pi_y(\mathcal{G}_v) = m(y) = m(x)$$

holds both (1) for $1 - \alpha$ of the points $x \in \mathcal{C}_v$, and (2) for every point $x \in \mathcal{C}_v$ such that $y = E_{b_{i+1}}(x) \in \mathbf{dst}(v)$.

By (1), and applying the consistency lemma (Lemma 7), we deduce a global super-LDF $\mathcal{G}_u$ over $\mathbf{dom}_u$ of norm $\|\mathcal{G}_u\| \leq 2^{L-i} \cdot \mathrm{Avg}(u)$ and of degree $\tilde{r}_i$ such that for $1 - \alpha/6$ of the cubes in $\mathcal{S}^*$,

$$\pi_{\mathcal{C}_v}(\mathcal{G}_u) = \tilde{\mathcal{G}}_v \qquad (*)$$

This constitutes at least $1 - \alpha/6 - 1/C = 1 - \alpha/2$ of all the cubes in $\mathbf{labels}(u)$ (there are no more than $1/C$ cubes outside $\mathcal{S}^*$, see Proposition 4). Now recall from the construction of the offspring-labels (see Definition 14) that at least one half of the offspring v of u are labelled by cubes in $\mathbf{labels}_1(u)$. We deduce that $1 - 2 \cdot \alpha/2 = 1 - \alpha$ of these cubes obey $(*)$. Similarly,

$\mathbf{labels}_2(u)$ make up at least one third of the total number of labels, thus $(*)$ holds for $\approx 1 - \frac{3}{2}\alpha$ of them (and in particular for at least one cube in $\mathbf{labels}_2(u)$, which is all we'll need).

Recall from the construction of the offspring-labels that the cubes $\mathcal{C}_v \in \mathbf{labels}_2(u)$ have $\mathbf{dst}(v) = E_{b_{i+1}}(\mathbf{dst}(u))$. By Definition 12, these points are mapped to the exact same variables: $\mathbf{var}_u(\mathbf{dst}(u)) = \mathbf{var}_v(\mathbf{dst}(v))$. As long as there exists one good offspring v of u, we'll have for every $x \in \mathbf{dst}(u)$, $y = E_{b_{i+1}}(x) \in \mathbf{dst}(v)$, and by (2), $\pi_x(\tilde{\mathcal{G}}_v) = m(x)$. We have shown,

$$\forall x \in \mathbf{dst}(u) \quad \pi_x(\mathcal{G}_u) = m(x)\,.$$

We have left to show that

$$\Pr_{x \in \mathbf{manf}(u)}(\pi_x(\mathcal{G}_u) = m(x)) \geq 1 - \alpha$$

Consider the second half of u's offspring, labelled by a label from

$$\mathbf{labels}_1(u) = \bigcup_{\mathrm{x} \in \mathbf{var}_u(\mathbf{dom}_u) \setminus V_u} \mathbf{labels}_{\mathrm{x}}(u)$$

These offspring are actually divided into $|\mathbf{manf}(u)|$ parts, one per each point $x \in \mathbf{manf}(u)$ (with the correspondence $\mathbf{var}_u(x) = \mathrm{x}$). By definition, the offspring v in x's sub-part have $\mathbf{dst}(v) \stackrel{def}{=} \left\{E_{b_{i+1}}(x)\right\}$. We have shown (recall (2) from before) that $\pi_x(\tilde{\mathcal{G}}_v) = m(x)$ holds for any x such that $y = E_{b_{i+1}}(x) \in \mathbf{dst}(v)$. Hence for every $x \in \mathbf{manf}(u)$, if there is a good node v labelled $\mathcal{C}_v \in \mathbf{labels}_{\mathbf{var}_u(x)}(u)$, then $\pi_x(\tilde{\mathcal{G}}_v) = m(x)$.

As shown above, $1 - \alpha$ of u's offspring labelled by cubes in $\mathbf{labels}_1(u)$ are both good, and obey $\pi_{\mathcal{C}_v}(\mathcal{G}_u) = \tilde{\mathcal{G}}_v$. Hence for $1 - \alpha$ of the points $x \in \mathbf{manf}(u)$ there must be a good offspring v labelled by $\mathcal{C}_v \in \mathbf{labels}_{\mathbf{var}_u(x)}(u)$ for which $\pi_{\mathcal{C}_v}(\mathcal{G}_u) = \tilde{\mathcal{G}}_v$. In this case,

$$\pi_x(\mathcal{G}_u) = \pi_x(\pi_{\mathcal{C}_v}(\mathcal{G}_u)) = \pi_x(\tilde{\mathcal{G}}_v) = m(x)$$

This establishes,

$$\Pr_{x \in \mathbf{manf}(u)}(\pi_x(\mathcal{G}_u) = m(x)) \geq 1 - \alpha$$

■

5.3 The Consistency Lemma

In this subsection we prove the consistency lemma, that allows us to deduce one global super-LDF for any good node, assuming "consistent" LDFs on its good offspring.

Lemma 7 (Consistency Lemma) *Let $u \in \mathbf{nodes}^*_i$ for some $0 \leq i < L$. Define $\mathcal{S}^*$ to be the multi-set of cubes that label good offspring of u, $\mathcal{S}^* \stackrel{def}{=} \left\{\mathcal{C}_v \in \mathbf{labels}(u) \,\middle|\, v \in \mathbf{nodes}^*_{i+1}\right\}$. If for every good offspring v of u there is a super-LDF $\tilde{\mathcal{G}}_v$ over $\mathcal{C}_v$, of total degree $\leq r = \tilde{r}_i/(D+2)$ and norm $\|\tilde{\mathcal{G}}_v\| \leq 2^{L-i-1} \cdot \mathrm{Avg}(v)$, such that*

$$\Pr_{x \in \mathcal{C}_v}(\pi_x(\tilde{\mathcal{G}}_v) = m(x)) \geq 1 - \alpha$$

then there is a super-LDF $\mathcal{G}_u$ over $\mathbf{dom}_u$ of total degree $\tilde{r}_i = r(D+2)$ and norm $\|\mathcal{G}_u\| \le 2^{L-i} \cdot$ $\text{Avg}(u)$ that obeys

$$\Pr_{\mathcal{C}_v \in \mathcal{S}^*}\left(\pi_{\mathcal{C}_v}(\mathcal{G}_u) = \tilde{\mathcal{G}}_v\right) \ge 1 - \alpha/6$$

Proof: Throughout the following proof, we make no effort to minimize the constants, but rather to shorten the mathematical expressions in which they appear.

