Spectral Method attacks Sparse LWE, Sparse LPN and Beyond

Shashwat Agrawal
Indian Institute of Technology Delhi
csz248012@iitd.ac.in

Amitabha Bagchi
Indian Institute of Technology Delhi
bagchi@cse.iitd.ac.in

Rajendra Kumar
Indian Institute of Technology Delhi
rajendra@cse.iitd.ac.in

Abstract

Given a set of k-sparse linear equations over a ring R, we give algorithms to determine whether the right-hand sides are random or have a secret assignment planted with noise. For a parameter $k/2 \leq l \leq n$, we give a spectral method to solve this problem in $\widetilde{O}\left(\binom{n}{l}|R|^l\right)$ time except with probability at most $n^{-\Omega(l)}$, provided the number of samples is roughly at least $\left(\frac{|R|n}{l}\right)^{k/2}$. This attack generalizes the Kikuchi method described by Wein et. al. (Journal of the ACM 2019) for $\mathbb{Z}_2$ to (commutative) rings of any finite size. We also give a simpler algorithm with better runtime than the spectral method and better sample complexity when $|R| = \omega(n/l)$. As a consequence, we obtain new sample-time tradeoffs for the decision problem of sparse LWE, sparse LPN over higher modulus q, and in general the distinguishing random vs planted $\mathbb{Z}_q$-linear equations for a large class of noise distributions. Our results imply a tightness of the hardness claims of Jain, Lin, Saha (Annual International Cryptology Conference, 2024) for sparse LWE.

1 Introduction

Solving systems of linear equations is one of the most fundamental problems in mathematics and computer science. Over a field, we can efficiently solve this via Gaussian elimination in polynomial time. The problem, however, becomes considerably harder when the equations are *noisy*: given m equations of the form $\langle \boldsymbol{a}_i, \boldsymbol{s} \rangle + e_i = b_i$ where e_i is an unknown error term, recovering the secret $\boldsymbol{s}$ is no longer straightforward. The difficulty of solving these noisy linear equations depends on the structure of the error distribution, and understanding precisely *when* efficient recovery is possible remains a central question in both learning theory and cryptography.

Two of the most well-studied average-case instantiations of noisy linear equations are the Learning With Errors (LWE) problem, introduced by Regev [Reg05], and its close cousin, the Learning Parity with Noise (LPN) problem [Ale03]; both serve as central hardness assumptions in modern cryptography. In LWE, one samples $(\boldsymbol{a}, b) \in \mathbb{Z}_q^n \times \mathbb{Z}_q$ by choosing $\boldsymbol{a} \leftarrow \mathbb{Z}_q^n$ uniformly, an error $e \leftarrow \Psi$ (typically a discrete Gaussian of width $r < q$), and setting $b = \langle \boldsymbol{a}, \boldsymbol{s} \rangle + e \bmod q$. The decision problem is to distinguish such samples from uniform over $\mathbb{Z}_q^n \times \mathbb{Z}_q$. The LPN problem is most commonly studied over $\mathbb{Z}_2$ with Bernoulli noise, or more generally over modulus q with sparse (low Hamming-weight) noise, and has a closely related structure. Most lattice-based cryptosystems rely on LWE's conjectured intractability against classical and quantum algorithms [Reg05; Pei09; Gen09; BV14; Pei16], while LPN's hardness rests on decades of cryptanalysis [Gol00; CM01; Fei02; MST03; AOW15a; AL16a; KMOW17a], offering competitive efficiency despite fewer constructions [Pie12; ADINZ17a].

Variants of these problems are considered to balance security with efficiency of cryptographic primitives. Different choices of modulus, noise distribution, or structural constraints (such as sparsity) can lead to primitives with faster key generation, smaller ciphertexts, or reduced computational overhead, while still relying on problems which are believed to have same hardness [LPR13; ABW10]. In this work, we focus on LWE and LPN with sparsity constraints, where the coefficient vectors $\boldsymbol{a} \in \mathbb{Z}_q^n$ are required to have only k non-zero entries for a parameter $k < n$. The motivation for this sparsity is that the samples can be computed and stored with a roughly $O(n/k)$ factor improvement in efficiency. Hardness of sparse LPN over $\mathbb{Z}_2$ and variants have been studied for as long as forty years in average-case complexity and have intimate connections to constraint satisfaction problems [Hås84; Gol11; AOW15b; AL16b; KMOW17b]. Sparse variants of LWE/LPN over large moduli q (where q is at least polynomial in the security parameter) are useful for designing efficient advanced cryptographic primitives, including FHE but remain underexplored. We summarize the current known attacks for sparse variants below.

The task in analyzing the hardness of these sparse variants is to determine the computational complexity as a function of m, i.e. the number of samples $(\boldsymbol{a}, b)$. For $m \leq n$, the random and the planted noisy instances are statistically indistinguishable. Raghavendra, Rao and Schramm [RRS17] showed that sparse LPN over $\mathbb{Z}_2$ can be solved in $n^{O(l)}$ time when m is a quantity of the form $l \left(\frac{n}{l}\right)^{k/2} \log^{c(k)}(n)$ using a degree l "Sum of squares" method for $l \geq k/2$. An improved (and the best known) bound of the form $l \left(\frac{n}{l}\right)^{k/2} \log n$ for m with the same time complexity can be obtained using the "Kikuchi method" by Wein, Alaoui, and Moore [WAM19] (independently by Hastings [Has20]).[1] Schmidhuber, O'Donnell, Kothari, and Babbush [SOKB25] gave a quantum algorithm that achieves a quartic speedup in time (i.e. $n^{l/4} \cdot \mathrm{poly}(n)$) over the Kikuchi method for certain choices of the sample complexity m, while using exponentially less space.

Despite the presence of numerous other variants of LWE, sparse LWE was not studied until

[1] The Kikuchi method is inspired from Kikuchi free energy, a marginal approximation to Gibbs free energy [Kik51]

a recent work by Jain, Lin, and Saha [JLS24]. Assuming hardness of sparse LWE, they construct secret key Linearly Homomorphic Encryption (LHE) schemes with slightly super-constant/constant overheads, and secret key homomorphic encryption for arbitrary constant-degree polynomials with slightly super-constant/constant overheads. They also perform some cryptanalysis on sparse LWE by suggesting two attacks for sparse LWE, namely the sparse vector in kernel attack and the dense minor attack. On analyzing the feasibility of these attacks for $k \geq \Omega(\log n)$ and $m = \text{poly}(n)$, it turns out that these attacks will take exponential time in n with high probability. Hence, the authors conjecture that sparse LWE with sparsity $k \geq \Omega(\log n)$ and polynomially many samples $m = \text{poly}(n)$ is as hard as the standard LWE with dimension $n' = \Theta(n)$ and polynomially many samples.

Subsequently, Bangachev et al. [BBTV25] give improved reductions from dense LWE to sparse LWE, and similarly from dense LPN to sparse LPN. Their reductions aim to establish $n^{\Omega(k)}$-hardness of sparse LWE under the assumption that dense LWE requires exponential time, but still incur an additional $1/\log k$ multiplicative factor in the exponent.

Our Contributions. In this work we provide algorithms for solving the analogue of sparse LPN over general rings R with $|R| > 2$, which we call k-LIN(R). Here the task is to determine whether a system of k-sparse linear equations over a finite ring R (having randomly chosen left-hand sides) has random right-hand sides or having a planted secret assignment with a noise distribution Ψ. We obtain analogous results of tradeoff between the sample complexity and runtime as those for sparse LPN.

Theorem 1.1. *(Informal version of Corollary 5.18) There exists an algorithm to distinguish whether a k-LIN(R) instance on n dimensions is random or planted in $\widetilde{O}\left(\binom{n}{l}(|R|-1)^l\right)$ time for some parameter $\lceil k/2 \rceil \leq l \leq n - \lfloor k/2 \rfloor$, provided the number of samples is $m \geq (|R|-1)^{\lfloor k/2 \rfloor}\left(\frac{n}{l}\right)^{\lceil k/2 \rceil} l \ln((|R|-1)n)$.*

Remark 1.1. *We note that the attack In the above theorem can achieve a speedup in runtime using quantum computation via the same algorithm as given by [SOKB25] for $R = \mathbb{Z}_2$. This speedup can be between quadratic to quartic depending on the choice of parameter l.*

We also have a simpler algorithm for the distinguishing problem with a better time complexity, and also a better sample complexity than above when $|R| = \omega(n/l)$.

Theorem 1.2. *(Informal version of Theorem 6.2) There exists an algorithm to distinguish whether a k-LIN(R) instance on n dimensions is random or planted in $\widetilde{O}\left((n/l)^{k+1} + |R|^l\right)$ time for some parameter $k \leq l \leq n$, provided the number of samples is $m \geq \left(\frac{n}{l}\right)^k l \ln((|R|-1)n)$.*

As can be seen, the parameter l in the above theorem trades sample complexity for time. In particular, we obtain tradeoffs between the sample complexity m and time complexity for sparse LWE and sparse LPN over higher modulus q. As an interesting case, if either $k = O(\log n/\log\log|R|)$ or we are allowed $m = n^{\Omega(\log\log|R|)}$ samples, we can choose l appropriately so that the time complexity is $\exp(\delta n)$ for arbitrarily small $\delta > 0$. Assuming standard LWE does not have an arbitrarily small exponential time, our result also establishes the tightness of the conjecture on the hardness of sparse LWE by [JLS24].

Moreover, the algorithms in the above theorems are applicable for a wide class of noise distributions Ψ. For the ring $\mathbb{Z}_q$, apart from the discrete Gaussian noise in LWE or the 'low Hamming weight' noise in LPN, the distinguishing problem is solvable for a large class of other plausible noise distributions. For example, one might consider a distribution where zero error is allowed only with

probability $1/q$, otherwise the error is rejection sampled again according to the discrete Gaussian. Another example can be an addition of noises, one from the discrete Gaussian and another from the LPN distribution noise, which was recently proposed in a bid for new hardness on noisy linear equations [GJLSV25]. In the sparse regime, our algorithms attack the problem for these noise distributions as well.

For Theorem 1.1 we extend the Kikuchi method used by [WAM19] for k-LIN($\mathbb{Z}_2$). This involves formulating a large graph corresponding to the k-LIN(R) instance, called the Kikuchi Graph. We determine whether the instance is random or planted based on the spectral norm of the adjacency matrix of this graph. For Theorem 1.2, we collect all subsamples of the original instance that have non-zero coefficients only on the first l variables, and then distinguish by naively checking all $|R|^l$ assignments on this sub-instance.

In [JLS24], Jain, Lin and Saha had defined sparse LWE in a manner which allows equations to be *at most* k-sparse instead of exactly k-sparse [2]. We analyse this variant for general rings as well, which we call k-ALIN(R). Intuitively, the distinguishing problem should be easier in this case, as left-hand sides here can have very few or even no non-zero entries. For large $|R|$, the above approach itself can work, however when $|R|$ is small, we can achieve a much better complexity.

Theorem 1.3. *(Informal version of Theorem 7.2) There exists an algorithm to distinguish whether a k-ALIN(R) instance on n dimensions is random or planted in $\widetilde{O}\left(|R|^k\right)$ time, provided the number of samples is $m \geq \widetilde{\Omega}\left(|R|^k\right)$.*

[JLS24] had also given a reduction from standard LWE on k dimensions to (at most) k-sparse LWE on $n \geq k$ dimensions. It is believed that distinguishing standard LWE takes $2^{\Omega(k)}$ time given arbitrary many samples. Our result above can be seen as a step towards matching this hardness result for (at most) k-sparse LWE.

Independent Work. Independent of our work, Kocurek and Manohar [KM25] have studied the problem of refuting random and semirandom instances of k-LIN(R)[3]. For this task, they define Kikuchi graphs for finite fields and abelian groups that are similar to our construction. Their sample-time tradeoffs are also similar to ours. However, our objective is different from the refutation task; ours is the task of distinguishing a random vs a planted instance. It turns out that the statistic they use for the refutation task, the *value* of an instance (maximum fraction of satisfiable constraints), cannot be extended to the distinguishing task when the noise distribution for the planted case has a high probability of introducing an error, i.e., has a low probability of making zero error. The statistic we use, which we call the *advantage* (Definition 3.5), is able to handle error distributions which satisfy a much milder condition than the one [KM25] need, and is thereby more widely applicable.

Moreover, they construct separate Kikuchi constructions depending on the sparsity k being even or odd, and also depending on the ring being a field or non-field, with significantly more involved construction in the odd sparsity case.

Our Kikuchi construction is able to mitigate both these issues; it is more amenable to the distinguishing task, and is a unified and simplified construction for all rings R and both sparsity parities when $|R| > 2$. Further, we are able to construct a simpler distinguisher which has a significantly lower runtime than that of their simple refutation algorithm when the size of the ring R is large.

[2] although in their Introduction, they mention that an equation has exactly k non-zero entries

[3] a semirandom instance is when the left-hand sides of a k-LIN(R) instance can be arbitrarily chosen, but the right-hand sides are uniformly and independently chosen

We elaborate on these aspects in Subsection 8.1.

Organization. Section 2 gives a high level overview of our spectral Kikuchi method. Section 3 introduces some basic notations, standard results, definitions of the k-LIN(R) problem and variants. Section 4 describes our Kikuchi graph construction for k-LIN(R). In Section 5, we utilize the spectral norm of the Kikuchi graph to distinguish planted vs random k-LIN(R) instances. In Section 6, we describe our simpler algorithm for distinguishing k-LIN(R) instances. In Section 7 we deal with the 'at most' k-sparse version, called k-ALIN(R), and give a simple algorithm to distinguish a planted vs random version of this problem. Finally, in Section 8 we discuss how our techniques and results compare with the works [KM25] and [JLS24].

2 Overview of the Kikuchi Method

In this section, we provide a high-level overview of our Kikuchi construction (Section 4) and the spectral method to distinguish planted vs random k-LIN(R) instances (Section 5). We first summarize how this technique was used for the k-XOR problem (k-sparse LWE/LPN for $\mathbb{Z}_2$), [WAM19] (and independently [Has20]), and then discuss the modifications we needed to make it work for general rings.

Let us review the steps of the Kikuchi method for $\mathbb{Z}_2$:

1. Each k-sparse equation is converted into a set of 2-sparse equations, i.e., equations in two variables. This is done by assigning variables to l-sized subsets of $[n]$ where $l \geq k/2$ is a parameter. For each original equation of the form $C : \sum_{j \in S} x_j = b$, we write down equations $C' : y_T + y_{T'} = b$ for all possible l-sized subsets $T, T' \subseteq [n]$ such that their symmetric difference equals S. For every vector $\boldsymbol{s} \in \mathbb{Z}_2^n$, one can define a natural vector $\boldsymbol{z} \in \mathbb{Z}_2^{\binom{[n]}{l}}$ such that $\boldsymbol{s}$ satisfies C iff $\boldsymbol{z}$ satisfies C'.

2. We note that such sets of 2-sparse equations correspond to a graph with $\binom{[n]}{l}$ as nodes and an edge between each pair of variables that appear in an equation. The weight of the edge is the value $(-1)^b$, where b is the RHS of the equation.

3. For the k-XOR problem an assignment $\boldsymbol{x} \in \mathbb{Z}_2^n$ has a function called the *advantage* (or *value*) associated with it. This advantage is defined in such a way that it is equal (up to scaling) to the quadratic form associated with the adjacency matrix of the graph defined in the previous item. Hence the optimal advantage can be directly related to the spectral norm of the adjacency matrix.

4. Distinguishing an instance as random or planted is done on the basis of spectral norm of the adjacency matrix of this graph. Appropriate tail bounds on the spectral norm are proved to show that it will be unambiguously separated in the two cases. The spectral norm of the expected adjacency matrix is 0 in the random case, and in the planted case is of the order of the average degree of the graph, which in turn is proportional to the number of samples of the k-XOR instance. Thus in order to get tail bounds which are separated in the two cases, it is essential that there are sufficiently many samples.