Unless otherwise mentioned, $\mathcal{C}_v$ will denote the cube labeling the node v.

For simplicity, assume $\mathbf{dom}_u = \mathcal{F}^d$ ($\mathcal{F}^{d_0}$ in case $i = 0$).

An $[r,d]$-LDF $P : \mathcal{F}^d \to \mathcal{F}$ is called *permissible with coefficient c_P* if $c_P \neq 0$ and for at least $2/3$ of the cubes $\mathcal{C}_v \in \mathcal{S}^*$, $\tilde{\mathcal{G}}_v[P|_{\mathcal{C}_v}] = c_P$. In this case we sometimes say that P appears with coefficient c_P in the cube $\mathcal{C}_v$. Obviously, an LDF P can be permissible with at most one coefficient c_P. We define the global super-LDF $\mathcal{G}_u$ by

$$\forall P \in LDF_{r,d}: \quad \mathcal{G}_u[P] \stackrel{def}{=} \begin{cases} c_P & P \text{ is permissible with } c_P \\ 0 & P \text{ isn't permissible} \end{cases}$$

We claim that $\mathcal{G}_u$ is the desired global super-LDF. We first claim that $\|\mathcal{G}_u\| \le 2 \cdot 2^{L-i-1}\text{Avg}(u)$.

Proposition 5 *The norm of $\mathcal{G}_u$ is bounded by $2 \cdot 2^{L-i-1}\text{Avg}(u) = 2^{L-i}\text{Avg}(u)$.*

Proof: Denote by $P_1, .., P_a$ the permissible LDFs (if $a = 0$ we're done), and denote $c_i \stackrel{def}{=} \mathcal{G}_u[P_i]$. Let us consider the average N of the norms $\|\tilde{\mathcal{G}}_v\|$,

$$\begin{aligned} N &\stackrel{def}{=} \frac{1}{|\mathcal{S}^*|}\sum_{\mathcal{C}_v \in \mathcal{S}^*}\|\tilde{\mathcal{G}}_v\| \le \frac{1}{|\mathcal{S}^*|}\sum_{\mathcal{C}_v \in \mathcal{S}^*} 2^{L-i-1}\text{Avg}(v) \le 2^{L-i-1}\text{Avg}(u) \\ &\le 2^{L-i-1}g_i = 2^{L-i-1} \cdot g \cdot C^{i+1} < |\mathcal{F}|^{c_g} \cdot C^L \ll \sqrt{|\mathcal{F}|} \end{aligned}$$

where the second inequality in the first line is true since averaging the norm over all of the offspring is at least as large as the average of the good offspring.

We will lower bound N as follows. P_1 appears with c_1 in the super-LDFs of $\ge 2/3$ of the good offspring, which means $N \ge 2/3 \cdot |c_1|$. P_2 appears with c_2 in the super-LDFs of $\ge 2/3$ of the good offspring, however some of its appearances can coincide with those of P_1. Denote by γ the maximal fraction of cubes in $\mathcal{S}^*$ on which possibly $P_1|_{\mathcal{C}} = P_2|_{\mathcal{C}}$. We know from Proposition 4 that $\mathcal{S}^*$ make up more than half of all cubes, and by Proposition 2 $P_1|_{\mathcal{C}} = P_2|_{\mathcal{C}}$ on no more than $\le \frac{rd}{|\mathcal{F}|}$ of all of the cubes, thus $\gamma \le \frac{2rd}{|\mathcal{F}|}$. Hence $N \ge 2/3 \cdot |c_1| + (2/3 - \gamma) \cdot |c_2|$. Continuing in this manner, P_3 adds at least $(2/3 - 2\gamma) \cdot |c_3|$ and we obtain

$$\forall 1 \le j \le a \qquad N \ge \sum_{i=1}^{j}\left(\frac{2}{3} - (i-1)\gamma\right) \cdot |c_i|$$

If $a \ge \frac{1}{6\gamma} \ge \frac{|\mathcal{F}|}{12rd}$, we get $N \ge \sum_{i=1}^{1/6\gamma}\left(\frac{2}{3} - \frac{1}{6\gamma}\gamma\right) \cdot |c_i| \ge \frac{1}{6\gamma} \cdot \frac{1}{2} \cdot 1 \gg \sqrt{|\mathcal{F}|}$, a contradiction. Thus $a < 1/6\gamma$, and

$$N \ge \sum_{i=1}^{a}\left(\frac{2}{3} - (i-1)\gamma\right) \cdot |c_i| \ge \frac{1}{2}\sum_{i=1}^{a}|c_i| = \frac{1}{2}\|\mathcal{G}_u\|$$

and indeed $\|\mathcal{G}_u\| \leq 2N \leq 2 \cdot 2^{L-i-1}\text{Avg}(u) = 2^{L-i}\text{Avg}(u)$. ■

We have left to show that for almost all of the cubes $\mathcal{C}_v \in \mathcal{S}^*$, $\pi_{\mathcal{C}_v}(\mathcal{G}_u) = \tilde{\mathcal{G}}_v$.

Let us define, for every good node v, the remainder super-LDF: $\mathcal{R}_v \stackrel{def}{=} \tilde{\mathcal{G}}_v - \pi_{\mathcal{C}_v}(\mathcal{G}_u)$ (the definition of $\mathcal{G}_u$ implies that every LDF P appearing in it has degree $\leq r$; subtraction is defined as usual subtraction of two vectors in $\mathbb{Z}^{LDF_{r,D+2}}$). Assume, for contradiction, that for at least an $\alpha/6$ fraction of the good nodes, $\mathcal{R}_v \neq \vec{0}$. We will derive a contradiction by finding an LDF P such that $P|_{\mathcal{C}_v}$ appears with the same coefficient $c_P \neq 0$ in $\mathcal{R}_v$ in at least $2/3 + \gamma\|\mathcal{G}_u\|$ fraction of the good nodes v. This LDF P can agree with another LDF in $\mathcal{G}_u$ on at most $\gamma\|\mathcal{G}_u\|$ fraction of the good cubes. Hence on at least $2/3$ of the good cubes, $c'_P \stackrel{def}{=} \tilde{\mathcal{G}}_v[P|_{\mathcal{C}_v}] = c_P + \mathcal{G}_u[P]$, which implies that P is permissible with coefficient c'_P, so by definition $\mathcal{G}_u[P] = c'_P$, hence $c_P = 0$, a contradiction.

For every $x \in \mathcal{F}^d$, define $m_R(x) \stackrel{def}{=} m(x) - \pi_x(\mathcal{G}_u)$. Obviously $m_R(x) = \pi_x(\mathcal{R}_v)$ if and only if $m(x) = \pi_x(\tilde{\mathcal{G}}_v)$. (This happens for at least $1 - \alpha$ of the points $x \in \mathcal{C}_v$ for every $\mathcal{C}_v \in \mathcal{S}^*$, by the conditions of the lemma).

Proposition 6 *Let $\mathcal{R}_v \stackrel{def}{=} \tilde{\mathcal{G}}_v - \pi_{\mathcal{C}_v}(\mathcal{G}_u)$ be as before. There exists an $[r, D+2]$-LDF P and a coefficient $c_P \neq 0$ such that*

$$\Pr_{\mathcal{C}_v \in \mathcal{S}^*}(\mathcal{R}_v[P|_{\mathcal{C}_v}] = c_P) > \delta$$

where $\delta = \Omega\left((\frac{\alpha}{s})^9\right)$ and $s \stackrel{def}{=} 2^{L-i}\text{Avg}(u)$.

Proof: Consider the following random procedure:

1. For every cube $\mathcal{C}_v \in \textbf{labels}(u)$ choose a random LDF from the set $\{Q \in LDF_{r,D+2} \mid \mathcal{R}_v[Q] \neq 0\}$. If $\mathcal{C}_v \notin \mathcal{S}^*$ or this set is empty, choose nothing.
2. For every point $x \in \textbf{dom}_u$ choose a random value from the set $\{a \in \mathcal{F} \mid m_R(x)[a] \neq 0\}$. If this set is empty, choose nothing.
3. Choose a random cube $\mathcal{C}_v \in \textbf{labels}(u)$ and a random point $x \in \mathcal{C}_v$. If no value is chosen for either the point or the cube, the procedure fails.

We are interested in pairs of good cube and point on it, on which the procedure doesn't fail, and that have relatively few possible values to choose from, and that are consistent. We eliminate 'bad' pairs as follows.

For a cube $\mathcal{C}_v \in \textbf{labels}(u)$ define $E_1(\mathcal{C}_v)$ to be the predicate that evaluates to `true` iff $\mathcal{C}_v \in \mathcal{S}^*$ and the set $\{Q \in LDF_{r,D+2} \mid \mathcal{R}_v[Q] \neq 0\}$ is non-empty. $\Pr_{\mathcal{C}_v \in \textbf{labels}(u)}(E_1(\mathcal{C}_v)) \geq (1 - 1/C) \cdot \alpha/6$ because $1 - 1/C$ of the cubes are in $\mathcal{S}^*$ (since u is good), and we assumed for contradiction that for $\alpha/6$ of these cubes $\mathcal{R}_v$ is non-trivial.

For a cube $\mathcal{C}_v \in \textbf{labels}(u)$ define $E_2(\mathcal{C}_v)$ to be the predicate that evaluates to `true` iff $E_1(\mathcal{C}_v)$ is true and also $\|\tilde{\mathcal{G}}_v\| \leq 2 \cdot \frac{6s}{\alpha(1-1/C)}$ where $s = 2^{L-i}\text{Avg}(u)$ bounds the average of $\|\tilde{\mathcal{G}}_v\|$ taken over nodes $v \in \mathcal{S}^*$. We note that the average norm $\|\tilde{\mathcal{G}}_v\|$ taken over cubes for which E_1 is true does not exceed $\frac{6s}{\alpha(1-1/C)}$. The standard Markov argument shows

$$\Pr_{\mathcal{C}_v \in \textbf{labels}(u)}(E_2(\mathcal{C}_v)) \geq \frac{1}{2} \cdot \Pr_{\mathcal{C}_v \in \textbf{labels}(u)}(E_1(\mathcal{C}_v)) \geq (1 - 1/C) \cdot \alpha/12$$

By the triangle inequality, $\|\mathcal{R}_v\| = \|\tilde{\mathcal{G}}_v - \pi_{\mathcal{C}_v}(\mathcal{G}_u)\| \le \|\tilde{\mathcal{G}}_v\| + \|\mathcal{G}_u\| \le \|\tilde{\mathcal{G}}_v\| + s$ hence the cubes $\mathcal{C}_v$ for which $E_2(\mathcal{C}_v) = \texttt{true}$ have $\|\mathcal{R}_v\| \le 12s/\alpha(1-1/C) + s < 13s/\alpha$.

For a cube $\mathcal{C}_v \in \mathbf{labels}(u)$, and a point $x \in \mathcal{C}_v$, define $E_3(\mathcal{C}_v, x)$ to be the predicate that evaluates to $\texttt{true}$ iff $E_2(\mathcal{C}_v) = \texttt{true}$ and also $x \in \mathcal{C}_v$ and x obeys $m_R(x) = \pi_x(\mathcal{R}_v)$ *and* $\mathcal{R}_v$ is not ambiguous on x. We say, in this case, that the point x and the cube $\mathcal{C}_v$ *agree non-ambiguously.* Since for every good cube $\mathcal{C}_v$ no more than α fraction of its points have $m_R(x) \neq \pi_x(\mathcal{R}_v)$, and no more than $\mathrm{amb}(r, D+2, \|\mathcal{R}_v\|)$ are ambiguous, it follows that

$$\begin{aligned}\Pr_{\substack{\mathcal{C}_v \in \mathbf{labels}(u)\\ x \in \mathcal{C}_v}}(E_3(\mathcal{C}_v, x)) &\ge \Pr_{\mathcal{C}_v}(E_2(\mathcal{C}_v)) \cdot (1 - \alpha - \mathrm{amb}(r, D+2, \|\mathcal{R}_v\|)) \\ &\ge (1-1/C) \cdot \alpha/12 \cdot (1 - \alpha - |\mathcal{F}|^{-\frac{1}{2}}) \\ &> \alpha/100\end{aligned}$$

(the second inequality follows from $\mathrm{amb}(r, D+2, \|\mathcal{R}_v\|) \le \frac{\tilde{r}_i/(D+2)\cdot(D+2)}{|\mathcal{F}|} \cdot \|\mathcal{R}_v\|^2 \ll |\mathcal{F}|^{-\frac{1}{2}}$).
The pairs of point x and cube $\mathcal{C}_v$ for which $E_3(\mathcal{C}_v, x) = \texttt{true}$ are pairs that agree non-ambiguously, and for which $\|\mathcal{R}_v\| \le 13s/\alpha$.