Now we discuss the specific modifications we made to generalize this for a ring R:

i. To successfully convert a k-sparse equation to a collection of specific 2-sparse equations, there are a couple of challenges in the $|R| > 2$ regime that we mitigate. Firstly, the LHS of an

equation here consists of a k-sparse vector $\boldsymbol{a} \in R^n$ instead of just the non-zero indices for $\mathbb{Z}_2$. Accordingly, the 2-sparse equations will have to be defined over variables indexed by the subset of R^n which has l-sparse vectors. Secondly, for defining the 2-sparse constraints in this space, we will generalize the notion of choosing pairs satisfying a symmetric difference: we choose all pairs $(\boldsymbol{u}, \boldsymbol{v})$ such that $\boldsymbol{v} - \boldsymbol{u} = \boldsymbol{a}$.

ii. We want to have $\mathbb{C}$-valued signs for each edge pair, so that we could utilize spectral method for random matrices. So the sign of $(\boldsymbol{u}, \boldsymbol{v})$ is $\chi(b)$ for the canonical character map of (the underlying additive group of) R, $\chi : R \to \mathbb{C}$. Our generalization of symmetric difference is asymmetric and yields a directed graph. To apply the spectral method, it is desirable that the adjacency matrix of the graph be *Hermitian*. We enable this by adding edges $\boldsymbol{v} \xrightarrow{\chi(-b)} \boldsymbol{u}$ for every original edge $\boldsymbol{u} \xrightarrow{\chi(b)} \boldsymbol{v}$ in the graph.

iii. We define a new notion of advantage of an assignment. For $R = \mathbb{Z}_q$ this captures the 'extent' to which the assignment violates a constraint (unlike the $\mathbb{Z}_2$ case where this can only capture whether a constraint got satisfied). We then show that it is equal (up to scaling) to the quadratic form of the (Hermitian) matrix from the previous item.

iv. Like the $\mathbb{Z}_2$ case, we can compute the spectral norm of the adjacency matrix, and prove appropriate tail bounds to ensure a separation between the spectral norms in the random vs planted instances. The method we employed for the tail bounds slightly deviates for that in the case of $\mathbb{Z}_2$, say in [SOKB25]. In the latter, the upper tail bound for the random case was obtained using Matrix Chernoff bounds for Rademacher series, and the lower tail bound for the planted case was obtained using properties of Skellam distribution. We use the *Matrix Bernstein inequality* to get both the tail bounds.

3 Preliminaries

3.1 Notations

Let $\mathbb{N}, \mathbb{R}, \mathbb{C}$ denote the set of naturals, reals, complex numbers respectively. Let $\iota = \sqrt{-1}$. For $n \in \mathbb{N}$, $[n] := \{1, 2, \ldots, n\}$. For $0 \leq r \leq n$, $\binom{[n]}{r}$ denotes the set of all r-size subsets of $[n]$. For a set S, $\mathrm{Unif}(S)$ denotes the uniform distribution on S. For $a \in \mathbb{C}$, $\mathrm{Re}(a)$ denotes its real part, a^* denotes its complex conjugate. $\omega_q := e^{\frac{2\pi\iota}{q}}$ denotes the qth primitive root of unity.

For a vector $\boldsymbol{x}$, unless stated otherwise, x_j denotes the jth coordinate of $\boldsymbol{x}$. For a matrix A, $\|A\|$ denotes its spectral norm. $A^\dagger$ denotes the conjugate transpose of A. $\boldsymbol{O}$ denotes a null matrix, and $\mathbf{0}$ denotes a null vector.

For a function $f : \mathbb{N} \to \mathbb{N}$, the notation $\widetilde{O}(f(n))$ denotes $O(f(n) \cdot \mathtt{poly}(\log f(n)))$ for some polynomial $\mathtt{poly}(\cdot)$.

In this paper, a ring R is assumed to be finite, commutative, and to have a multiplicative identity $1 \in R$. For $t \leq n$, R^n_t denotes the subset of R^n consisting of all vectors with exactly t coordinates in $R \setminus \{0\}$, and $R^n_{\leq t}$ is the subset of R^n consisting of all vectors with at most t coordinates in $R \setminus \{0\}$. For $\boldsymbol{x}, \boldsymbol{y} \in R^n$, $\langle \boldsymbol{x}, \boldsymbol{y} \rangle := x_1 y_1 + \ldots + x_n y_n$ denotes the sum of their coordinate-wise products (the ring product operation).

For a vector $\boldsymbol{x} \in R^n$, we denote $\boldsymbol{x}^{\odot l} \in R^{R^n_l}$ as the vector whose $\boldsymbol{u}^{\text{th}}$ coordinate (for $\boldsymbol{u} \in R^n_l$) is given by $x^{\odot l}_{\boldsymbol{u}} := \langle \boldsymbol{u}, \boldsymbol{x} \rangle$.

For a directed graph $\mathcal{G}$, let out-deg$_{\mathcal{G}}(T)$ denote the number of outgoing edges of a vertex T of $\mathcal{G}$, and let $D(\mathcal{G})$ denote the maximum outgoing degree of $\mathcal{G}$. We may skip writing $\mathcal{G}$ in these notations when the graph is clear in the context.

3.2 Canonical Character of an Abelian Group

For a ring R, its underlying additive group is finite and abelian, so it can be assumed to be group isomorphic to $\mathbb{Z}_{q_1} \times \ldots \times \mathbb{Z}_{q_t}$. Let $\psi : (R, +) \to \mathbb{Z}_{q_1} \times \ldots \times \mathbb{Z}_{q_t}$ be this group isomorphism, which we can find in $\widetilde{O}(|R|)$ time. Define the canonical character of R as the group homomorphism

$$\begin{aligned} \chi_R :& (R, +) \to (\mathbb{C}, \cdot) \\ & x \mapsto \omega_{q_1}^{\psi(x)_1} \cdots \omega_{q_t}^{\psi(x)_t} \end{aligned}$$

We write χ_R as simply χ when the ring is clear in the context.

Observation 3.1. $\sum_{a \in R} \chi(a) = 0$.

3.3 Concentration Inequalities

Let us state tail bound results for random variables (both scalar and matrix) which we will use in this work.

Theorem 3.1. *(Chernoff Bound) Let $X = \sum_{i=1}^{n} X_i$ be the sum of independent random variables on $[0, 1]$ with $E[X] = \mu$. Then*

$$\forall\ \delta \geq 0,\ \Pr\left[X \geq (1+\delta)\mu\right] \leq e^{-\frac{\delta^2 \mu}{2+\delta}}$$
$$\forall\ 0 \leq \delta \leq 1,\ \Pr\left[X \leq (1-\delta)\mu\right] \leq e^{-\frac{\delta^2 \mu}{2}}$$

Theorem 3.2. *(Hoeffding Inequality) Let $X_1, \ldots, X_n$ be independent and identically distributed random variables on $[a, b]$, let $\overline{X} = \frac{1}{n}\sum_{i \in [n]} X_i$, and let $\mu := \mathbb{E}[\overline{X}]$. Then*

$$\Pr\left[\overline{X} \leq \mu - t\right] < \exp\left(-\frac{2nt^2}{(b-a)^2}\right)$$
$$\Pr\left[\overline{X} \geq \mu + t\right] < \exp\left(-\frac{2nt^2}{(b-a)^2}\right)$$

The Matrix Bernstein Inequality bounds the norm of a sum of random Hermitian matrices (see, e.g., [Tro12]).

Theorem 3.3. *(Theorem 1.4 in [Tro12]) Consider a finite sequence $\{\boldsymbol{X}_i\}$ of independent, random, Hermitian matrices with dimension N. Assume that each random matrix satisfies $\mathbb{E}[\boldsymbol{X}_i] = \boldsymbol{O}$ and $\|\boldsymbol{X}_i\| \leq B$ with probability 1. Then, for all $t \geq 0$,*

$$\mathbb{P}\left[\left\|\sum_i \boldsymbol{X}_i\right\| \geq t\right] \leq N \cdot \exp\left(\frac{-t^2/2}{\sigma^2 + (Bt)/3}\right),$$

where

$$\sigma^2 := \left\|\sum_i \mathbb{E}\left[\boldsymbol{X}_i^2\right]\right\|.$$

3.4 k-LIN(R)

We now state definitions and problems related to sparse systems of linear equations over a ring.

Definition 3.2. *(exact k-sparse instance) Let R be a ring. A k-LIN(R) instance $\mathcal{I} = \{(\boldsymbol{a}_i, b_i)_{i\in[m]}\}$ consists of vectors $\boldsymbol{a}_i \in R_k^n$ and $b_i \in R$.*

Definition 3.3. *(at most k-sparse instance) Let R be a ring. A k-ALIN(R) instance $\mathcal{I} = \{(\boldsymbol{a}_i, b_i)_{i\in[m]}\}$ consists of vectors $\boldsymbol{a}_i \in R_{\leq k}^n$ and $b_i \in R$.*

For a k-LIN(R) or k-ALIN(R) instance $\mathcal{I}$, we view $(\boldsymbol{a}, b) \in \mathcal{I}$ as a R-linear homogeneous equation in n variables $x_1, \ldots, x_n$: $\sum_{j\in[n]} a_j x_j = b$.

Remark 3.4. *The exact k-sparse instance has been defined and analysed in [KM25] for abelian groups and in numerous other works for $\mathbb{Z}_2$ (e.g. [Hås84; Gol11; AOW15b; RRS17; WAM19]). On the other hand, [JLS24] defined sparse LWE as an* at most *k-sparse instance over $\mathbb{Z}_p$ for prime p. We consider and analyse both versions, and compare with their works in Section 8.*

Next, we define a notion of *advantage* of an assignment for an instance.

Definition 3.5. *Given a k-LIN(R) (or k-ALIN(R)) instance $\mathcal{I}$ and an assignment $\boldsymbol{x} \in R^n$, the advantage of assignment $\boldsymbol{x}$ for $\mathcal{I}$ is given by*

$$\mathtt{adv}_{\mathcal{I}}(\boldsymbol{x}) := \frac{1}{|\mathcal{I}|} \sum_{(\boldsymbol{a},b)\in\mathcal{I}} Re\left(\chi\left(b - \langle \boldsymbol{a}, \boldsymbol{x}\rangle\right)\right) \in [-1, 1]$$

We denote $\mathtt{adv}(\mathcal{I}) := \max_{\boldsymbol{x}\in R^n} |\mathtt{adv}_{\mathcal{I}}(\boldsymbol{x})|$.

When $R = \mathbb{Z}_2$, Definition 3.5 relates to the more popular notion of fraction of satisfied constraints by an assignment; indeed $\mathrm{adv}_{\mathcal{I}}(\boldsymbol{x})$ is simply the fraction of satisfied minus the fraction of violated constraints.

When $R = \mathbb{Z}_q$, the advantage takes into account the 'extent' to which a constraint is violated. For $q > 2$, if $\boldsymbol{x}$ violates $(\boldsymbol{a}, b)$ (i.e. $b \neq \langle \boldsymbol{a}, \boldsymbol{x}\rangle$), $b - \langle \boldsymbol{a}, \boldsymbol{x}\rangle$ can take any value in $\mathbb{Z}_q^*$. For $i \in [m]$, let $b_i = \langle \boldsymbol{a}_i, \boldsymbol{x}\rangle + e_i$ (i.e. e_i is the error term for $\boldsymbol{x}$ with respect to i^{th} constraint in the linear equation version; w.l.o.g., $-q/2 \leq e_i \leq q/2$). We can rewrite the advantage as follows:

$$\mathtt{adv}_{\mathcal{I}}(\boldsymbol{x}) = \frac{1}{m} \sum_{i\in[m]} \cos\left(\frac{2\pi\,|e_i|}{q}\right) = 1 - \frac{2}{m}\left(\sum_{i\in[m]} \sin^2\left(\frac{\pi\,|e_i|}{q}\right)\right) \tag{1}$$

The $\sin(\cdot)$ function increases in the range of $|e_i|$, so in some sense, each term in the summand measures the 'extent' to which $\boldsymbol{x}$ violates $(\boldsymbol{a}_i, b_i)$.

Now we will define random and planted versions of k-LIN(R) and k-ALIN(R). To do this for k-ALIN(R) we first define a distribution for choosing left-hand sides, generalizing the distribution described in [JLS24].

Definition 3.6. *(analogous to Definition 4.1 in [JLS24]) For a dimension $n \in \mathbb{N}$, a sparsity parameter k, a ring R, we define the distribution $D_{\mathtt{coeff},n,k,R}$ that first samples a uniformly random set $S \in \binom{[n]}{k}$. Then, the distribution samples a random vector as follows: Sample $\boldsymbol{a} \in R^n$ so that each coordinate a_j is set to 0 if $j \notin S$. Otherwise, a_j is chosen at random from R.*

Observation 3.7. *While sampling a vector from the distribution $D_{\texttt{coeff},n,k,R}$, the probability of sampling $\mathbf{0} \in R^n$ is $p_0 = 1/|R|^k$.*

Definition 3.8. *(Random instance) A k-LIN(R) (resp. k-ALIN(R)) instance $\mathcal{I} = \{(\boldsymbol{a}_i, b_i)_{i\in[m]}\}$ is said to be a random instance if $\boldsymbol{a}_i \xleftarrow{iid} Unif(R_k^n)$ (resp. $\boldsymbol{a}_i \xleftarrow{iid} D_{\texttt{coeff},n,k,R}$) and $b_i \xleftarrow{iid} Unif(R)$.*

Definition 3.9. *(Planted instance) A k-LIN(R) (resp. k-ALIN(R)) instance $\mathcal{I} = \{(\boldsymbol{a}_i, b_i)_{i\in[m]}\}$ is said to be a planted instance with respect to secret vector $\boldsymbol{s} \in R^n$ and noise distribution Ψ on R if $\boldsymbol{a}_i \xleftarrow{iid} Unif(R_k^n)$ (resp. $\boldsymbol{a}_i \xleftarrow{iid} D_{\texttt{coeff},n,k,R}$) and for each $i \in [m]$, $b_i = \langle \boldsymbol{a_i}, \boldsymbol{s}\rangle + e_i$ where $e_i \xleftarrow{iid} \Psi$.*

3.4.1 Sparse LWE and Sparse LPN.

We first recall the discrete Gaussian distribution on the integers and $\mathbb{Z}_q$.

Definition 3.10. *The discrete Gaussian distribution on the integers with width r, denoted $\Psi_{\mathbb{Z},r}$, is the distribution having probability mass at $x \in \mathbb{Z}$ proportional to $\exp\left(-\pi x^2/r^2\right)$.*

Definition 3.11. *The discrete Gaussian distribution on $\mathbb{Z}_q$ with width r, denoted $\Psi_{\mathbb{Z}_q,r}$ is the distribution having probability mass at $j \in \mathbb{Z}_q$ proportional to $\sum_{x\in q\mathbb{Z}+j} e^{-\pi x^2/r^2}$.*

Let us state a couple of observations related to the discrete Gaussian, which shall be useful later.

Observation 3.12. *For $1 < r < q$, we have $\Psi_{\mathbb{Z}_q,r}(0) = \Theta(1/r)$ (see Appendix A.1).*

Observation 3.13. *Let $\chi : \mathbb{Z}_q \to \mathbb{C}$ be the canonical character of $\mathbb{Z}_q$. For $1 < r < q$, we have $\mathbb{E}[\chi(\Psi_{\mathbb{Z}_q,r})] \geq e^{-\pi r^2/q^2}$ (see Appendix A.2).*

We now recall the k-sparse LWE problem from [JLS24].

Definition 3.14. *(Definition 4.3 in [JLS24]) For an integer dimension n, sparsity parameter k, sample complexity m, modulus q, and noise parameter r, we define the k-LWE$_{n,m,q,r}$ (resp. k-ALWE$_{n,m,q,r}$) problem as a distinguishing problem that is given m samples from either of two k-LIN($\mathbb{Z}_q$) distributions, and is required to guess the distribution. The first distribution is a planted k-LIN($\mathbb{Z}_q$) (resp. k-ALIN($\mathbb{Z}_q$)) distribution with noise distribution $\Psi_{\mathbb{Z}_q,r}$ and the second distribution is a random k-LIN($\mathbb{Z}_q$) (resp. k-ALIN($\mathbb{Z}_q$)) distribution.*

The standard decisional LWE problem is identical to k-ALWE$_{n,m,q,r}$ problem for $k = n$.

We also define a sparse version of the Learning Parity with Noise(LPN) problem for higher modulus, which (the non-sparse version) was defined in [ADINZ17b].

Definition 3.15. *For an integer dimension n, sparsity parameter k, sample complexity m, modulus q, and noise parameter μ, we define the k-LPN$_{n,m,q,\mu}$ (resp. k-ALPN$_{n,m,q,\mu}$) problem as a distinguishing problem that is given m samples from either of two k-LIN($\mathbb{Z}_q$) distributions, and is required to guess the distribution. The first distribution is a planted k-LIN($\mathbb{Z}_q$) (resp. k-ALIN($\mathbb{Z}_q$)) distribution with noise distribution $\zeta_{\mathbb{Z}_q,\mu}$ on $\mathbb{Z}_q$ where $\zeta_{\mathbb{Z}_q,\mu}(0) = 1-\mu$, $\forall j \in \mathbb{Z}_q \setminus \{0\}$ $\zeta_{\mathbb{Z}_q,\mu}(j) = \mu/(q-1)$. The second distribution is a random k-LIN($\mathbb{Z}_q$) (resp. k-ALIN($\mathbb{Z}_q$)) distribution.*

Let us compute $\mathbb{E}[\chi(\zeta_{\mathbb{Z}_q,\mu})]$ which will be useful for us.

Observation 3.16. *Let $\chi : \mathbb{Z}_q \to \mathbb{C}$ be the canonical character of $\mathbb{Z}_q$. Then $\mathbb{E}[\chi(\zeta_{\mathbb{Z}_q,\mu})] = 1 - \frac{q\mu}{(q-1)}$.*

Proof.

$$\mathbb{E}[\chi(\zeta_{\mathbb{Z}_q,\mu})] = \sum_{j\in\mathbb{Z}_q} \zeta_{\mathbb{Z}_q,\mu}(j)\chi(j) \;= 1-\mu+\frac{\mu}{q-1}\sum_{j\in\mathbb{Z}_q}\chi(j) \;= 1-\mu-\frac{\mu}{(q-1)}$$

□

4 A Kikuchi Graph Construction

In this section, our objective is to distinguish a random instance of k-LIN(R) from a planted such instance.

We will develop the insight of viewing a k-sparse R-linear equation as the edges of a graph in a large vertex space. Given a k-LIN(R) instance $\mathcal{I}$, we wish to produce a directed graph, namely $\mathcal{K}_l(\mathcal{I})$ (for some parameter l). This technique was introduced for the k-XOR problem by Wein. et. al. [WAM19] (and implicitly by Hastings [Has20]). Having a graph representative of the instance allows us to use spectral properties of the graph for the distinguishing problem.

For a head start, here is some intuition on how this is done: Let us consider a transformed space whose dimensions are indexed by $R^n_{k/2}$ (assume for simplicity for now that k is even). Then an equation over indeterminates $\{x_i\}_{i\in[n]}$

$$C : a_{i_1}x_{i_1} + \ldots + a_{i_k}x_{i_k} = b_i$$

can be converted to an equation over indeterminates $\{y_{\boldsymbol{u}}\}_{\boldsymbol{u}\in R^n_{k/2}}$

$$C' : y_{\boldsymbol{v}} - y_{\boldsymbol{u}} = b_i$$

where $\boldsymbol{u}, \boldsymbol{v}$ satisfy

$$u_j = \begin{cases} -a_j & \text{for } j = i_1, \ldots, i_{k/2} \\ 0 & \text{otherwise} \end{cases} \qquad v_j = \begin{cases} a_j & \text{for } j = i_{k/2+1}, \ldots, i_k \\ 0 & \text{otherwise} \end{cases}$$

The rationale behind this conversion is this: For $\boldsymbol{s} \in R^n$, consider $\boldsymbol{s}^{\odot k/2} \in R^{R^n_{k/2}}$. As mentioned in Section 3, this is defined as follows: $s^{\odot k/2}_{\boldsymbol{u}} := \sum_{j\in[n]} u_j s_j$. Then $\boldsymbol{s}$ satisfies equation C iff $\boldsymbol{s}^{\odot k/2}$ satisfies C'.

Now we can define a graph over the vertex set $R^n_{k/2}$, where C' can be viewed as an edge $\boldsymbol{u} \to \boldsymbol{v}$ with a sign that can be a function of b_i. We will define this sign to be $\chi(b_i)$; this will make the graph's adjacency matrix complex valued and will enable us to utilize the well-studied spectral theory of Hermitian matrices.

The above choice of non-zero indices for $\boldsymbol{u}, \boldsymbol{v}$ is only one of $\binom{k}{k/2}$ possible choices; we can choose any of the $k/2$ indices for $\boldsymbol{u}$. Furthermore, note that the following generalizations can also be made:

- We need not restrict ourselves to $(k/2)$ non-zero indices, we can have l non-zero indices in $\boldsymbol{u}, \boldsymbol{v}$ for an appropriate $l \geq k/2$.
- We also have some freedom to choose the values u_j, v_j, for instance we could have defined $u_{i_1} = 1 - a_{i_1}, v_{i_1} = 1$ and still maintained that $\boldsymbol{s}$ satisfies C iff $\boldsymbol{s}^{\odot k/2}$ satisfies C'.

Construction 4.1. *Let $\mathcal{I}$ be a k-LIN(R) instance, and $\lceil k/2 \rceil \leq l \leq n$ be a parameter. $\mathcal{K}_l(\mathcal{I})$ is a directed graph over vertices indexed by R^n_l. For every constraint $(\boldsymbol{a}, b) \in \mathcal{I}$, $\mathcal{K}_l(\mathcal{I})$ has edges $\boldsymbol{u} \xrightarrow{\chi(b)} \boldsymbol{v}$ and $\boldsymbol{v} \xrightarrow{\chi(-b)} \boldsymbol{u}$ for all $\boldsymbol{u}, \boldsymbol{v} \in R^n_l$ which satisfy $\boldsymbol{v} - \boldsymbol{u} = \boldsymbol{a}$.*

Remark 4.2. *Our Kikuchi construction differs from that of [KM25] (their case of k even, R field) in two aspects. Firstly, they had edges $\boldsymbol{u} \xrightarrow{\chi(\beta b)} \boldsymbol{v}$ for $\boldsymbol{v} - \boldsymbol{u} = \beta \boldsymbol{a}$ for all $\beta \in R \setminus \{0\}$. Secondly, they imposed an additional restriction on these edges that the symmetric difference of supports of $\boldsymbol{u}$ and $\boldsymbol{v}$ should be equal to the support of $\boldsymbol{a}$. When we compare our results with theirs in Subsection 8.1, we elaborate on how differing in the first aspect allows us to distinguish random instances from planted ones for a large class of noise distributions, and on how differing in the second aspect allows us to have a unified and simplified construction for odd k and R being any ring when $|R| > 2$.*

For a constraint $(\boldsymbol{a}, b) \in \mathcal{I}$, how many corresponding edges does $\mathcal{K}_l(\mathcal{I})$ have? Before venturing to compute this, let us rewrite Construction 4.1 in a way that is more amenable to computing this.

Construction 4.3. *(Equivalent to Construction 4.1) For every constraint $(\boldsymbol{a}, b) \in \mathcal{I}$, $\mathcal{K}_l(\mathcal{I})$ has all edges $(\boldsymbol{u}, \boldsymbol{v}, \chi(b))$, $(\boldsymbol{v}, \boldsymbol{u}, \chi(-b))$ which satisfy the following:*

1. *For every $j \in [n]$ such that $a_j = 0$, exactly one of the following holds:*
   - *(a) $u_j = v_j = 0$.*
   - *(b) $u_j = v_j \neq 0$.*
2. *For every $j \in [n]$ such that $a_j \in R \setminus \{0\}$, exactly one of the following holds:*
   - *(a) $u_j = -a_j, v_j = 0$.*
   - *(b) $u_j = 0, v_j = a_j$.*
   - *(c) $u_j, v_j \in R \setminus \{0\}$ such that $v_j - u_j = a_j$.*

We will now compute $\theta^{(n,l,k)}$ which is the number of edges added in $\mathcal{K}_l(\mathcal{I})$ corresponding to the constraint $(\boldsymbol{a}, b) \in \mathcal{I}$.

Let $r_{1a}, r_{1b}, r_{2a}, r_{2b}, r_{2c}$ denote the number of indices on which conditions 1(a), 1(b), 2(a), 2(b), 2(c) are applied respectively. It can be seen that the following must be satisfied:

$$r_{1a} + r_{1b} + r_{2a} + r_{2b} + r_{2c} = n \quad \text{(total indices)} \tag{2}$$
$$r_{1b} + r_{2a} + r_{2c} = l \quad \text{(non-zero entries of } \boldsymbol{u}\text{)} \tag{3}$$
$$r_{1b} + r_{2b} + r_{2c} = l \quad \text{(non-zero entries of } \boldsymbol{v}\text{)} \tag{4}$$
$$r_{2a} + r_{2b} + r_{2c} = k \quad \text{(non-zero entries of } \boldsymbol{a}\text{)} \tag{5}$$

If $r_{2c} = r \in \{0, 1, \ldots, k\}$, we obtain $r_{2a} = r_{2b} = (k - r)/2$ from (3), (4), (5), $r_{1b} = l - (k + r)/2$ from (3), $r_{1a} = n - l - (k - r)/2$ from (2). For indices j of condition 1(b), we have $(|R| - 1) = |R| - 1$ choices of (u_j, v_j), and for indices j of condition 2(c), we have $|R| - 2$ choices of (u_j, v_j).

Thus we get the number of constraints for each original constraint as: [4]

$$\theta^{(n,l,k)} = 2 \sum_{r=0}^{t} \binom{k}{r} (|R| - 2)^r \binom{k - r}{(k - r)/2} \binom{n - k}{l - (k + r)/2} (|R| - 1)^{l - (k+r)/2} \tag{6}$$

Example 4.1. *Let us consider a couple of toy examples to get a better idea on the construction. Let $R = \mathbb{Z}_5$, dimension $n = 4$ and $l = 3$.*

1. *Consider $\boldsymbol{a} = (1, 2, 0, 0) \in R_k^n$ for $k = 2$. Two vectors $\boldsymbol{u}, \boldsymbol{v} \in R_l^n$ corresponding to $\boldsymbol{a}$ (satisfying $\boldsymbol{v} - \boldsymbol{u} = \boldsymbol{a}$) are:*
   - *(a) $\boldsymbol{u} = (-1, 0, 3, 4), \boldsymbol{v} = (0, 2, 3, 4)$. This corresponds to the case $r_{2c} = 0$.*
   - *(b) $\boldsymbol{u} = (2, 2, 1, 0), \boldsymbol{v} = (3, 4, 1, 0)$. This corresponds to the case $r_{2c} = 2$.*
2. *Consider $\boldsymbol{a} = (1, 2, 3, 0) \in R_k^n$ for $k = 3$. Two vectors $\boldsymbol{u}, \boldsymbol{v} \in R_l^n$ corresponding to $\boldsymbol{a}$ (satisfying $\boldsymbol{v} - \boldsymbol{u} = \boldsymbol{a}$) are:*
   - *(a) $\boldsymbol{u} = (2, 0, -3, 4), \boldsymbol{v} = (3, 2, 0, 4)$. This corresponds to the case $r_{2c} = 1$.*

[4] in this equation we use the convention that if s is not a non-negative integer or $s > t$, then $\binom{t}{s} = 0$

(b) $\boldsymbol{u} = (1,1,1,0), \boldsymbol{v} = (2,3,4,0)$. *This corresponds to the case* $r_{2c} = 3$.

Note that in (6), only the terms where r has same parity as k are positive.

We can obtain two lower bounds for $\theta^{(n,l,k)}$. One of them will come from the summand of $r = 0$ or $r = 1$, depending on whether k is even or odd, respectively.

$$\text{Assuming } \lceil k/2 \rceil \le l \le n - \lfloor k/2 \rfloor, \quad \theta^{(n,l,k)} \ge \frac{2^{k-1}}{(k+1)} \binom{n-k}{l-\lceil k/2 \rceil} (|R|-1)^{l-\lfloor k/2 \rfloor} \tag{7}$$

The other bound will come from the $r = k$ summand term.

$$\text{Assuming } k \le l \le n, \quad \theta^{(n,l,k)} \ge 2^{1-k} \binom{n-k}{l-k} (|R|-1)^{l} \tag{8}$$

Depending on whether $|R|$ is small or large, we will use these bounds for our analysis.

5 Distinguishing by spectral method over Kikuchi Graphs

The Kikuchi Graph $\mathcal{K}_l(\mathcal{I})$ can be naturally associated with its adjacency matrix (for simplicity, we use the same notation to denote the adjacency matrix)[5]. By Construction 4.1, for every edge $\boldsymbol{u} \xrightarrow{\chi(b)} \boldsymbol{v}$, we have the edge $\boldsymbol{v} \xrightarrow{\chi(-b)} \boldsymbol{u}$ present in the graph. Therefore $\mathcal{K}_l(\mathcal{I})$ is Hermitian.

Our certificate to determine whether $\mathcal{I}$ is planted or random will be the *spectral norm* of $\mathcal{K}_l(\mathcal{I})$. To give some intuition for this, let us relate the optimal advantage of $\mathcal{I}$ with the spectral norm of $\mathcal{K}_l(\mathcal{I})$.

Proposition 5.1. *Let* $\mathcal{I} = \{(\boldsymbol{a}_i, b_i)\}_{i\in[m]}$ *be a* k*-LIN*(R) *instance. For every assignment* $\boldsymbol{x} \in R^n$*, define* $\boldsymbol{y} \in \mathbb{C}^{R_l^n}$ *as* $y_{\boldsymbol{u}} := \chi\left(-\sum_{j\in[n]} u_j x_j\right)$*. Then* $adv_{\mathcal{I}}(\boldsymbol{x}) = \frac{1}{m\theta^{(n,l,k)}} \boldsymbol{y}^\dagger \mathcal{K}_l(\mathcal{I}) \boldsymbol{y}$*. Further,* $\mathtt{adv}(\mathcal{I}) \le \frac{N}{m\theta^{(n,l,k)}} \|\mathcal{K}_l(\mathcal{I})\|$ *for* $N = \binom{n}{l}(|R|-1)^l$.