For a cube $\mathcal{C}_v \in \mathbf{labels}(u)$, a point $x \in \mathcal{C}_v$, an $[r, D+2]$-LDF Q (viewed as an LDF over $\mathcal{C}_v$) and a value $a \in \mathcal{F}$, define $E_4(\mathcal{C}_v, x, Q, a)$ to be the predicate that evaluates to $\texttt{true}$ iff $E_3(\mathcal{C}_v, x) = \texttt{true}$ and also $Q(x) = a$. We will lower bound the probability $\Pr_{\mathcal{C}_v, x, Q, a}(E_4(\mathcal{C}_v, x, Q, a))$ where $\mathcal{C}_v \in \mathbf{labels}(u)$, $x \in \mathcal{C}_v$, and Q and a are chosen according to the random procedure described in the beginning of the proof (i.e. Q is chosen uniformly from the set $\{Q \in LDF_{r,D+2} \mid \mathcal{R}_v[Q] \neq 0\}$, and a uniformly from the set $\{a \in \mathcal{F} \mid m_R(x)[a] \neq 0\}$. Note that when E_4 is true, there are no more than $\frac{13s}{\alpha}$ LDFs that appear in $\mathcal{R}_v$. Since $E_4(\mathcal{C}_v, x, Q, a) = \texttt{true}$ implies by definition $E_3(\mathcal{C}_v, x) = \texttt{true}$ we know that $m_R(x) = \pi_x(\mathcal{R}_v)$, hence any value a randomly chosen for x has a "matching value" in the set $\{Q \in LDF_{r,D+2} \mid \mathcal{R}_v[Q] \neq 0\}$. This value is the chosen one with probability at least $\frac{\alpha}{13s}$. Thus,

$$\Pr(E_4(\mathcal{C}_v, x, Q, a)) \ge \Pr(E_3(\mathcal{C}_v, x)) \cdot \frac{\alpha}{13s}$$

Finally, note that if $E_4(\mathcal{C}_v, x, Q, a) = \texttt{true}$, then $\mathcal{R}_v[Q] = m_R(x)[a]$ because the cube and point agree non-ambiguously. Also, since in this case $\|\mathcal{R}_v\| \le 13s/\alpha$, the coefficient $\mathcal{R}_v[Q]$ can be any value from the set $B \stackrel{def}{=} \{\pm 1, ..., \pm 13s/\alpha\}$, $26s/\alpha$ values in all. Denote by $E_c(\mathcal{C}_v, x, Q, a)$ the predicate that is the same as E_4 except that it evaluates to true only if in addition, $\mathcal{R}_v[Q] = m_R(x)[a] = c$. There must be at least one value $c_0 \in B$ for which

$$\Pr(E_{c_0}(\mathcal{C}_v, x, Q, a)) \ge \frac{\alpha}{26s} \cdot \Pr(E_4(\mathcal{C}_v, x, Q, a)) \ge \frac{\alpha}{26s} \cdot \frac{\alpha}{13s} \cdot \Pr(E_3(\mathcal{C}_v, x)) \ge \frac{\alpha^2}{338s^2} \cdot \frac{\alpha}{100} = \Omega(\frac{\alpha}{s})^3$$

We now apply the following corollary of [RS97],

Lemma 8 *Let* $\rho = (\frac{rd}{\mathcal{F}})^c$ *for some constant* $c > 0$, *and let* $\mathcal{S} = \mathbf{labels}(u)$ *for labels*(u) *as above. Let* $\mathcal{A} : \mathcal{S} \to \mathrm{LDF}_{r,D+2}$ *be an assignment of* $[r, D+2]$*-LDF per cube, and let* $\mathcal{A}_0 : \mathcal{F}^d \to \mathcal{F}$ *be an assignment of value per point. If*

$$\Pr_{\mathcal{C} \in_R \mathcal{S}, x \in_R \mathcal{C}}(\mathcal{A}[\mathcal{C}](x) = \mathcal{A}_0[x]) \ge \rho$$

then there is an $[r, d]$*-LDF* P *for which* $\Pr_{\mathcal{C} \in \mathcal{S}}(P|_{\mathcal{C}} = \mathcal{A}[\mathcal{C}]) \ge \rho^3$.

We omit the proof of this lemma, and note that a very similar cube vs. point version appears in [DFK+99]. We apply this lemma as follows. We take $\mathcal{S} = \mathbf{labels}(u)$. For every cube $\mathcal{C}_v$ whose selected value Q has $\mathcal{R}_v[Q] = c_0$, assign $\mathcal{A}[\mathcal{C}_v] = Q$, otherwise let $\mathcal{A}[\mathcal{C}_v]$ be a totally random value. For each point $x \in \mathcal{F}^d$, we define $\mathcal{A}_0[x]$ to be the value selected for it in the random procedure. Again, if no value was selected, we assign a totally random value. We have

$$\Pr_{\mathcal{C}\in\mathcal{S}, x\in\mathcal{C}}(\mathcal{A}[\mathcal{C}](x) = \mathcal{A}_0[x]) \geq \Pr(E_{c_0}(\mathcal{C}, x, \mathcal{A}[\mathcal{C}], \mathcal{A}_0[x]))$$

The probability on the right hand side is taken over a random choice of cube $\mathcal{C} \in \mathcal{S}$ and point $x \in \mathcal{C}$, and over the random choices made when defining $\mathcal{A}[\mathcal{C}]$ and $\mathcal{A}_0[x]$. It follows easily that this probability is at least $\geq \Pr(E_{c_0})$. Thus we obtain using Lemma 8 an *LDF* P that agrees with $\geq (\Pr(E_{c_0}))^3 \geq \Omega\left((\frac{\alpha}{s})^9\right)$ fraction of the cubes and their chosen values. Of the randomly assigned cubes, P is expected to appear in less than $1/|\mathcal{F}|$. Thus at least half of the cubes in which P appears also obey $\mathcal{R}_v[P|_{\mathcal{C}_v}] = c_0$. These cubes make up at least $\delta = \frac{1}{2}\cdot(\Pr(E_{c_0}))^3 = \Omega\left((\frac{\alpha}{s})^9\right)$ of the good cubes. ■

We have found a polynomial P that appears (with the same coefficient $c_0 \neq 0$) in a non-negligible fraction of the cubes (i.e. $\mathcal{R}_v[P|_{\mathcal{C}_v}] = c_0$ for a non-negligible $\geq \delta$ fraction of the good offspring v of u). We now show that P, in fact, appears in most of the points with coefficient c_0,

Proposition 7 *For at least half of the points $x \in \mathbf{dom}_u$, $m_R(x)[P(x)] = c_0$.*

Proof: Let $N = \{x \in \mathbf{dom}_u \mid m_R(x)[P(x)] \neq c_0\}$ be the set of points where P does not appear with coefficient c_0. We shall prove that $\mu \stackrel{def}{=} \frac{|N|}{|\mathbf{dom}_u|} < \frac{1}{2}$. We now state a hitting lemma that shows that if N is not too small, then almost all of the cubes must hit a non-negligible fraction of the points in N.