[5] if there are parallel edges, we add their signs in the corresponding matrix entry

Proof.

$$
\begin{aligned}
\boldsymbol{y}^\dagger \mathcal{K}_l(\mathcal{I})\boldsymbol{y} &= \sum_{\boldsymbol{u},\boldsymbol{v}\in R_l^n} y^*_{\boldsymbol{u}} \cdot \left(\mathcal{K}_l(\mathcal{I})\right)_{\boldsymbol{u}\boldsymbol{v}} \cdot y_{\boldsymbol{v}} \\
&= \sum_{i\in[m]} \sum_{\substack{\boldsymbol{v}-\boldsymbol{u}=\boldsymbol{a}_i \\ \boldsymbol{u}-\boldsymbol{v}=\boldsymbol{a}_i}} y^*_{\boldsymbol{u}} \cdot \left(\mathcal{K}_l(\mathcal{I})\right)_{\boldsymbol{u}\boldsymbol{v}} \cdot y_{\boldsymbol{v}} \\
&= \sum_{i\in[m]} \sum_{\boldsymbol{v}-\boldsymbol{u}=\boldsymbol{a}_i} \chi\left(\sum_{j\in[n]} u_j x_j\right) \cdot \chi\left(b_i\right) \cdot \chi\left(-\sum_{j\in[n]} v_j x_j\right) + \\
&\quad \sum_{i\in[m]} \sum_{\boldsymbol{u}-\boldsymbol{v}=\boldsymbol{a}_i} \chi\left(\sum_{j\in[n]} u_j x_j\right) \cdot \chi\left(-b_i\right) \cdot \chi\left(-\sum_{j\in[n]} v_j x_j\right) \\
&= \sum_{i\in[m]} \left(\sum_{\boldsymbol{v}-\boldsymbol{u}=\boldsymbol{a}_i} \chi\left(b_i - \langle \boldsymbol{v}-\boldsymbol{u}, \boldsymbol{x}\rangle\right) + \sum_{\boldsymbol{u}-\boldsymbol{v}=\boldsymbol{a}_i} \chi\left(-b_i + \langle \boldsymbol{u}-\boldsymbol{v}, \boldsymbol{x}\rangle\right) \right) \\
&= \sum_{i\in[m]} \frac{\theta^{(n,l,k)}}{2} \cdot \left(\chi\left(b_i - \langle \boldsymbol{a}_i, \boldsymbol{x}\rangle\right) + \chi\left(-b_i + \langle \boldsymbol{a}_i, \boldsymbol{x}\rangle\right)\right) \quad \text{(each } (\boldsymbol{a}_i, b_i) \text{ had } \theta^{(n,l,k)}/2 \text{ edge pairs)} \\
&= \theta^{(n,l,k)} \sum_{i\in[m]} \mathrm{Re}\left(\chi\left(b_i - \langle \boldsymbol{a}_i, \boldsymbol{x}\rangle\right)\right) \\
&= m\theta^{(n,l,k)} \mathtt{adv}_{\mathcal{I}}(\boldsymbol{x})
\end{aligned}
$$

Further, note that for every $\boldsymbol{y} \in \mathbb{C}^{R_l^n}$ with $|y_{\boldsymbol{u}}| = 1 \ \forall \ \boldsymbol{u} \in R_l^n$ we have $\|\boldsymbol{y}\| = \sqrt{N}$ and thus

$$
\left|\boldsymbol{y}^\dagger \mathcal{K}_l(\mathcal{I})\boldsymbol{y}\right| = N \cdot \left| \left(\frac{\boldsymbol{y}^\dagger}{\|\boldsymbol{y}^\dagger\|}\right) \mathcal{K}_l(\mathcal{I}) \left(\frac{\boldsymbol{y}}{\|\boldsymbol{y}\|}\right) \right| \leq N \left\|\mathcal{K}_l(\mathcal{I})\right\|
$$

and the second part follows. □

Proposition 5.1 gives some evidence that estimating the spectral norm of $\|\mathcal{K}_l(\mathcal{I})\|$ may be helpful for distinguishing a random vs planted instance. One consequence of this proposition is that we can hope $\|\mathcal{K}_l(\mathcal{I})\|$ to be high when $\mathcal{I}$ is planted. We will show that this is indeed the case with high probability. Moreover, to separate the random instance using this quantity, we also need to show that $\|\mathcal{K}_l(\mathcal{I})\|$ is low with high probability when $\mathcal{I}$ is random. The rest of the argument in this section is about proving tail bounds for $\|\mathcal{K}_l(\mathcal{I})\|$ in both cases.

5.1 Independent Summands of $\mathcal{K}_l(\mathcal{I})$

Firstly for a k-LIN(R) instance $\mathcal{I} = \{(\boldsymbol{a}_i, b_i)\}_{i\in[m]}$ we write $\mathcal{K}_l(\mathcal{I})$ as a sum of independent Hermitian matrices:

$$
\mathcal{K}_l(\mathcal{I}) = \sum_{i\in[m]} \boldsymbol{K}_i \tag{9}
$$

$$
\boldsymbol{K}_i(\boldsymbol{u},\boldsymbol{v}) := \begin{cases} \chi(b_i) & \text{if } \boldsymbol{v}-\boldsymbol{u} = \boldsymbol{a}_i \\ \chi(-b_i) & \text{if } \boldsymbol{u}-\boldsymbol{v} = \boldsymbol{a}_i \\ 0 & \text{otherwise} \end{cases} \tag{10}
$$

By definition, $\boldsymbol{K}_i$ are Hermitian, and independent as the constraints in $\mathcal{I}$ are independent.

We now prepare to establish some key properties of $\boldsymbol{K}_i$. Observe that the graph corresponding to $\boldsymbol{K}_i$ is a collection of disjoint *chains* (see illustration in Figure 1). This is because the graph corresponding to $\boldsymbol{K}_i$ never contains constraints $(\boldsymbol{u}, \boldsymbol{v}, \chi(b_i))$ and $(\boldsymbol{u}', \boldsymbol{v}', \chi(b_i))$ with $(\boldsymbol{u} = \boldsymbol{u}' \ \wedge \ \boldsymbol{v} \neq \boldsymbol{v}')$, or $(\boldsymbol{u} \neq \boldsymbol{u}' \ \wedge \ \boldsymbol{v} = \boldsymbol{v}')$.

$$\cdots \underset{\chi(-b_i)}{\overset{\chi(b_i)}{\rightleftarrows}} \boldsymbol{u} \underset{\chi(-b_i)}{\overset{\chi(b_i)}{\rightleftarrows}} \boldsymbol{v} \underset{\chi(-b_i)}{\overset{\chi(b_i)}{\rightleftarrows}} \cdots$$

Figure 1: A chain in the graph of $\boldsymbol{K}_i$

Observation 5.1. *The graph corresponding to $\boldsymbol{K}_i$ has maximum out-degree 2.*

Proof. $\boldsymbol{K}_i$ never contains constraints $(\boldsymbol{u}, \boldsymbol{v}, \chi(b_i))$ and $(\boldsymbol{u}', \boldsymbol{v}', \chi(b_i))$ with $(\boldsymbol{u} = \boldsymbol{u}' \ \wedge \ \boldsymbol{v} \neq \boldsymbol{v}')$, or $(\boldsymbol{u} \neq \boldsymbol{u}' \ \wedge \ \boldsymbol{v} = \boldsymbol{v}')$. So, for a vertex $\boldsymbol{u}$, there are only 4 possible edges incident to $\boldsymbol{u}$:

$$(\boldsymbol{u}, \boldsymbol{v}, \chi(b_i)), (\boldsymbol{v}, \boldsymbol{u}, \chi(-b_i)), (\boldsymbol{v}', \boldsymbol{u}, \chi(b_i)), (\boldsymbol{u}, \boldsymbol{v}', \chi(-b_i))$$

In particular, $\boldsymbol{u}$ has an out-degree of at most 2. □

Let us prove an upper bound on the spectral norm of the adjacency matrix of a signed graph. This is a generalization of the well known spectral norm bound of an unsigned graph.

Lemma 5.2. *Let $\mathcal{G}$ be a directed graph (possibly with multiple parallel edges and self loops) where each edge is of the form (I, J, β) with $\beta \in \mathbb{C}, |\beta| \leq 1$. Then $\|\mathcal{G}\| \leq D(\mathcal{G})$.*

Proof. Let $\boldsymbol{\phi}$ be an eigenvector corresponding to the eigenvalue of $\mathcal{G}$ whose absolute value is $\|\mathcal{G}\|$. Let I be a vertex in $\mathcal{G}$ such that $|\phi_I| \geq |\phi_J|$ for all vertices J of $\mathcal{G}$. Since $\boldsymbol{\phi} \neq \boldsymbol{0}$, we have $|\phi_I| \neq 0$. Now

$$\begin{aligned}\|\mathcal{G}\| = \frac{|(\mathcal{G}\phi)_I|}{|\phi_I|} &\leq \frac{\sum_{(I,J,\beta)\in E(\mathcal{G})} |\beta \cdot \phi_J|}{|\phi_I|} \\ &\leq \frac{\sum_{(I,J,\beta)\in E(\mathcal{G})} |\phi_J|}{|\phi_I|} \leq \text{out-deg}(I) \leq D(\mathcal{G})\end{aligned}$$

□

The following upper bound for $\|\boldsymbol{K}_i\|$ follows as a corollary of Observation 5.1 and Lemma 5.2.

Corollary 5.3. *Let $\mathcal{I} = \{(\boldsymbol{a}_i, b_i)\}_{i\in[m]}$ be any k-LIN(R) instance, and $\boldsymbol{K}_i$ be as defined in (10). Then $\|\boldsymbol{K}_i\| \leq 2$.*

Next, we state a result regarding sum of squares of $\{\boldsymbol{K}_i\}_{i\in[m]}$. Looking ahead, this will be useful to compute the variance statistic corresponding to $\mathcal{K}_l(\mathcal{I})$. The key observation is that $\boldsymbol{K}_i^2$ can be viewed as an adjacency matrix of a graph that captures "2-hops" of the original graph of $\boldsymbol{K}_i$.

Lemma 5.4. *Let $\mathcal{I} = \{(\boldsymbol{a}_i, b_i)\}_{i\in[m]}$ be any k-LIN(R) instance, and $\boldsymbol{K}_i$ be as defined in (10). Then $\left\|\sum_{i\in[m]} \boldsymbol{K}_i^2\right\| \leq 2D(\mathcal{K}_l(\mathcal{I}))$, $\left\|\sum_{i\in[m]} \mathbb{E}\left[\boldsymbol{K}_i^2\right]\right\| \leq 2D(\mathcal{K}_l(\mathcal{I}))$ (where the expectation is over the randomness of $\{b_i\}_i$)*[6].

[6] $D(\mathcal{K}_l(\mathcal{I}))$ does not depend on $\{b_i\}_i$ as the latter only contribute to signs of the edges

Proof. In the graph of $\boldsymbol{K}_i$, consider any edge pair

$$\dots \ \boldsymbol{u} \ \underset{\xleftarrow{\chi(-b_i)}}{\xrightarrow{\chi(b_i)}} \ \boldsymbol{v} \dots$$

which is part of a chain. By Observation 5.1, each vertex in the graph of $\boldsymbol{K}_i$ has out-degree at most 2, so the edge $(\boldsymbol{u}, \boldsymbol{v}, \chi(b_i))$ can contribute to at most 2 outgoing edges of $\boldsymbol{u}$ in the graph corresponding to $\boldsymbol{K}_i^2$ (One is $\boldsymbol{u} \xrightarrow{\chi(b_i)} \boldsymbol{v} \xrightarrow{\chi(-b_i)} \boldsymbol{u}$ and the other is possibly $\boldsymbol{u} \xrightarrow{\chi(b_i)} \boldsymbol{v} \xrightarrow{\chi(b_i)} \boldsymbol{w}$).

Let $\mathcal{G}$ be the graph corresponding to $\sum_{i\in[m]} \boldsymbol{K}_i^2$. We then have $\text{out-deg}_{\mathcal{G}}(\boldsymbol{u}) \le 2 \cdot \text{out-deg}_{\mathcal{K}_l(\mathcal{I})}(\boldsymbol{u})$ for each vertex $\boldsymbol{u}$. By applying Lemma 5.2 on $\mathcal{G}$, we obtain $\left\|\sum_{i\in[m]} \boldsymbol{K}_i^2\right\| \le 2D(\mathcal{K}_l(\mathcal{I}))$.

Further, for any edge $\boldsymbol{u} \xrightarrow{\beta} \boldsymbol{v}$ of $\boldsymbol{K}_i^2$, $|\mathbb{E}[\beta]| \le \mathbb{E}[|\beta|] = 1$ by Jensen's inequality. Thus the graph $\mathbb{E}[\mathcal{G}] = \sum_{i\in[m]} \mathbb{E}\left[\boldsymbol{K}_i^2\right]$ also satisfies the premise of Lemma 5.2 with $\text{out-deg}_{\mathbb{E}[\mathcal{G}]}(\boldsymbol{u}) = \text{out-deg}_{\mathcal{G}}(\boldsymbol{u}) \le 2 \cdot \text{out-deg}_{\mathcal{K}_l(\mathcal{I})}(\boldsymbol{u})$ for each vertex $\boldsymbol{u}$, and the result follows. □

Let us now obtain an upper bound on $D(\mathcal{K}_l(\mathcal{I}))$, for which we use the distribution of the 'left-hand sides', i.e. the fact that $\boldsymbol{a}_i \xleftarrow{\text{iid}} \text{Unif}(R_k^n)$. Recall that this is the setting in both the random and planted instances.

Lemma 5.5. *Let $\mathcal{I} = \{(\boldsymbol{a}_i, b_i)\}_{i\in[m]}$ be a k-LIN(R) instance where $\boldsymbol{a}_i \xleftarrow{iid} Unif(R_k^n)$. Let $N = \binom{n}{l}(|R|-1)^l$, and $\theta^{(n,l,k)}$ be as defined in (6). For any $\boldsymbol{u} \in R_l^n$, let $d_{\boldsymbol{u}}^{(i)}$ be the random variable denoting the number of edges (at most 2) corresponding to $(\boldsymbol{a}_i, b_i)$ outgoing from $\boldsymbol{u}$ in $\mathcal{K}_l(\mathcal{I})$. Then $\mathbb{E}\left[d_{\boldsymbol{u}}^{(i)}\right] = \frac{\theta^{(n,l,k)}}{N}$.*

Proof. $\boldsymbol{a}_i$ is equally likely to be any vector in R_k^n. By symmetry, for every $\boldsymbol{u}' \in R_l^n$, there is an equal number of vertex pairs $(\boldsymbol{u}', \boldsymbol{v})$ such that $\boldsymbol{v} - \boldsymbol{u}' \in R_k^n$. Since there are (deterministic) $\theta^{(n,l,k)}$ edges added in $\mathcal{K}_l(\mathcal{I})$ for $(\boldsymbol{a}_i, b_i)$, we get from the above two observations that $\mathbb{E}\left[d_{\boldsymbol{u}}^{(i)}\right] = \frac{\theta^{(n,l,k)}}{N}$. □

Proposition 5.6. *Let $\mathcal{I} = \{(\boldsymbol{a}_i, b_i)\}_{i\in[m]}$ be a k-LIN(R) instance where $\boldsymbol{a}_i \xleftarrow{iid} Unif(R_k^n)$. Let $N = \binom{n}{l}(|R|-1)^l$, $\Delta = \frac{m\theta^{(n,l,k)}}{N}$. Then for any $\delta > 0$, we have $D(\mathcal{K}_l(\mathcal{I})) \le (1+\delta)\Delta$ except with probability at most $N \cdot \exp\left(-\frac{\delta^2\Delta}{2(2+\delta)}\right)$.*

Proof. As a consequence of Lemma 5.5, $\text{out-deg}(\boldsymbol{u}) = \sum_{i\in[m]} d_{\boldsymbol{u}}^{(i)}$ satisfies

$$\mathbb{E}[\text{out-deg}(\boldsymbol{u})] = \frac{m\theta^{(n,l,k)}}{N} = \Delta$$

Further, $\text{out-deg}(\boldsymbol{u})$ is a sum of independent random variables $\left\{d_{\boldsymbol{u}}^{(i)}\right\}_{i\in[m]}$ supported on $\{0,1,2\}$, and
$D(\mathcal{K}_l(\mathcal{I})) = \max_{\boldsymbol{u}\in R_l^n} \text{out-deg}(\boldsymbol{u})$. Thus we can apply a Chernoff bound (Theorem 3.1) followed by a union bound to prove the proposition. □

5.2 An upper bound for the random instance

We will show that when $\mathcal{I}$ is a random k-LIN(R) instance, an upper bound on $\|\mathcal{K}_l(\mathcal{I})\|$ holds with high probability. To establish this, let us apply the Matrix Bernstein inequality to our case.

We start with analysing the summands $\boldsymbol{K}_i$'s further when $\mathcal{I}$ is a random instance.

Lemma 5.7. *Let $\mathcal{I}$ be a random k-LIN(R) instance, and $\boldsymbol{K}_i$ be as defined in (10). Then $\mathbb{E}[\boldsymbol{K}_i] = \boldsymbol{O}$.*

Proof. An entry of $\boldsymbol{K}_i$ can be $\chi(b_i), \chi(-b_i)$ or 0. Now $\mathbb{E}[\chi(b_i)] = \frac{1}{|R|}\sum_{a\in R}\chi(a) = 0$ from Observation 3.1. Similarly $\mathbb{E}[\chi(-b_i)] = 0$ and the result follows. □

With this, the antecedent of Theorem 3.3 is satisfied (taking $\boldsymbol{X}_i = \boldsymbol{K}_i$) with $\|\boldsymbol{K}_i\| \le 2$ (from Corollary 5.3) for all $i \in [m]$. Also, from Lemma 5.4, we have

$$\sigma^2 = \left\| \sum_{i\in[m]} \mathbb{E}\left[\boldsymbol{K}_i^2\right] \right\| \le 2D(\mathcal{K}_l(\mathcal{I})) \tag{11}$$

Now we choose an appropriate value of t to limit the probability in Theorem 3.3 to be an inverse polynomial of $N = |R_l^n| = \binom{n}{l}(|R|-1)^l$, say $N^{-\epsilon}$ for some choice of $\epsilon > 0$. Thus we must satisfy

$$N \cdot \exp\left(\frac{-t^2/2}{2D(\mathcal{K}_l(\mathcal{I})) + 2t/3}\right) \le N^{-\epsilon} \tag{12}$$

By simplifying (12) we have the following upper bound for the maximum eigenvalue:

Corollary 5.8. *Let $\mathcal{I} = \{(\boldsymbol{a}_i, b_i)\}_{i\in[m]}$ be a random k-LIN(R) instance of dimension n, let $N = \binom{n}{l}(|R|-1)^l$. Then for any $\epsilon > 0$, except with probability at most $N^{-\epsilon}$, we have*

$$\|\mathcal{K}_l(\mathcal{I})\| \le \frac{2\sqrt{(1+\epsilon)\ln N}}{3}\left(\sqrt{(1+\epsilon)\ln N} + \sqrt{9D(\mathcal{K}_l(\mathcal{I})) + (1+\epsilon)\ln N}\right) \tag{13}$$

When the left-hand sides $\boldsymbol{a}_i \xleftarrow{\text{iid}} \text{Unif}(R_k^n)$ (which is the case for a random instance), Proposition 5.6 gives us that $D(\mathcal{K}_l(\mathcal{I}))$ is upper bounded by $(1+\delta)\Delta$ with high probability, where $\Delta = \frac{m\theta^{(n,l,k)}}{N}$ is the average out-degree of $\mathcal{K}_l(\mathcal{I})$, and $\delta > 0$ is a constant. We can now combine Corollary 5.8 and Proposition 5.6 to get an upper bound on the spectral norm with respect to Δ. We would like to choose Δ (which in turn means choosing the number of constraints m in $\mathcal{I}$) so that we get $\|\mathcal{K}_l(\mathcal{I})\| \le \delta\Delta$. Observe that in (13) we can impose

$$\Delta \ge \frac{16\delta + 12}{3\delta^2}(1+\epsilon)\ln N$$

to get this desired bound for all $\delta > 0$. Also, the probability of failure in Proposition 5.6 is upper bounded by $N^{-\epsilon}$ for this choice.

Theorem 5.9. *Let $\mathcal{I} = \{(\boldsymbol{a}_i, b_i)\}_{i\in[m]}$ be a random k-LIN(R) instance of dimension n, let $N = \binom{n}{l}(|R|-1)^l$. Suppose $\Delta := \frac{m\theta^{(n,l,k)}}{N} \ge \frac{16\delta+12}{3\delta^2}(1+\epsilon)\ln N$ for some $\delta, \epsilon > 0$. Then we have $\|\mathcal{K}_l(\mathcal{I})\| \le \delta\Delta$ except with probability at most $2N^{-\epsilon}$.*

Finally let us restate this result in terms of the parameters of $\mathcal{I}$.

Here we will use the lower bounds of $\theta^{(n,l,k)}$ from (7), (8) depending on the value of $|R|$ and l.

If $|R| = O(n/l)$, then (7) will give a better bound than (8). On the other hand, if $|R| = \omega(n/l)$, (8) gives a better bound.

Corollary 5.10. *Let k, n, l be positive integers such that $\lceil k/2 \rceil \leq l \leq n - \lfloor k/2 \rfloor$. Let $0 < \delta < 1$. Consider a k-LIN(R) instance $\mathcal{I}$ of dimension n and m constraints, $m \geq \frac{l \ln((|R|-1)n)}{\delta^2} \cdot (|R|-1)^{\lfloor k/2 \rfloor} \left(\frac{\alpha n}{l}\right)^{\lceil k/2 \rceil}$ for an absolute constant $\alpha > 0$. Let $\Delta := \frac{m\theta^{(n,l,k)}}{\binom{n}{l}(|R|-1)^l}$. Then if $\mathcal{I}$ is random, we get $\|\mathcal{K}_l(\mathcal{I})\| \leq \delta\Delta$ except with probability at most $n^{-\Omega(l)}$.*

Corollary 5.11. *Let k, n, l be positive integers such that $k \leq l \leq n$. Let $0 < \delta < 1$. Consider a k-LIN(R) instance $\mathcal{I}$ of dimension n and m constraints, $m \geq \frac{l \ln((|R|-1)n)}{\delta^2} \cdot \left(\frac{\alpha n}{l}\right)^k$ for an absolute constant $\alpha > 0$. Let $\Delta := \frac{m\theta^{(n,l,k)}}{\binom{n}{l}(|R|-1)^l}$. Then if $\mathcal{I}$ is random, we get $\|\mathcal{K}_l(\mathcal{I})\| \leq \delta\Delta$ except with probability at most $n^{-\Omega(l)}$.*

5.3 A lower bound for the planted instance

Recall that the only difference in a planted instance from a random instance of k-LIN(R) is that the 'right-hand side' $b_i \in R$ of a constraint $(\boldsymbol{a}_i, b_i)$ is correlated with $\boldsymbol{a}_i$.

We will impose a mild assumption on the noise distribution Ψ for the planted instance: the quantity $\rho := \mathbb{E}[\chi(\Psi)] \in \mathbb{R} \setminus \{0\}$.[7]

The heart of our analysis will still be the Matrix Bernstein inequality, but with a slightly different set of matrices. This is because Lemma 5.7 ($\mathbb{E}[\boldsymbol{K}_i] = \boldsymbol{O}$) will no longer hold for the same definition of $\boldsymbol{K}_i$ when $\mathcal{I}$ is planted. Let us calculate $\mathbb{E}[\boldsymbol{K}_i]$.

Notation 5.2. *Let $\boldsymbol{a} \in R_k^n$ and $\boldsymbol{s} \in R^n$. We define a square matrix $A_{\boldsymbol{s},\boldsymbol{a}}$ indexed by R_l^n as follows:*

$$A_{\boldsymbol{s},\boldsymbol{a}}(\boldsymbol{u}, \boldsymbol{v}) := \begin{cases} \chi(\langle \boldsymbol{a}, \boldsymbol{s} \rangle) & \text{if } \boldsymbol{v} - \boldsymbol{u} = \boldsymbol{a} \\ \chi(-\langle \boldsymbol{a}, \boldsymbol{s} \rangle) & \text{if } \boldsymbol{u} - \boldsymbol{v} = \boldsymbol{a} \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

Lemma 5.12. *Let $\mathcal{I} = \{(\boldsymbol{a}_i, b_i)\}_{i \in [m]}$ be a planted k-LIN(R) instance with respect to secret vector $\boldsymbol{s} \in R^n$ and noise distribution Ψ on R with $\rho = \mathbb{E}[\chi(\Psi)] \in \mathbb{R}$. Let $\boldsymbol{K}_i$'s be the same as defined in (10). Then $\mathbb{E}[\boldsymbol{K}_i] = \rho A_{\boldsymbol{s},\boldsymbol{a}_i}$.*

Proof. This follows from the fact that $b_i = \langle \boldsymbol{a_i}, \boldsymbol{s} \rangle + e_i$, $\chi(b_i) = \chi(\langle \boldsymbol{a_i}, \boldsymbol{s} \rangle) \cdot \chi(e_i)$ and $\mathbb{E}[\chi(e_i)] = \mathbb{E}[\chi(-e_i)] = \rho$. □

Notice that $\sum_{i \in [m]} A_{\boldsymbol{s},\boldsymbol{a}_i}$ can be viewed as a signed adjacency matrix of an appropriate graph. This graph has the same edges as in $\mathcal{K}_l(\mathcal{I})$ but with different signs. Analogous to the spectral norm of an adjacency matrix being lower bounded by the average degree of the graph, we will set a lower bound for the spectral norm of $\sum_{i \in [m]} A_{\boldsymbol{s},\boldsymbol{a}_i}$ which will be useful later.

Proposition 5.13. *Let $\mathcal{I} = \{(\boldsymbol{a}_i, b_i)\}_{i \in [m]}$ be a planted k-LIN(R) instance with respect to secret vector $\boldsymbol{s} \in R^n$ and noise distribution Ψ on R. Let $N := |R_l^n| = \binom{n}{l}(|R|-1)^l$, $\Delta := \frac{m\theta^{(n,l,k)}}{N}$. Then for any $\delta > 0$ we have $\left\| \sum_{i \in [m]} A_{\boldsymbol{s},\boldsymbol{a}_i} \right\| \geq \Delta$.*

Proof.

$$\left\| \sum_{i \in [m]} A_{\boldsymbol{s},\boldsymbol{a}_i} \right\| = \sup_{\boldsymbol{x} \in \mathbb{C}^{R_l^n} \setminus \{\boldsymbol{0}\}} \frac{\boldsymbol{x}^\dagger \left(\sum_{i \in [m]} A_{\boldsymbol{s},\boldsymbol{a}_i} \right) \boldsymbol{x}}{\boldsymbol{x}^\dagger \boldsymbol{x}} \tag{15}$$

[7] the noise for sparse LWE, sparse LPN and other nice distributions shall satisfy this assumption.

Let us consider the quadratic form in (15) for a specific $\boldsymbol{x}$. Define $\boldsymbol{z} \in \mathbb{C}^{R_l^n}$ as follows: For $\boldsymbol{u} \in R_l^n$, $z_{\boldsymbol{u}} := \chi(-\langle \boldsymbol{u}, \boldsymbol{s} \rangle)$. It can be seen that for $\boldsymbol{u}, \boldsymbol{v} \in R_l^n$ satisfying $\boldsymbol{v} - \boldsymbol{u} = \boldsymbol{a}$, we have $(z_{\boldsymbol{u}})^{-1} z_{\boldsymbol{v}} = \chi(-\langle \boldsymbol{a}, \boldsymbol{s} \rangle)$.

With this, we have

$$\begin{aligned}
\left\| \sum_{i \in [m]} A_{\boldsymbol{s}, \boldsymbol{a}_i} \right\| &\geq \frac{\left(\sum_{i \in [m]} \boldsymbol{z}^\dagger \cdot A_{\boldsymbol{s}, \boldsymbol{a}_i} \cdot \boldsymbol{z} \right)}{\boldsymbol{z}^\dagger \boldsymbol{z}} \\
&= \frac{\sum_{i \in [m]} \sum_{(\boldsymbol{u}, \boldsymbol{v})} z_{\boldsymbol{u}}^{-1} \cdot A_{\boldsymbol{s}, \boldsymbol{a}_i}(\boldsymbol{u}, \boldsymbol{v}) \cdot z_{\boldsymbol{v}}}{N} \\
&= \frac{\sum_{i \in [m]} (\#\text{ of non-zero entries in } A_{\boldsymbol{s}, \boldsymbol{a}_i})}{N} \\
&= \Delta
\end{aligned}$$

where the last inequality holds because the graph of $\sum_{i \in [m]} A_{\boldsymbol{s}, \boldsymbol{a}_i}$ has the same edges as in $\mathcal{K}_l(\mathcal{I})$. □

Let us now define $\widetilde{\boldsymbol{K}_i} := \boldsymbol{K}_i - \mathbb{E}[\boldsymbol{K}_i] = \boldsymbol{K}_i - \rho A_{\boldsymbol{s}, \boldsymbol{a}_i}$. By definition, $\mathbb{E}[\widetilde{\boldsymbol{K}_i}] = \boldsymbol{O}$. Let us also obtain a bound on its spectral norm.

Lemma 5.14. *For $\widetilde{\boldsymbol{K}_i}$ defined as above, $\left\| \widetilde{\boldsymbol{K}_i} \right\| \leq 2(1 + |\rho|)$.*

Proof. $\left\| \widetilde{\boldsymbol{K}_i} \right\| \leq \|\boldsymbol{K}_i\| + |\rho| \, \|A_{\boldsymbol{s}, \boldsymbol{a}_i}\| \leq 2(1 + |\rho|)$ where the second inequality can be obtained by an analogous argument for $A_{\boldsymbol{s}, \boldsymbol{a}_i}$ as for $\boldsymbol{K}_i$ in Corollary 5.3. □

To apply the Matrix Bernstein inequality, we are left to compute

$$\sigma^2 := \left\| \sum_{i \in [m]} \mathbb{E}\left(\widetilde{\boldsymbol{K}_i}^2 \right) \right\|$$

$$\sigma^2 = \left\| \sum_{i \in [m]} \mathbb{E}\left((\boldsymbol{K}_i - \rho A_{\boldsymbol{s}, \boldsymbol{a}_i})^2 \right) \right\| = \left\| \sum_{i \in [m]} \left(\mathbb{E}\left(\boldsymbol{K}_i^2 \right) - \rho^2 A_{\boldsymbol{s}, \boldsymbol{a}_i}^2 \right) \right\| \tag{16}$$

$$\leq \left\| \sum_{i \in [m]} \mathbb{E}\left(\boldsymbol{K}_i^2 \right) \right\| + |\rho|^2 \left\| \sum_{i \in [m]} A_{\boldsymbol{s}, \boldsymbol{a}_i}^2 \right\| \tag{17}$$

$$\leq 2(1 + |\rho|^2) D(\mathcal{K}_l(\mathcal{I})) \quad \text{(using Lemma 5.4 and same idea for } A_{\boldsymbol{s}, \boldsymbol{a}_i}^2\text{)} \tag{18}$$

Like the random case, we need to choose an appropriate value of t to limit the probability in Theorem 3.3 to be essentially an inverse polynomial of $N = |R_l^n| = \binom{n}{l}(|R| - 1)^l$, say $N^{-\epsilon}$ for some choice of $\epsilon > 0$. Thus we must satisfy

$$N \cdot \exp\left(\frac{-t^2/2}{2\left(1 + |\rho|^2\right) D(\mathcal{K}_l(\mathcal{I})) + 2(1 + |\rho|)t/3} \right) \leq N^{-\epsilon} \tag{19}$$

It can be seen that (19) is satisfied for (denote $Z = (1+\epsilon)\ln N$ in below equation)

$$t \geq \frac{2}{3}\left((1+|\rho|)Z + \sqrt{\left((1+|\rho|)Z\right)^2 + 9\left(1+|\rho|^2\right)D(\mathcal{K}_l(\mathcal{I}))Z}\right) \tag{20}$$

Like the random case, the next step would have been to bound $D(\mathcal{K}_l(\mathcal{I}))$. Fortunately, we can again assume $D(\mathcal{K}_l(\mathcal{I})) \leq (1+\gamma)\Delta$ from Proposition 5.6; even when $\mathcal{I}$ is planted, since $\boldsymbol{a}_i \xleftarrow{\text{iid}} \text{Unif}(R_k^n)$.

So, let us now proceed to get an upper bound on $\left\|\sum_{i\in[m]} \widetilde{\boldsymbol{K}_i}\right\|$ with respect to Δ. Like before, we would like to choose Δ so that we get $\left\|\sum_{i\in[m]} \widetilde{\boldsymbol{K}_i}\right\| \leq \gamma\Delta$. Observe that in equation (20) we can impose

$$\Delta \geq \frac{12(1+|\rho|^2)(1+\gamma) + 4\gamma(1+|\rho|)}{3\gamma^2}(1+\epsilon)\ln N$$

to get this desired bound for all $\gamma > 0$. Again, the failure probability of Proposition 5.6 is bounded by $N^{-\epsilon}$ for this choice.

We now have $\left\|\sum_{i\in[m]} \widetilde{\boldsymbol{K}_i}\right\| \leq \gamma\Delta$. How do we proceed to get a lower bound for $\left\|\sum_{i\in[m]} \boldsymbol{K}_i\right\| = \|\mathcal{K}_l(\mathcal{I})\|$?

$$\begin{aligned}
\|\mathcal{K}_l(\mathcal{I})\| = \left\|\sum_{i\in[m]} \boldsymbol{K}_i\right\| &\geq \left\|\sum_{i\in[m]} \mathbb{E}[\boldsymbol{K}_i]\right\| - \left\|\sum_{i\in[m]} \widetilde{\boldsymbol{K}_i}\right\| \\
&\geq |\rho| \left\|\sum_{i\in[m]} A_{\boldsymbol{s},\boldsymbol{a}_i}\right\| - \gamma\Delta \quad \text{(from Lemma 5.12)} \\
&\geq |\rho|\,\Delta - \gamma\Delta \quad \text{(from Proposition 5.13)}
\end{aligned}$$

Theorem 5.15. *Let $\mathcal{I} = \{(\boldsymbol{a}_i, b_i)\}_{i\in[m]}$ be a planted k-LIN(R) instance of dimension n, noise distribution Ψ on R with $\rho = \mathbb{E}[\chi(\Psi)]$. Let $N = \binom{n}{l}(|R|-1)^l$. Suppose*

$$\Delta := \frac{m\theta^{(n,l,k)}}{N} \geq \frac{12(1+|\rho|^2)(1+\gamma) + 4\gamma(1+|\rho|)}{3\gamma^2}(1+\epsilon)\ln N$$

Then we have $\|\mathcal{K}_l(\mathcal{I})\| \geq (|\rho|-\gamma)\Delta$ except with probability at most $2N^{-\epsilon}$.