Lemma 9 (Hitting Lemma) *Let $0 < \beta < 1$ and let $\mathcal{D} = \mathcal{F}^d$. Let $N \subset \mathcal{D}$ be a set of points, $|N| \geq \beta|\mathcal{D}|$. Most $(1 - \frac{8}{\beta|\mathcal{F}|})$ cubes in* $\mathbf{labels}(u)$ *(for u as above) have at least $\frac{\beta}{2}$ of their points in N.*

The proof of this lemma is easily obtained using the pairwise independence of points in a random cube, and is omitted (special care should be given to the fact that the points in these cubes are distributed only *almost* uniformly: certain points – e.g. $\mathrm{span}(\mathbf{dst}(u))$ – appear more often than others).

We now know that $1 - \frac{8}{\mu|\mathcal{F}|}$ of the cubes in $\mathbf{labels}(u)$ ($1 - \frac{16}{\mu|\mathcal{F}|}$ fraction of $\mathcal{S}^*$, since $|\mathcal{S}^*| > \frac{1}{2}|\mathbf{labels}(u)|$) have $\frac{\mu}{2}$ of their points from N. Consider only cubes $\mathcal{C}_v$ whose norm isn't too large – $\|\mathcal{R}_v\| \leq 2s/(\delta/2)$ (the average of $\|\mathcal{R}_v\|$ over all nodes $v \in \mathcal{S}^*$ is $\leq s + s = 2s$, hence we are ignoring a $\delta/2$ fraction). If $\frac{\mu}{2} > \alpha + \mathrm{amb}(r, D+2, 2s/(\delta/2))$ then every such cube must agree non-ambiguously with at least one point from N. This implies that P does not appear in these cubes (that constitute at least $1 - \frac{16}{\mu|\mathcal{F}|} - \delta/2$ fraction of $\mathcal{S}^*$) with coefficient c_0, and hence, $\delta \leq \frac{16}{\mu|\mathcal{F}|} + \delta/2$. Altogether we have that

$$\mu \leq \max\left(2(\mathrm{amb}(r, D+2, 2s/(\delta/2)) + \alpha), \frac{32}{\delta|\mathcal{F}|}\right) < \frac{1}{2}$$

■

Having P appearing in most points, we now show that P appears in most cubes with coefficient c_0.

Proposition 8 *For at least* $3/4$ *of the cubes* $\mathcal{C}_v \in \mathcal{S}^*$*,* $\mathcal{R}_v[P|_{\mathcal{C}_v}] = c_0$.

Proof: Let $N = \{x \in \mathbf{dom}_u \mid m_R(x)[P(x)] = c_0\}$. N has, by Proposition 7, most of the points in $\mathcal{F}^d$. According to the hitting lemma, all except $\frac{16}{|\mathcal{F}|}$ of the cubes in $\mathbf{labels}(u)$ ($\frac{32}{|\mathcal{F}|}$ of $\mathcal{S}^*$), have $\frac{1}{4}$ of their points from N.

By the Markov inequality, at most $1/10$ of the cubes in $\mathcal{S}^*$ have norm $\|\mathcal{R}_v\| \leq 10 \cdot 2s = 20s$, and thus no more than $20s$ LDFs appearing in them. Therefore $1 - 1/10 - \frac{32}{|\mathcal{F}|} > 3/4$ of the cubes in $\mathcal{S}^*$ have $\frac{1}{4}$ of their points from N, and are assigned no more than $20s$ LDFs. Denote these cubes $\mathcal{S}^*(P)$. We will show that for every cube $\mathcal{C}_v \in \mathcal{S}^*(P)$, $\mathcal{R}_v[P|_{\mathcal{C}_v}] = c_0$.

Let $\mathcal{C}_v$ be a cube in $\mathcal{S}^*(P)$. The fraction of points of $\mathcal{C}_v$ on which m_R agrees with $\mathcal{R}_v$ non-ambiguously and the point belongs to N is at least $\frac{1}{4} - \alpha - \mathrm{amb}(r, D+2, 20s) > \frac{1}{5}$ (recall $\alpha \leq 1/100$). For each such point $x \in \mathcal{C}_v$, there is an LDF Q, $Q(x) = P(x)$ with $\mathcal{R}_v[Q] = c_0$. For every such point there are no more than $20s$ candidates, hence there is at least one *LDF* Q with $\mathcal{R}_v[Q] = c_0$ that is equal to P on at least

$$\frac{1}{5} \cdot \frac{1}{20s} > \frac{r(D+2)}{|\mathcal{F}|}$$

of $\mathcal{C}_v$'s points. This LDF is therefore equal to $P|_{\mathcal{C}_v}$ (two distinct $[r, D+2]$-LDFs can agree on at most $\frac{r(D+2)}{|\mathcal{F}|}$ fraction of their domain).

We have shown that $\mathcal{R}_v[P|_{\mathcal{C}_v}] = c_0$ for all cubes $\mathcal{C}_v \in \mathcal{S}^*(P)$, which make up at least $3/4$ of the cubes in $\mathcal{S}^*$. ■

We unveiled an LDF P that appears with the same coefficient $c_0 \neq 0$ in $\mathcal{R}_v$ for at least $3/4 > 2/3 + \gamma\|\mathcal{G}_u\|$ of the good nodes v. Hence P appears with the same ($c' = c_0 + \mathcal{G}_u[P]$) coefficient in $\tilde{\mathcal{G}}_v$ for at least $2/3$ of the good nodes v. Thus, P is permissible with coefficient c', and by our definition of $\mathcal{G}_u$, $\mathcal{G}_u[P] = c'$. Thus $c_0 = 0$, a contradiction. ■

6 g-CVP is NP-hard

We begin by defining the Closest Vector Problem (CVP), and its gap version g-CVP. We then define an intermediate problem called Shortest Integer Solution (SIS), and show a reduction from g-SIS to g-CVP. We then show the simple reduction from g-$\mathcal{SSAT}$ to g-SIS and therefore to g-CVP. We restrict ourselves to l_1 norm, although the results can be easily translated to any l_p norm, $1 \leq p < \infty$.