Finally let us restate this result in terms of the parameters of $\mathcal{I}$. Once again we will use the lower bounds of $\theta^{(n,l,k)}$ from (7), (8); the better result can be used depending on whether $|R| = O(n/l)$ or $|R| = \omega(n/l)$.

Corollary 5.16. *Let k, n, l be positive integers such that $\lceil k/2 \rceil \leq l \leq n - \lfloor k/2 \rfloor$. Let $0 < \gamma < 1$. Consider a k-LIN(R) instance $\mathcal{I}$ of dimension n and m constraints, $m \geq \frac{l\ln((|R|-1)n)}{\gamma^2} \cdot (|R|-1)^{\lfloor k/2 \rfloor}\left(\frac{\alpha n}{l}\right)^{\lceil k/2 \rceil}$ where $\alpha > 0$ is an absolute constant. Let $\Delta := \frac{m\theta^{(n,l,k)}}{\binom{n}{l}(|R|-1)^l}$. Then if $\mathcal{I}$ is planted with noise distribution Ψ on R such that $\rho := \mathbb{E}[\chi(\Psi)] \in \mathbb{R}\setminus\{0\}$, we get that $\|\mathcal{K}_l(\mathcal{I})\| \geq (|\rho|-\gamma)\Delta$ except with probability at most $n^{-\Omega(l)}$.*

Corollary 5.17. *Let k, n, l be positive integers such that $k \leq l \leq n$. Let $0 < \gamma < 1$. Consider a k-LIN(R) instance $\mathcal{I}$ of dimension n and m constraints, $m \geq \frac{l\ln((|R|-1)n)}{\gamma^2} \cdot \left(\frac{\alpha n}{l}\right)^k$ where $\alpha > 0$ is an absolute constant. Let $\Delta := \frac{m\theta^{(n,l,k)}}{\binom{n}{l}(|R|-1)^l}$. Then if $\mathcal{I}$ is planted with noise distribution Ψ on R such that $\rho := \mathbb{E}[\chi(\Psi)] \in \mathbb{R}\setminus\{0\}$, we get that $\|\mathcal{K}_l(\mathcal{I})\| \geq (|\rho|-\gamma)\Delta$ except with probability at most $n^{-\Omega(l)}$.*

5.4 Distinguishing between a planted and a random instance

Corollaries 5.10, 5.11 and Corollaries 5.16, 5.17 together suggest that we can solve the decisional k-LIN(R) problem with high probability if we had the ability to approximate the maximum eigenvalue of $\mathcal{K}_l(\mathcal{I})$. For this, we need to ensure that the thresholds established in these corollaries are separated. In particular, we must have $(|\rho| - \gamma)\Delta > \delta\Delta$, which can be easily achieved by setting, say, $\gamma = \delta = |\rho|/4$.

We can then use a standard eigenvalue estimation technique, say the Power Iteration method, to approximate $\|\mathcal{K}_l(\mathcal{I})\|$. Note that $\mathcal{K}_l(\mathcal{I})$ is a $N \times N$ matrix for $N = \binom{n}{l}(|R|-1)^l$. Moreover, it is a sparse matrix since each row has at most $D(\mathcal{K}_l(\mathcal{I})) = O(\Delta)$ non-zero entries, where $\Delta = O(\ln N)$ suffices for our purpose. Thus, the Power Iteration takes $\widetilde{O}(N) = \widetilde{O}\left(\binom{n}{l}(|R|-1)^l\right)$ time to converge.

Corollary 5.18. *Let k, n, l be positive integers such that $\lceil k/2 \rceil \leq l \leq n - \lfloor k/2 \rfloor$. Consider a k-LIN(R) instance $\mathcal{I}$ of dimension n and m constraints which is either random or is planted with noise distribution Ψ on R with $\rho = \mathbb{E}[\chi(\Psi)] \in \mathbb{R} \setminus \{0\}$. Let $m \geq \frac{l \ln((|R|-1)n)}{|\rho|^2} \cdot (|R|-1)^{\lfloor k/2 \rfloor} \left(\frac{\alpha n}{l}\right)^{\lceil k/2 \rceil}$ where $\alpha > 0$ is an absolute constant. Then we can determine the distribution of $\mathcal{I}$ in $\widetilde{O}\left(\binom{n}{l}(|R|-1)^l\right)$ time except with probability at most $n^{-\Omega(l)}$.*

Corollary 5.19. *Let k, n, l be positive integers such that $k \leq l \leq n$. Consider a k-LIN(R) instance $\mathcal{I}$ of dimension n and m constraints which is either random or is planted with noise distribution Ψ on R with $\rho = \mathbb{E}[\chi(\Psi)] \in \mathbb{R} \setminus \{0\}$. Let $m \geq \frac{l \ln((|R|-1)n)}{|\rho|^2} \cdot \left(\frac{\alpha n}{l}\right)^k$ where $\alpha > 0$ is an absolute constant. Then we can determine the distribution of $\mathcal{I}$ in $\widetilde{O}\left(\binom{n}{l}(|R|-1)^l\right)$ time except with probability at most $n^{-\Omega(l)}$.*

An important aspect of this distinguisher is that it works for any (nice enough) noise distribution of the planted instance. The sample complexity required by the algorithm is inversely proportional to $|\rho|^2$ where $\rho = \mathbb{E}[\chi(\Psi)]$. The sample and time complexities do not depend on any other property of the noise distribution. Also, as long as $|\rho|$ is sufficiently large (say at least inverse polynomial in n), it won't affect the sample complexity by a lot. As immediate consequences of Corollaries 5.18, 5.19 we have algorithms for solving k-sparse LWE and sparse LPN over modulus q.

For sparse LWE, we can compute $1/\rho^2 \leq e^{2\pi r^2/q^2}$ (Observation A.1). Since the width r for the LWE distribution is chosen to be less than q, this dependency will only contribute a constant term.

Theorem 5.20. *Let n, m, q, r be positive integers, $r < q/2$. Let $\alpha > 0$ be an absolute constant. There exists an algorithm that solves the decisional k-LWE$_{n,m,q,r}$ problem such that:*

- *Let $l \in \mathbb{N}$ be a parameter such that $\lceil k/2 \rceil \leq l \leq n - \lfloor k/2 \rfloor$. Then provided that the number of samples $m \geq l \ln(qn) \cdot (q-1)^{\lfloor k/2 \rfloor} \left(\frac{\alpha n}{l}\right)^{\lceil k/2 \rceil}$, the algorithm takes time $\widetilde{O}\left(\binom{n}{l}(q-1)^l\right)$ and is correct, except with probability at most $n^{-\Omega(l)}$.*
- *Let $l \in \mathbb{N}$ be a parameter such that $k \leq l \leq n$. Then provided that the number of samples $m \geq l \ln(qn) \cdot \left(\frac{\alpha n}{l}\right)^k$, the algorithm takes time $\widetilde{O}\left(\binom{n}{l}(q-1)^l\right)$ and is correct except with probability at most $n^{-\Omega(l)}$.*

For sparse LPN, $\rho = 1 - \frac{q\mu}{q-1}$ from Definition 3.15. (For $\mu \to \frac{q-1}{q}$, the planted distribution approaches the random distribution.)

Theorem 5.21. *Let n, m, q be positive integers, $0 \leq \mu < (q-1)/q$. Let $\alpha > 0$ be an absolute constant. There exists an algorithm that solves the decisional k-LPN$_{n,m,q,\mu}$ problem such that:*

- *Let $l \in \mathbb{N}$ be a parameter such that $\lceil k/2 \rceil \leq l \leq n - \lfloor k/2 \rfloor$. Then provided that the number of samples $m \geq \left(1 - \frac{q\mu}{q-1}\right)^{-2} (l \ln(qn)) \cdot (q-1)^{\lfloor k/2 \rfloor} \left(\frac{\alpha n}{l}\right)^{\lceil k/2 \rceil}$, the algorithm takes time $\widetilde{O}\left(\binom{n}{l}(|R|-1)^l\right)$ time and is correct except with probability at most $n^{-\Omega(l)}$.*
- *Let $l \in \mathbb{N}$ be a parameter such that $k \leq l \leq n$. Then provided that the number of samples $m \geq \left(1 - \frac{q\mu}{q-1}\right)^{-2} (l \ln(qn)) \cdot \left(\frac{\alpha n}{l}\right)^{k}$, the algorithm takes time $\widetilde{O}\left(\binom{n}{l}(|R|-1)^l\right)$ time and is correct except with probability at most $n^{-\Omega(l)}$.*

5.5 A quantum speedup

In essence, to distinguish a random k-LIN(R) instance from a planted one, the Kikuchi method performs estimation of the spectral norm of a sparse Hermitian matrix in a space of dimension $N = \binom{n}{l}(|R|-1)^l$. For any known classical method, such as the power iteration, eigenvalue estimation requires matrix multiplications, so it is unlikely for the time complexity to go below $\widetilde{\Omega}(N)$.

However, with quantum algorithmic techniques, we can do better. We can use the same idea as given by Schmidhuber et. al. in [SOKB25] for $R = \mathbb{Z}_2$: Using a *guiding state* created from the k-LIN(R) instance, we are able to determine whether the eigenspace corresponding to eigenvalues $\geq \lambda$ (for some threshold $\lambda > 0$) is non-empty. This enjoys a speedup on Power iteration which can be from quadratic to quartic depending on the choice of parameter l. We elaborate on this intuition in Appendix B. For even more details, the reader may refer to Appendix C of [SOKB25].

6 A Simple Algorithm

In this section, we describe another algorithm for distinguishing whether a k-LIN(R) instance is random or planted. This algorithm is simpler than the spectral method and outperforms the latter when $|R|$ is large, specifically $|R| = \omega(n/l)$; like the spectral method, l here is a parameter used for sample-time complexity tradeoff. This algorithm in essence, is the dense-minor attack proposed in [JLS24]. It is also similar to the simple refutation algorithm proposed in Section 8 of [KM25], but is more efficient since we process only one dense minor instead of all $\binom{n}{l}$ dense minors [8]. We describe this in Algorithm 1.

Algorithm 1 DENSE MINOR

Require: k-LIN(R) instance $\mathcal{I}$ which is either random or is planted with respect to distribution Ψ on R with $\rho := \mathbb{E}[\chi(\Psi)] \neq 0$, tradeoff parameter $k \leq l \leq n$

1: $\mathcal{I}' := \{(\boldsymbol{a}, b) \in \mathcal{I} : \boldsymbol{a}$ has non-zero entries only on the first l coordinates$\}$

2: $\texttt{adv} := \max \left\{ \left| \frac{1}{|\mathcal{I}'|} \sum_{(\boldsymbol{a},b) \in \mathcal{I}'} \chi\left(b - \langle \boldsymbol{a}, \boldsymbol{x} \rangle\right) \right| : \boldsymbol{x} \in R^n \text{ has non-zero entries only on the first } l \text{ coordinates} \right\}$

3: **if** $\texttt{adv} < |\rho|/2$ **then**

4: **return** `RANDOM`

5: **end if**

6: **return** `PLANTED`

Essentially, the algorithm distinguishes by (naively) computing the advantage for a sub-instance $\mathcal{I}'$ of $\mathcal{I}$. This is efficient enough for us as it only needs to check for assignments $\boldsymbol{x}$ having only first

[8] for the refutation task, the latter was required

l coordinates non-zero. In particular, Algorithm 1 takes time $O(mn + |R|^l\,|\mathcal{I}'|\,k\log|R|)$ time since this is the time taken by steps 1 and 2.

Firstly we state that $\mathcal{I}'$ defined in step 1 of the algorithm has sufficiently many constraints. This can be shown via a routine calculation.

Lemma 6.1. *Let $\mathcal{I}$ be a k-LIN(R) instance with dimension n and m constraints. Let $S \in \binom{[n]}{l}$ be a fixed set of indices, and let $\mathcal{I}_S := \{(\boldsymbol{a}, b) \in \mathcal{I} : \boldsymbol{a}$ has non-zero entries only on indices of $S\}$. Suppose $m \geq \frac{l\ln((|R|-1)n)}{\gamma^2}\cdot\left(\frac{\alpha n}{l}\right)^k$ for some constants $\alpha, \gamma > 0$. Then $|\mathcal{I}_S| \geq \Omega(\gamma^{-2})l\log(|R|\,n)$ except with probability $n^{-\Omega(l)}$.*

Proof. Since a left-hand side $\boldsymbol{a}$ is chosen uniformly and independently from R_k^n, the probability its support lies in S is given by $p = \frac{\binom{l}{k}}{\binom{n}{k}} \geq \left(\frac{l}{en}\right)^k$. The expected number of constraints in S are $mp \geq \frac{C^k l\ln((|R|-1)n)}{\gamma^2}$ for some constant C. We can apply a Chernoff Bound (Theorem 3.1):

$$\Pr\left[|\mathcal{I}_S| < (1-\delta)\frac{C^k l\ln((|R|-1)n)}{\gamma^2}\right] < e^{-\delta^2\frac{C^k l\ln((|R|-1)n)}{2\gamma^2}} = n^{-\Omega(l)}$$

□

We proceed by assuming $\mathcal{I}'$ has $m' \geq \frac{l\ln(|R|n)}{|\rho|^2}$ constraints, where $\rho := \mathbb{E}[\chi(\Psi)] \neq 0$ for noise Ψ. For $\boldsymbol{x} \in R^n$, let us denote $\mathtt{adv}_{\boldsymbol{x}} := \frac{1}{|\mathcal{I}'|}\sum_{(\boldsymbol{a},b)\in\mathcal{I}'}\chi\left(b - \langle\boldsymbol{a},\boldsymbol{x}\rangle\right)$.

We first analyze the random case. If $\mathcal{I}$ is random, for every $(\boldsymbol{a}, b) \in \mathcal{I}'$, b is uniformly random in R, and so is $b - \langle\boldsymbol{a},\boldsymbol{x}\rangle$ for an arbitrary $\boldsymbol{x} \in R^n$. From Observation 3.1, we have $\mathbb{E}\left[\chi(b - \langle\boldsymbol{a},\boldsymbol{x}\rangle)\right] = 0$. So

$$\forall\ \boldsymbol{x} \in R^n,\ \mathbb{E}[\mathtt{adv}_{\boldsymbol{x}}] = \frac{1}{m'}\sum_{(\boldsymbol{a},b)\in\mathcal{I}'}\mathbb{E}\left[\chi(b - \langle\boldsymbol{a},\boldsymbol{x}\rangle)\right] = 0 \tag{21}$$

$$\Rightarrow \Pr\left[|\mathtt{adv}_{\boldsymbol{x}}| > |\rho|/2\right] < 2e^{-m'|\rho|^2/8} \tag{22}$$

$$\Rightarrow \Pr\left[\mathtt{adv} > |\rho|/2\right] < 2\,|R|^l\,e^{-m'|\rho|^2/8} \leq 2n^{-l/8} \tag{23}$$

where the second inequality follows from applying Hoeffding inequality (Theorem 3.2) over random variables $\{\chi(b - \langle\boldsymbol{a},\boldsymbol{x}\rangle)\}_{(\boldsymbol{a},b)\in\mathcal{I}'}$, and the last inequality follows by a union bound over all $\boldsymbol{x} \in R^n$ having non-zero entries only on first l coordinates.

Now let us see what happens in the planted case. Suppose $\mathcal{I}$ is planted with respect to the secret vector $\boldsymbol{s} \in R^n$. Let $\boldsymbol{s}' \in R^n$ be defined as follows:

$$s'_j = \begin{cases} s_j \text{ if } j \leq l \\ 0 \text{ if } j > l \end{cases}$$

Then for every $(\boldsymbol{a}, b) \in \mathcal{I}'$, $b = \langle\boldsymbol{a},\boldsymbol{s}\rangle + e = \langle\boldsymbol{a},\boldsymbol{s}'\rangle + e$ such that $\mathbb{E}[\chi(e)] = \rho$.
Thus

$$\mathbb{E}[\mathtt{adv}_{\boldsymbol{s}'}] = \frac{1}{m'}\sum_{(\boldsymbol{a},b)\in\mathcal{I}'}\mathbb{E}\left[\chi(b - \langle\boldsymbol{a},\boldsymbol{s}'\rangle)\right] = \frac{1}{m'}\sum_{(\boldsymbol{a},b)\in\mathcal{I}'}\mathbb{E}\left[\chi(e)\right] = \rho \tag{24}$$

$$\Rightarrow \Pr\left[\mathtt{adv} < |\rho|/2\right] \leq \Pr\left[|\mathtt{adv}_{\boldsymbol{s}'}| < |\rho|/2\right] \tag{25}$$

$$\leq \Pr\left[|\mathtt{adv}_{\boldsymbol{s}'} - \rho| > |\rho|/2\right] < 2e^{-m'\rho^2/8} \leq 2(|R|\,n)^{-l/8} \quad \text{(using Theorem 3.2)} \tag{26}$$

Hence we get the following result:

Theorem 6.2. *Let k, n, l be positive integers such that $k \leq l \leq n$. Consider a k-LIN(R) instance $\mathcal{I}$ of dimension n and m constraints which is either random or is planted with noise distribution Ψ on R with $\rho := \mathbb{E}[\chi(\Psi)]$. Let $m \geq \frac{l \ln((|R|-1)n)}{|\rho|^2} \cdot \left(\frac{\alpha n}{l}\right)^k$ where $\alpha > 0$ is an absolute constant. Then Algorithm 1 correctly distinguishes the distribution of $\mathcal{I}$ in $\widetilde{O}\left(|\rho|^{-2}\left(\left(\frac{\alpha n}{l}\right)^{k+1} + |R|^l\right)\right)$ time except with probability at most $n^{-\Omega(l)}$.*

Remark 6.1. *By comparing Theorem 6.2 and Corollary 5.19, Algorithm 1 outperforms the spectral method over the Kikuchi graph when $|R| = \omega(n/l)$.*

So the advantage of spectral method over the Kikuchi graph is when $|R| = O(n/l)$. For instance, if $|R| = n^\delta$ for some $0 \leq \delta < 1$, then Algorithm 1 will require roughly $(n/l)^k$ samples to solve in $\left(n^{\delta l}\right)$ time[9]. On the other hand, it can be seen from Corollary 5.18 that the spectral method requires roughly $\left(\frac{(1+\delta)\cdot n^{(1+\delta)k/2}}{\delta \cdot l^{k/2}}\right)$ samples to distinguish in the same time.

7 Distinguishing random vs planted k-ALIN(R) instances

Let us also analyze the distinguishing problem for k-ALIN(R), i.e. when the left-hand sides are at most k-sparse and follow the distribution in Definition 3.6. As it may seem intuitively, we will get substantially better time and sample complexities in this version of the problem compared to k-LIN(R).

Firstly, it can be seen that Algorithm 1 will work just as well if $\mathcal{I}$ is a k-ALIN(R) instance. Hence the problem enjoys the sample-time complexity tradeoff as laid out in Theorem 6.2.

Theorem 7.1. *Let k, n, l be positive integers such that $k \leq l \leq n$. Consider a k-ALIN(R) instance $\mathcal{I}$ of dimension n and m constraints which is either random or is planted with noise distribution Ψ on R with $\rho = \mathbb{E}[\chi(\Psi)] \neq 0$. Let $m \geq \frac{l \ln((|R|-1)n)}{|\rho|^2} \cdot \left(\frac{\alpha n}{l}\right)^k$ where $\alpha > 0$ is an absolute constant. Then Algorithm 1 correctly distinguishes the distribution of $\mathcal{I}$ in $\widetilde{O}\left(|\rho|^{-2}\left(\left(\frac{\alpha n}{l}\right)^{k+1} + |R|^l\right)\right)$ time except with probability at most $n^{-\Omega(l)}$.*

We now describe another strategy. Recall from Observation 3.7 that the probability of sampling $\mathbf{0} \in R^n$ as a left-hand side is $p_0 = |R|^{-k}$. What if we have enough samples m in $\mathcal{I}$ so that there are $m' \geq l \ln n / |\rho|^2$ constraints in $\mathcal{I}$ of the form $(\mathbf{0}, b)$? Let us call this collection $\mathcal{I}' = \{(\mathbf{0}, b)\} \subseteq \mathcal{I}$. Similar to Algorithm 1, we can compute

$$\texttt{adv} = \frac{1}{m'} \sum_{(\mathbf{0},b) \in \mathcal{I}'} \chi(b)$$

If $\mathcal{I}$ is random, we have

$$\mathbb{E}[\texttt{adv}] = \frac{1}{m'} \sum_{(\mathbf{0},b) \in \mathcal{I}'} \mathbb{E}\left[\chi(b)\right] = 0 \tag{27}$$

$$\Rightarrow \Pr\left[|\texttt{adv}| > |\rho|/2\right] < 2e^{-m'|\rho|^2/8} \leq 2n^{-l/8} \quad \text{(using Theorem 3.2)} \tag{28}$$

[9] assuming $k \ll l$, ignoring ρ and logarithmic terms for simplicity

On the other hand, if $\mathcal{I}$ is planted, for $(\mathbf{0}, b) \in \mathcal{I}'$ we have $b = e \xleftarrow{\text{iid}} \Psi$ such that $\mathbb{E}[\chi(e)] = \rho$.

$$\mathbb{E}[\texttt{adv}] = \frac{1}{m'} \sum_{(\mathbf{0},b)\in\mathcal{I}'} \mathbb{E}\left[\chi(b)\right] = \frac{1}{m'} \sum_{(\mathbf{0},b)\in\mathcal{I}'} \mathbb{E}\left[\chi(e)\right] = \rho \tag{29}$$

$$\Rightarrow \Pr\left[|\texttt{adv}| < |\rho|/2\right] \leq \Pr\left[|\texttt{adv} - \rho| > |\rho|/2\right] < 2e^{-m'\rho^2/8} \leq 2n^{-l/8} \quad \text{(using Theorem 3.2)} \tag{30}$$

Hence, like the dense minor idea, we can distinguish on the basis of whether $|\texttt{adv}| \leq |\rho|/2$.

We know that the probability of a left-hand side being $\mathbf{0}$ is $p_0 = |R|^{-k}$, so a standard application of Chernoff bound (Theorem 3.1) gives that we need $m \geq 2|R|^k |\rho|^{-2} (l \ln n)$ samples to obtain the required number of $(\mathbf{0}, b)$ samples ($m' \geq |\rho|^{-2} l \ln n$) with probability at least $1 - n^{-\Omega(l)}$.

Finally, the time complexity of this strategy is $\widetilde{O}\left(|\rho|^{-2} |R|^k\right)$ as all we need to do is collect all samples with left-hand sides $\mathbf{0}$ and compute the sum of the corresponding (real parts of) right-hand sides.

Theorem 7.2. *Let k, n, l be positive integers. Consider a k-ALIN(R) instance $\mathcal{I}$ of dimension n and m constraints which is either random or is planted with noise distribution Ψ on R with $\rho := \mathbb{E}[\chi(\Psi)] \neq 0$. Let $m \geq 2|R|^k |\rho|^{-2} (l \ln n)$ where $\alpha > 0$ is an absolute constant. Then there exists an algorithm that correctly distinguishes the distribution of $\mathcal{I}$ in $\widetilde{O}\left(|\rho|^{-2} |R|^k\right)$ time except with probability at most $n^{-\Omega(l)}$.*

Once again, we can choose the best algorithm depending on whether $|R| = \omega(n/l)$ (Theorem 7.1) or $|R| = O(n/l)$ (Theorem 7.2).

8 Discussion

8.1 Comparison with [KM25]

Independent of our work, Kocurek and Manohar [KM25] have studied the problem of refuting random and semirandom instances of k-LIN(R)[10]. For this task, they define Kikuchi graphs for finite fields and abelian groups that are similar to our construction. Their sample-time tradeoffs are also similar to ours. However, our objective is different from the refutation task; ours is the task of distinguishing a random vs a planted instance. The reason why their methods cannot be extended to our problem is that the statistic they use, the *value* of an instance (maximum fraction of satisfiable constraints) doesn't allow us to distinguish the random instance from the planted instance when the noise in the planted distribution is too high. Our distinguishing statistic, the *advantage* (Definition 3.5), doesn't suffer from this problem. As we have seen in Subsection 5.4, the advantage can help us distinguish the two cases for any noise distribution satisfying a mild condition.

Moreover, they construct separate Kikuchi constructions depending on the sparsity k being even or odd, and also depending on the ring being a field or non-field, with significantly more involved construction in the odd sparsity case.

Our Kikuchi construction is able to mitigate both these issues; it is more amenable to the distinguishing task, and is a unified and simplified construction for all rings R and both sparsity parities when $|R| > 2$. Further, we are able to construct a simpler distinguisher which has a

[10] a semirandom instance is when the left-hand sides of a k-LIN(R) instance can be arbitrarily chosen, but the right-hand sides are uniformly and independently chosen

significantly lower runtime than that of their simple refutation algorithm when the size of the ring R is large.

Let us elaborate on these aspects:

8.1.1 Distinguishing random vs planted instances for any planted noise

In the task of refutation of a k-LIN(R) instance $\mathcal{I}$, the goal is to output a certificate that a given instance is highly unsatisfiable, meaning no assignment can satisfy more than a certain fraction of the constraints. To solve the task of refuting a random/semirandom k-LIN(R) instance $\mathcal{I}$, the algorithm of [KM25] returns a certificate $\texttt{alg-val}(\mathcal{I}) \in [0,1]$ with the following guarantees:

1. $\texttt{alg-val}(\mathcal{I}) \geq \texttt{val}(\mathcal{I})$, where $\texttt{val}(\mathcal{I})$ is the maximum fraction of satisfied constraints by any assignment.
2. If $\mathcal{I}$ is random, then $\texttt{alg-val}(\mathcal{I}) \leq \frac{1}{|R|} + \varepsilon$ for an arbitrary small $\varepsilon > 0$ with high probability, which is close to the expected value $\left(\mathbb{E}\left[\texttt{val}(\mathcal{I})\right] = \frac{1}{|R|}\right)$.
3. The algorithm requires $m \geq \varepsilon^{-2} O(n) \left((|R|-1)n/l\right)^{k/2-1}$ samples for the refutation task.

To utilize their algorithm for the distinguishing task, a natural thought is to utilize Item 1 of above for the planted case, hoping that $\texttt{val}(\mathcal{I})$ will be high for the planted case. However, we elaborate below why this will not hold for noise distributions with low probability ($\sim 1/q$) of zero error, and so for those cases we won't be able to use their construction for the distinguishing task. We also elaborate how our construction mitigates this.

Let us see what the quantity $\texttt{val}(\mathcal{I})$ looks like if $\mathcal{I}$ is a planted instance with respect to a secret assignment $\boldsymbol{s} \in R^n$ and noise distribution Ψ on R. For any sample $(\boldsymbol{a}, b) \in \mathcal{I}$, we have $b = \langle \boldsymbol{a}, \boldsymbol{s} \rangle + e$ for $e \xleftarrow{\text{iid}} \Psi$, so the probability that $\boldsymbol{s}$ satisfies the constraint $(\boldsymbol{a}, b)$ is $\Psi(0)$. This in turn means that the expected number of satisfied constraints by $\boldsymbol{s}$ is $\Psi(0)$, and $\mathbb{E}[\texttt{val}(\mathcal{I})] \geq \Psi(0)$.

This will actually suffice to solve (distinguish) the sparse LPN problem over the ring $\mathbb{Z}_q$, as $\Psi(0) = 1 - \mu$ (see Definition 3.15) which can be assumed to be non-negligibly larger than the expected value in the random case $1/q$ (say $\mu \geq \frac{q-1}{q} - \frac{1}{n^d}$ for some constant $d > 0$). Due to this, the ε parameter will contribute a factor of n^{2d} in the sample complexity cost. Note that this matches our result for sparse LPN (Theorem 5.21).

For sparse LWE, $\Psi = \Psi_{\mathbb{Z}_q,r}$ for which $\Psi_{\mathbb{Z}_q,r}(0) = O(1/r)$ when the width r is super constant (from Observation 3.12). For instance, if $r = q/10$, we must have

$$\texttt{alg-val}(\mathcal{I}) \leq 1/q + \varepsilon < \Psi(0) = 10/q \;\; \Rightarrow \varepsilon < 9/q$$

In this case, ε contributes a factor of q^2 to the sample complexity, making the total contribution of the modulus $q^{k/2+1}$. In contrast, our result (Theorem 5.20) has a factor of $q^{k/2}$ in the sample complexity.

But what if the noise distribution is slightly perturbed, so that the mass at $0 \in \mathbb{Z}_q$ is $1/q$? Indeed, we can have a noise Ψ' resembling the discrete gaussian, except that $\Psi'(0) = 1/q$, and $\Psi'(j) = \Psi_{\mathbb{Z}_q,r}(j) + \Psi_{\mathbb{Z}_q,r}(0)/(q-1) - 1/(q(q-1))$ for $j \in \mathbb{Z}_q \setminus \{0\}$. In this case, distinguishing on the basis of $\texttt{val}(\mathcal{I})$ is not possible.

On the other hand, our distinguisher examines a different quantity $\texttt{adv}(\mathcal{I})$ (recall Definition 3.5) to separate the random and planted instances. In the random case, this quantity has expectation 0.

For the planted case we have $\mathbb{E}\left[\mathtt{adv}(\mathcal{I})\right] \geq \mathbb{E}\left[\mathrm{adv}_{\mathcal{I}}(\boldsymbol{s})\right] = \mathbb{E}[\mathrm{Re}\left(\chi(\Psi')\right)] =: \rho'$. To estimate ρ', firstly note that $\rho := \mathbb{E}[\chi(\Psi_{\mathbb{Z}_q,r})] \geq e^{-\pi r^2/q^2} \geq e^{-\pi/4}$ when $r < q/2$ (Observation 3.13). Then

$$\begin{aligned}
\rho' := \mathbb{E}[\mathrm{Re}\left(\chi(\Psi')\right)] &= \sum_{j\in\mathbb{Z}_q} e^{2\pi\iota j/q}\Psi'(j) \quad (\text{note that } \Psi'(j) = \Psi'(-j)) \\
&= 1/q + \sum_{j\in\mathbb{Z}_q\setminus\{0\}} e^{2\pi\iota j/q}\left(\Psi_{\mathbb{Z}_q,r}(j) + \Psi_{\mathbb{Z}_q,r}(0)/(q-1) - 1/(q(q-1))\right) \\
&= \rho - \left(\Psi_{\mathbb{Z}_q,r}(0) - 1/q\right)(q/(q-1)) \\
&\geq \rho - 2/r \ = \Omega(1)
\end{aligned}$$

Hence, in this case, ρ' is still a constant, and the same sample complexity as for sparse LWE suffices for the distinguishing task. In fact, our method is *oblivious* to the noise distribution Ψ, as long as it satisfies a mild condition ($\mathbb{E}[\mathrm{Re}\left(\chi(\Psi)\right)] \in (0,1]$).