A lattice $L = L(v_1, .., v_n)$, for linearly independent vectors $v_1, .., v_n \in R^k$ is the set of all integral linear combinations of $v_1, .., v_n$, $L = \{\sum a_i v_i \mid a_i \in \mathbb{Z}\}$.
The closest-vector problem is defined as follows:

CVP. Given (L, y) where $L = L(v_1, .., v_n)$ is a lattice and $y \in R^k$, find a lattice vector closest to y (i.e. a lattice vector $v \in L$ that minimizes $\|v - y\|$.

Approximating CVP to within factor $g = g(n)$ means finding a lattice vector v whose distance from y, $\|v - y\|$, is no more than g times the minimal distance. The gap version of CVP is a decision problem as follows,

g-CVP. Given (L, y, d) for a lattice L, a vector $y \in R^k$, and a number d, distinguish between the following two cases:

Yes: There exists a lattice vector $v \in L$ for which $\|v - y\| \leq d$.

No: For every lattice vector $v \in L$, $\|v - y\| > g \cdot d$.

Proving that g-CVP is NP-hard means that having an approximation algorithm to within factor g would imply $P = NP$.

6.1 Shortest Integer Solution - SIS

Definition of SIS and g-SIS

We define a variant of CVP named Shortest Integer Solution (SIS) and its gap version, g-SIS. We then show a simple reduction from g-SIS to g-CVP.

SIS: Given (B, t) for an integer matrix B with columns $b_1, .., b_n$ and a target vector $t \in L(b_1, ..., b_n)$, such that there exists $(a_1, .., a_n)$ with $\sum_{i=1}^{n} a_i b_i = t$, find such a vector $(a_1, .., a_n)$ that minimizes $\sum |a_i|$. In other words, find the shortest integer solution for the linear system $B \cdot x = t$.

The gap version of SIS is as follows,

g-SIS: Given (B, t, d) with B and t as before, and a number d, distinguish between the following two cases:

Yes: The shortest integer solution has length d or less.

No: The shortest integer solution has length $> g \cdot d$.

Reducing g-SIS to g-CVP

Given an instance of g-SIS, (B, t, d), we efficiently construct a lattice L and a target vector y such that 'yes' instances of g-SIS are translated into 'yes' instances of g-CVP and 'no' instances are translated into 'no' instances. The lattice L is constructed by multiplying the matrix B by a very large number w, and adding a distinct 1-coordinate to each column. The vector y (that we are to approximate from within the lattice) will be t multiplied by w with zeros in the n additional coordinates:

$$L = \begin{pmatrix} & wB & \\ & & \\ 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \qquad y = \begin{pmatrix} \vdots \\ wt \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

To see that 'yes' instances map into 'yes' instances just note that any solution a to the system $Bx = t$, gives a lattice vector $L \cdot a$ such that $\|L \cdot a - y\| = \|a\|$. Let w be such that the entries in the upper half of the matrix are all integer multiples of $g \cdot d + 1$. The next lemma will show that 'no' instances of g-SIS (where the shortest solution has length $> g \cdot d$) map into 'no' instances of g-CVP.

Lemma 10 *If there is a lattice vector, $L \cdot a$, such that $r \stackrel{def}{=} \|L \cdot a - y\| \leq g \cdot d$, then there is an integer solution to (B, t) of length r.*

Proof: $r \leq gd$ means that $L \cdot a = y$ in all but the lower n coordinates, otherwise the distance r would be at least $g \cdot d + 1$. In other words, a is a solution to the g-SIS instance. The lower n coordinates of $L \cdot a$ are exactly equal to a, and therefore $\|a\| = r$. ■

6.2 From g-$\mathcal{SSAT}$ to g-SIS

We shall prove that g-SIS is NP hard for $g = n^{c/\log\log n}$ (for some constant $c > 0$) by reducing g-$\mathcal{SSAT}$ to it.

We begin with a g-$\mathcal{SSAT}$ test system $I = \langle \Psi = \{\psi_1, .., \psi_n\}, V = \{v_1, .., v_m\}, \{\mathcal{R}_{\psi_1}, .., \mathcal{R}_{\psi_n}\} \rangle$ where Ψ is a set of tests over variables V, and for each $\psi \in \Psi$, $\mathcal{R}_\psi$ is the set of satisfying assignments for ψ. We (efficiently) construct from it an instance of g-SIS, (B, t, d). We then show that the 'yes' instances of g-$\mathcal{SSAT}$ are mapped to 'yes' instanceof g-SIS and 'no' instances to 'no' instances.

We show that a consistent natural super-assignment to Ψ translates to a short (i.e. of l_1 norm $|\Psi|$) solution for (B, t). On the other hand we show that any solution that is shorter than $g \cdot |\Psi|$, translates to a consistent super-assignment of norm $< g$ for Ψ.

The General Construction. The matrix B will have a column for every pair of test $\psi \in \Psi$ and a satisfying assignment $r \in \mathcal{R}_\psi$ for it. The upper rows of B will take care of consistency, and the lower rows will take care of non-triviality.

Non-Triviality Rows. There will be a row designated to each test. In the row of ψ all of the columns associated with ψ will have a 1, and all other columns will have zero.

Consistency Rows. We shall have $|\mathcal{F}|$ rows for each pair of tests ψ_i and ψ_j and common variable x (there will be $a \cdot |\mathcal{F}|$ rows if ψ_i and ψ_j share a variables). These rows contain a consistency-ensuring gadget and only the columns associated with ψ_i and ψ_j will have non-zero values in these rows. The gadget will ensure that the super-assignments to ψ_i and ψ_j *are consistent* on their common variable x.
The **target vector** t will be an all-1 vector. We set $d \stackrel{def}{=} |\Psi|$.
We now turn to describe the structure of the gadget itself. This will complete the description of the g-SIS instance.

The Gadget. Let's concentrate on the gadget for the pair of tests ψ_i and ψ_j with common variable x. This is a pair of matrices G_1 of dimension $(|\mathcal{F}| \times |\mathcal{R}_{\psi_i}|)$ and G_2 of dimension $(|\mathcal{F}| \times \left|\mathcal{R}_{\psi_j}\right|)$. The matrices G_1 and G_2 have $|\mathcal{F}|$ rows, each corresponding to a possible assignment for the variable x. The r-th column in G_1 is the 'characteristic function' of $r|_x$, i.e. zeros everywhere except for a 1 in the $r|_x$-th coordinate. Similarly, the column in G_2 corresponding to r' is the

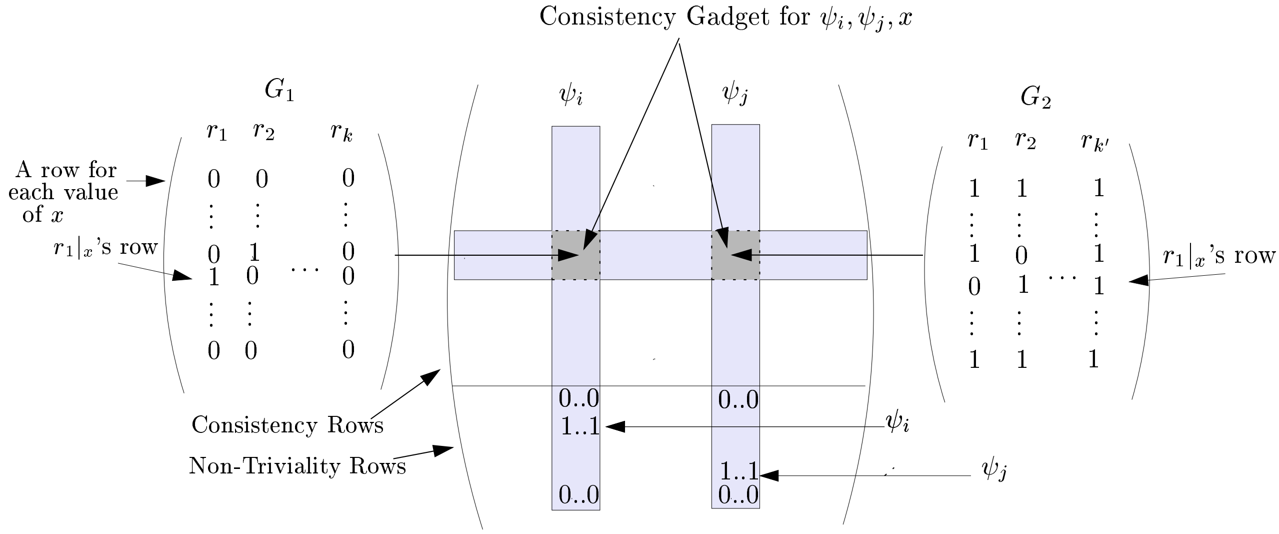

Figure 1: The SIS matrix B

negation of the characteristic function of $r'|_x$, i.e. 1 everywhere except for one 0 in the $r'|_x$-th coordinate (see Figure 1).

Proving Correctness. Let us now show that 'yes' instances of the g-$\mathcal{SSAT}$ map to 'yes' instances of the g-SIS.

Lemma 11 *If there is a consistent natural super-assignment to the g-$\mathcal{SSAT}$ test system Ψ, then there is a solution of l_1 norm $|\Psi|$ to the above g-SIS instance.*

Proof: We take the consistent natural super-assignment S and construct from it a solution to the g-SIS. We will concatenate the vectors $S(\psi_1)S(\psi_2)$... (turning n $|\mathcal{R}_{\psi_i}|$-coordinate vectors into one long vector with $\sum_i |\mathcal{R}_{\psi_i}|$ coordinates) to obtain our alleged solution to g-SIS. The target vector is reached in the non-triviality rows because S is natural i.e. it assigns a $+1$ coefficient to exactly one column of every test.

To show that the target vector is reached in the consistency rows, consider the set of $|\mathcal{F}|$ rows belonging to an arbitrary pair of tests ψ_i and ψ_j with common variable x. Suppose $S(\psi_i)[r_1], S(\psi_j)[r_2]$ are the single 1's in $S(\psi_i), S(\psi_j)$ respectively (S is natural). S is consistent so $r_1|_x = r_2|_x$. By the construction of B we see that

$$\begin{array}{c} \\ 1 \\ \vdots \\ r_1|_x \\ \vdots \\ |\mathcal{F}| \end{array}\begin{array}{c} r_1 \\ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \end{array} + \begin{array}{c} \\ 1 \\ \vdots \\ r_2|_x \\ \vdots \\ |\mathcal{F}| \end{array}\begin{array}{c} r_2 \\ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \end{array} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

and the target vector is reached in these rows.

The length of the solution is the sum of the lengths of the $S(x)$'s, and since $\|S\| = 1$, it is exactly $|\Psi|$. ■

We will now show that 'no' instances of the g-$\mathcal{SSAT}$ map to 'no' instances of the g-SIS by showing that if we ended up with an instance that isn't a 'no' instance, then we must have started with a non-'no' instance.

Lemma 12 *Let s be a solution to the above g-SIS instance, $\|s\| \le g\,|\Psi|$. There exists a non-trivial consistent super-assignment S of norm $\le g$ for the g-$\mathcal{SSAT}$ instance.*

Proof: We show how to construct S from s: we 'break' s into $|\Psi|$ pieces of length $|\mathcal{R}_\psi|$, one for each test $\psi \in \Psi$. We obtain a super-assignment S whose norm is $\frac{1}{|\Psi|}\|s\|$.

For any arbitrary $\psi \in \Psi$, the target vector is reached in the ψ-th row of the non-triviality rows. This implies that

$$\sum_{r \in \mathcal{R}_\psi} S(\psi)[r] = 1 \tag{1}$$

and in particular S is non-trivial (the sum of the coordinates in $S(\psi)$ remains the same under projection to any single variable).

Let $\psi_i, \psi_j \in \Psi$ be arbitrary tests with a common variable x. We shall show that $\pi_x(S(\psi_i)) = \pi_x(S(\psi_j))$. Consider the $|\mathcal{F}|$ rows that correspond to ψ_i, ψ_j, x. In each of these rows the sum of the vectors is 1, in other words, for any $f \in \mathcal{F}$,

$$\sum_{r\,:\,r|_x = f} S(\psi_i)[r] + \sum_{r\,:\,r|_x \neq f} S(\psi_j)[r] = 1 \tag{2}$$

Subtracting (1) for ψ_j from (2) gives,

$$\sum_{r\,:\,r|_x = f} S(\psi_i)[r] = \sum_{r\,:\,r|_x = f} S(\psi_j)[r]$$

which, by definition of the projection means $\pi_x(S(\psi_i)) = \pi_x(S(\psi_j))$. We hence have a consistent super-assignment of norm $\frac{1}{|\Psi|}\|s\| \le g$.

■

The two above lemmas complete the reduction of g-$\mathcal{SSAT}$ to g-SIS.

6.3 Other l_p norms

Our result actually holds for CVP with any l_p norm for $1 < p < \infty$, as seen by the following reduction.

Let us begin by observing that in our reduction from g-$\mathcal{SSAT}$ to CVP via g-SIS, a 'yes' instance (a test-system with a consistent natural super-assignment), was transformed to a g-SIS instance having a solution of length $|\Psi|$, which was transformed to a CVP instance $(L, y, dist = |\Psi|)$ such that there is a lattice vector $v \in L$ with $\|v - y\|_1 = |\Psi|$, and such that the vector $v - y$ is a zero-one vector, thus $\|v - y\|_p = \sqrt[p]{\|v - y\|_1} = \sqrt[p]{|\Psi|}$.

Now take the same lattice L and target vector y as a CVP_p instance with distance parameter $\sqrt[p]{|\Psi|}$, $(L, y, dist_p = \sqrt[p]{|\Psi|})$. The above observation simply says that a 'yes' instance has a solution whose distance is $\sqrt[p]{|\Psi|}$.

On the other hand, if $(L, y, |\Psi|)$ is a 'no' CVP_1 instance, then every lattice vector $v \in L$, has $\|v - y\|_1 > g \cdot |\Psi|$. Since $\|x\|_p \geq \sqrt[p]{\|x\|_1}$ for any integer-vector x, we have $\|v - y\|_p > \sqrt[p]{g \cdot |\Psi|} = \sqrt[p]{g} \cdot \sqrt[p]{|\Psi|}$.

This establishes that it is NP-hard to approximate CVP_p to within a factor of $\sqrt[p]{g} = n^{c_p/\log\log n}$ for some constant $c_p > 0$.

7 Discussion

Our result for the Closest Vector Problem was obtained via g-$\mathcal{SSAT}$ using recursive composition, that alternates between two types of algebraic encodings: the embedding extension, and the low-degree extension. This technique was adapted from the proof of a low error-probability PCP characterization of NP [DFK+99], and proved to be useful in this setting as well.

Two interesting open problems remain. The first is the Shortest Vector Problem, the homogeneous counterpart of CVP. This problem is easier to approximate than CVP, as an approximation algorithm for CVP yields an approximation for SVP [GMSS99], yet currently the best approximation algorithms for it give no better factors than those for CVP. However, where hardness results go, the SVP lags behind, with known hardness of approximation for a factor no larger than some constant.

The hardness of approximating SVP is of special interest in cryptography, where the hardness of this problem serves as the basic assumption of a crypto-system of Ajtai and Dwork, see [AD97].

The second open problem is to achieve hardness of approximation factors for CVP that are polynomial in n, say n^{ε} for some $\varepsilon > 0$. Our technique seems incapable of doing this, as the recursive structure requires super-constant depth, limiting the blow-up allowable at each level.

References

[ABSS93] S. Arora, L. Babai, J. Stern, and Z. Sweedyk. The hardness of approximate optima in lattices, codes and linear equations. In *Proc. 34th IEEE Symp. on Foundations of Computer Science*, pages 724–733, 1993.

[AD97] Miklós Ajtai and Cynthia Dwork. A public-key cryptosystem with worst-case/average-case equivalence. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, pages 284–293, El Paso, Texas, 4–6 May 1997.

[Ajt96] M. Ajtai. Generating hard instances of lattice problems. In *Proc. 28th ACM Symp. on Theory of Computing*, pages 99–108, 1996.

[Ajt98] Miklós Ajtai. The shortest vector problem in *L2* is NP-hard for randomized reductions. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing (STOC-98)*, pages 10–19, New York, May 23–26 1998. ACM Press.

[ALM+98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM*, 45(3):501–555, May 1998.

[AS98] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: a new characterization of NP. *Journal of the ACM*, 45(1):70–122, January 1998.

[Bab86] L. Babai. On Lovász's lattice reduction and the nearest lattice point problem. *Combinatorica*, 6:1–14, 1986.

[CN98] J.Y. Cai and A. Nerurkar. Approximating the SVP to within a factor $(1 + 1/\dim^{\varepsilon})$ is NP-hard under randomized reductions. In *Proc. of the 13th Annual IEEE Conference on Computational Complexity*, pages 46–55. 1998.

[Coo71] S. Cook. The complexity of theorem-proving procedures. In *Proc. 3rd ACM Symp. on Theory of Computing*, pages 151–158, 1971.

[DFK+99] I. Dinur, E. Fischer, G. Kindler, R. Raz, and S. Safra. PCP characterizations of NP: Towards a polynomially-small error-probability. In *Proc. 31st ACM Symp. on Theory of Computing*, 1999.

[GG98] O. Goldreich and S. Goldwasser. On the limits of non-approximability of lattice problems. In *Proc. 30th ACM Symp. on Theory of Computing*, pages 1–9, 1998.

[GMSS99] O. Goldreich, D. Micciancio, S. Safra, and J.-P. Seifert. Approximating shortest lattice vectors is not harder than approximating closest lattice vectors. *Inform. Process. Lett.*, 71(2):55–61, 1999.

[Lev73] L. Levin. Universal'nyĭe perebornyĭe zadachi (universal search problems : in Russian). *Problemy Peredachi Informatsii*, 9(3):265–266, 1973.

[LLL82] A.K. Lenstra, H.W. Lenstra, and L. Lovász. Factoring polynomials with rational coefficients. *Math. Ann.*, 261:513–534, 1982.

[LLS90] J. Lagarias, H.W. Lenstra, and C.P. Schnorr. Korkine-Zolotarev bases and successive minima of a lattice and its reciprocal lattice. *Combinatorica*, 10:333–348, 1990.

[Mic98] D. Micciancio. The shortest vector in a lattice is hard to approximate to within some constant. In *Proc. 39th IEEE Symp. on Foundations of Computer Science*, 1998.

[RS97] R. Raz and S. Safra. A sub-constant error-probability low-degree test, and a sub-constant error-probability PCP characterization of NP. In *Proc. 29th ACM Symp. on Theory of Computing*, pages 475–484, 1997.

[Sch85] C.P. Schnorr. A hierarchy of polynomial-time basis reduction algorithms. In *Proceedings of Conference on Algorithms, Pécs (Hungary)*, pages 375–386. North-Holland, 1985.

[vEB81] P. van Emde Boas. Another NP-complete problem and the complexity of computing short vectors in a lattice. Technical Report 81–04, Math. Inst. Univ. Amsterdam, 1981.