In Observation 3.6 of [KM25], Kocurek and Manohar relate the value of the instance with a quadratic form of their Kikuchi construction. One key aspect of their Kikuchi construction (for their case of even k, R finite field) is that for a constraint $(\boldsymbol{a}, b) \in \mathcal{I}$, they consider all edges $(\boldsymbol{u}, \boldsymbol{v}) \in R_l^n$ such that $\boldsymbol{v} - \boldsymbol{u} = \beta\boldsymbol{a}$ for all $\beta \in R \setminus \{0\}$. This is equivalent to adding constraints $(\beta\boldsymbol{a}, \beta b) \in \mathcal{I}$ for all $\beta \in R \setminus \{0\}$ for each original constraints. Informally, equation (31) below captures the contribution of this constraint to the quadratic form for some assignment $\boldsymbol{x} \in R^n$:

$$\sum_{\beta\in R\setminus\{0\}} \chi\left(\beta b - \beta\langle\boldsymbol{a},\boldsymbol{x}\rangle\right) = \sum_{\beta\in R\setminus\{0\}} \chi\left(\beta\left(b - \langle\boldsymbol{a},\boldsymbol{x}\rangle\right)\right) = \begin{cases} (|R|-1) & \text{if } b = \langle\boldsymbol{a},\boldsymbol{x}\rangle \\ -1 \text{ if } b \neq \langle\boldsymbol{a},\boldsymbol{x}\rangle \end{cases} \tag{31}$$

So the extent of the noise (how far is b from $\langle\boldsymbol{a},\boldsymbol{x}\rangle$) is not captured in the quadratic form.

On the other hand, we only keep edges of the form $\boldsymbol{v} - \boldsymbol{u} = \pm\boldsymbol{a}$. By this, we will achieve the appropriate contribution of this constraint in the advantage:

$$\sum_{\beta\in\{\pm1\}} \chi\left(\beta b - \beta\langle\boldsymbol{a},\boldsymbol{x}\rangle\right) = 2\mathrm{Re}\left(\chi\left(b - \langle\boldsymbol{a},\boldsymbol{x}\rangle\right)\right) \tag{32}$$

We show in Proposition 5.1 how the quadratic form of our Kikuchi construction relates to our notion of advantage.

8.1.2 A unified construction

[KM25] produce separate Kikuchi constructions and analyses depending on the sparsity parameter k being odd or even, and also depending on the ring R being a field or $\mathbb{Z}_{q_1} \times \ldots \times \mathbb{Z}_{q_t}$.

- Given a left-hand side $\boldsymbol{a} \in R_k^n$, their construction for even k only considers edges $(\boldsymbol{u}, \boldsymbol{v}) \in R_l^n \times R_l^n$ wherein the symmetric difference of supports of $\boldsymbol{u}$ and $\boldsymbol{v}$ equal to the support of $\boldsymbol{a}$ (see Remark 4.2). Such pairs will not exist when k is odd. So for odd k, they revert to a much more involved setup which includes partitioning $\mathcal{I}$ into multiple sub-instances, building a Kikuchi graph for each sub-instance (which differs from the even case), and performing a further processing step of deleting edges from high-degree vertices before taking the spectral norm.

  In our construction, we do not restrict our edges to satisfy the above restriction for even k. This means that when $|R| > 2$, it is possible that for $\boldsymbol{v} - \boldsymbol{u} = \boldsymbol{a}$, we have $u_j, v_j, a_j \neq 0$ for some

$j \in [n]$ (condition 2(c) of Construction 4.3). This also allows valid pairs $(\boldsymbol{u}, \boldsymbol{v})$ for $\boldsymbol{a} \in R_k^n$ when k is odd and $|R| > 2$.

- When R is the abelian group $G = \mathbb{Z}_{q_1} \times \ldots \times \mathbb{Z}_{q_t}$, they reduce the instance to a quotient group G/H which is either small, is of prime order (i.e. a field) or does not have small subgroups. This helps them to save a factor of $|R|$ in their sample complexity dependence. Here too, their Kikuchi matrix (for even k too) differs from the field case.

  For reasons stated above, our construction does not consider edges of the form $\boldsymbol{v} - \boldsymbol{u} = \beta \boldsymbol{a}$ for $\beta \in R \setminus \{0\} \setminus \{\pm 1\}$, which helped them (when R was a field) save a factor of $|R|$ in their sample complexity. Thus using our construction for non-fields as well does not increase dependence on $|R|$ in our sample complexity.

We also remark that our analysis works when R is any ring apart from a finite field or $\mathbb{Z}_{q_1} \times \ldots \times \mathbb{Z}_{q_t}$, i.e. any arbitrary multiplication operation on the ring is allowed.

8.1.3 Other Differences

- They use the Trace Moment method to bound the spectral norm of the Kikuchi matrix, as they deal with the semirandom instance, where the left-hand sides are arbitrary. Since we deal with instances where the left-hand sides are also random (in both random and planted instances), we use the Matrix Bernstein inequality whose analysis is simpler.

- They give a simpler algorithm for the refutation task which outperforms the Kikuchi method when $|R|$ is large, specifically $|R| = \omega(n/l)$. This algorithm takes same time as their Kikuchi method which is $(|R|\, n)^{O(l)}$. This algorithm involves checking all assignments in R^l for each $\binom{n}{l}$ l-sized subsets of $[n]$.

  We too give a simple algorithm in the large $|R|$ case, which takes time $\widetilde{O}\left(\rho^{-2}\left(\left(\frac{\alpha n}{l}\right)^{k+1} + |R|^l\right)\right)$. We are able to achieve this significant advantage in time, as processing a fixed l-sized subset suffices for the task of distinguishing a random vs planted instance.

8.2 Comparison with [JLS24]

Jain, Lin and Saha in [JLS24] introduced the sparse LWE problem over the ring $\mathbb{Z}_p$ for a prime p. There is an ambiguity in their definition, as they mention that each sample has *exactly* k-non-zero entries, yet each non-zero entry is sampled uniformly from $\mathbb{Z}_p$ rather than $\mathbb{Z}_p^*$. In this work, we have defined and analyzed both versions (k-LIN(R) and k-ALIN(R)).

They attempt to claim hardness of sparse LWE, both by a reduction from standard LWE, and via a conjecture. We now discuss the implications of our algorithms on these hardness assumptions.

8.2.1 On hardness of k-ALWE$_{n,m,p,r}$ assuming hardness of LWE

Over the ring $R = \mathbb{Z}_p$, they show a reduction from standard LWE over dimension k to k-sparse LWE over dimension $n \geq k$ (Theorem 5.1 of [JLS24]) for the at most k-non-zero entries version, i.e. k-ALWE$_{n,m,p,r}$ as per our notation in Definition 3.14. The reduction preserves the number of samples m. It is a standard assumption in cryptography that solving standard LWE on dimension k takes time $2^{\Omega(k)}$ time given arbitrary many samples (even as many as $p^{O(k)}$). Based on this assumption, their result establishes that with these many samples, solving k-ALWE$_{n,m,p,r}$ also takes $2^{\Omega(k)}$ time.

On the other hand, in Section 7 we show that solving k-ALWE$_{n,m,p,r}$ is achievable in $\widetilde{O}\left(p^k\right)$ time given enough ($m = \widetilde{O}\left(p^k\right)$) samples. So when the number of samples is large and the modulus p is small, our result is towards matching the hardness assumption for k-ALWE$_{n,m,p,r}$ from their reduction.

8.2.2 On conjectured hardness of k-LWE$_{n,m,p,r}$, k-ALWE$_{n,m,p,r}$

As an interesting case, let us analyze the sample complexity required by our dense minor algorithm (see Theorem 6.2) if we allow a time of $2^{\delta n}$ for an arbitrarily small $\delta > 0$. We have roughly $m \geq \left(\frac{n}{l}\right)^k$ samples required for a time of roughly $(n/l)^k + p^l$. We must have

$$(n/l)^k + p^l \approx 2^{\delta n} \quad \Rightarrow l \approx \frac{\delta n}{\log p}$$
$$\Rightarrow m = O\left((\log p)^k\right) \text{ samples are sufficient}$$

In particular, if $k = O(\log n / \log\log p)$ then polynomially many samples suffice. If $k = O(\log n)$, then we require $(\log p)^{O(\log n)}$ samples which is polynomial for constant modulus p, and slightly more than polynomial for $p = O\left(\exp(n)\right)$.

Conjecture 7.1 in [JLS24] suggests that k-LWE$_{n,m,p,r}$ is as hard as standard LWE over $\Theta(n)$ dimension, poly(n) samples and the same modulus and width parameters, when the sparsity $k = \Omega(\log n)$ and $m = \text{poly}(n)$. Assuming hardness of standard LWE, the algorithm from the Kikuchi method nearly closes the gap of this conjecture, either when the sparsity is slightly less, or the polynomial samples requirement is slightly relaxed.

References

[ABW10] Benny Applebaum, Boaz Barak, and Avi Wigderson. "Public-key cryptography from different assumptions". In: *Proceedings of the forty-second ACM symposium on Theory of computing*. 2010, 171–180.

[ADINZ17a] Benny Applebaum, Ivan Damgård, Yuval Ishai, Michael Nielsen, and Lior Zichron. "Secure arithmetic computation with constant computational overhead". In: *Annual International Cryptology Conference*. Springer. 2017, pp. 223–254.

[ADINZ17b] Benny Applebaum, Ivan Damgård, Yuval Ishai, Michael Nielsen, and Lior Zichron. "Secure arithmetic computation with constant computational overhead". In: *Annual International Cryptology Conference*. Springer. 2017, pp. 223–254.

[AL16a] Benny Applebaum and Shachar Lovett. "Algebraic attacks against random local functions and their countermeasures". In: *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*. 2016, pp. 1087–1100.

[AL16b] Benny Applebaum and Shachar Lovett. "Algebraic attacks against random local functions and their countermeasures". In: *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*. 2016, 1087–1@phdthesismanohar2019new, title=New Spectral Techniques in Algorithms, Combinatorics, and Coding Theory: The Kikuchi Matrix Method, author=Manohar, Peter, year=2019, school=Weizmann Institute 100.

[Ale03] Michael Alekhnovich. "More on average case vs approximation complexity". In: *44th Annual IEEE Symposium on Foundations of Computer Science, 2003. Proceedings.* IEEE. 2003, pp. 298–307.

[AOW15a] Sarah R Allen, Ryan O'Donnell, and David Witmer. "How to refute a random CSP". In: *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*. IEEE. 2015, pp. 689–708.

[AOW15b] Sarah R Allen, Ryan O'Donnell, and David Witmer. "How to refute a random CSP". In: *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*. IEEE. 2015, pp. 689–708.

[BBTV25] Kiril Bangachev, Guy Bresler, Stefan Tiegel, and Vinod Vaikuntanathan. "Near-optimal time-sparsity trade-offs for solving noisy linear equations". In: *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*. 2025, pp. 1910–1920.

[BV14] Zvika Brakerski and Vinod Vaikuntanathan. "Lattice-based FHE as secure as PKE". In: *Proceedings of the 5th conference on Innovations in theoretical computer science*. 2014, pp. 1–12.

[CM01] Mary Cryan and Peter Bro Miltersen. "On pseudorandom generators in NC0". In: *International Symposium on Mathematical Foundations of Computer Science*. Springer. 2001, pp. 272–284.

[Fei02] Uriel Feige. "Relations between average case complexity and approximation complexity". In: *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*. 2002, pp. 534–543.

[Gen09] Craig Gentry. "Fully homomorphic encryption using ideal lattices". In: *Proceedings of the forty-first annual ACM symposium on Theory of computing*. 2009, pp. 169–178.

[GJLSV25] Riddhi Ghosal, Aayush Jain, Paul Lou, Amit Sahai, and Neekon Vafa. "Post-quantum PKE from Unstructured Noisy Linear Algebraic Assumptions: Beyond LWE and Alekhnovich's LPN". In: *Annual International Conference on the Theory and Applications of Cryptographic Techniques*. Springer. 2025, pp. 64–93.

[GLG22] Sevag Gharibian and François Le Gall. "Dequantizing the quantum singular value transformation: hardness and applications to quantum chemistry and the quantum PCP conjecture". In: *Proceedings of the 54th annual ACM SIGACT symposium on theory of computing*. 2022, pp. 19–32.

[Gol00] Oded Goldreich. "Candidate One-Way Functions Based on Expander Graphs". In: (2000).

[Gol11] Oded Goldreich. "Candidate one-way functions based on expander graphs". In: *Studies in Complexity and Cryptography. Miscellanea on the Interplay between Randomness and Computation: In Collaboration with Lidor Avigad, Mihir Bellare, Zvika Brakerski, Shafi Goldwasser, Shai Halevi, Tali Kaufman, Leonid Levin, Noam Nisan, Dana Ron, Madhu Sudan, Luca Trevisan, Salil Vadhan, Avi Wigderson, David Zuckerman*. Springer, 2011, pp. 76–87.

[Has20] Matthew B Hastings. "Classical and quantum algorithms for tensor principal component analysis". In: *Quantum* 4 (2020), p. 237.

[Hås84] Johan Håstad. *An NP-complete problem: some aspects of its solution and some possible applications*. Institut Mittag-Leffler, 1984.

[JLS24] Aayush Jain, Huijia Lin, and Sagnik Saha. "A systematic study of sparse LWE". In: *Annual International Cryptology Conference*. Springer. 2024, pp. 210–245. URL: https://eprint.iacr.org/2024/1589.

[Kik51] Ryoichi Kikuchi. "A theory of cooperative phenomena". In: *Physical review* 81.6 (1951), p. 988.

[KM25] Nicholas Kocurek and Peter Manohar. "Spectral Refutations of Semirandom k-LIN over Larger Fields". In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2025)*. Schloss Dagstuhl–Leibniz-Zentrum für Informatik. 2025, pp. 17–1. URL: https://arxiv.org/pdf/2508.18185.

[KMOW17a] Pravesh K Kothari, Ryuhei Mori, Ryan O'Donnell, and David Witmer. "Sum of squares lower bounds for refuting any CSP". In: *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*. 2017, pp. 132–145.

[KMOW17b] Pravesh K Kothari, Ryuhei Mori, Ryan O'Donnell, and David Witmer. "Sum of squares lower bounds for refuting any CSP". In: *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*. 2017, pp. 132–145.

[LPR13] Vadim Lyubashevsky, Chris Peikert, and Oded Regev. "On ideal lattices and learning with errors over rings". In: *Journal of the ACM (JACM)* 60.6 (2013), pp. 1–35.

[MST03] Elchanan Mossel, Amir Shpilka, and Luca Trevisan. "On epsilon-Biased Generators in NC^ 0". In: *Annual Symposium on Foundations of Computer Science*. Vol. 44. IEEE Computer Society Press. 2003, pp. 136–145.

[Pei09] Chris Peikert. "Public-key cryptosystems from the worst-case shortest vector problem". In: *Proceedings of the forty-first annual ACM symposium on Theory of computing* (2009), pp. 333–342.

[Pei16] Chris Peikert. "A decade of lattice cryptography". In: *Foundations and trends® in theoretical computer science* 10.4 (2016), pp. 283–424.

[Pie12] Krzysztof Pietrzak. "Cryptography from learning parity with noise". In: *International Conference on Current Trends in Theory and Practice of Computer Science*. Springer. 2012, pp. 99–114.

[Reg05] Oded Regev. "On Lattices, Learning with Errors, Random Linear Codes, and Cryptography". In: *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing* (2005), pp. 84–93.

[RRS17] Prasad Raghavendra, Satish Rao, and Tselil Schramm. "Strongly refuting random CSPs below the spectral threshold". In: *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*. 2017, pp. 121–131.

[SOKB25] Alexander Schmidhuber, Ryan O'Donnell, Robin Kothari, and Ryan Babbush. "Quartic quantum speedups for planted inference". In: *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. SIAM. 2025, pp. 905–913. URL: https://arxiv.org/abs/2406.19378.

[Tro12] Joel A Tropp. "User-friendly tail bounds for sums of random matrices". In: *Foundations of computational mathematics* 12.4 (2012), pp. 389–434. URL: https://arxiv.org/abs/1004.4389.

[WAM19] Alexander Wein, Ahmed El Alaoui, and Cristopher Moore. "The Kikuchi hierarchy and tensor PCA". In: *Journal of the ACM* (2019).

A Facts on Discrete Gaussian

Recall the discrete Gaussian distribution $\Psi_{\mathbb{Z}_q,r}$ on $R = \mathbb{Z}_q$ with width parameter r (Definition 3.11). We will state a couple of facts regarding this distribution which we use.

Recall the Poisson Summation Formula for $f : \mathbb{R} \to \mathbb{R}$ where $\hat{f}$ denotes its Fourier transform:

$$\sum_{x\in\mathbb{Z}} f(x) = \sum_{y\in\mathbb{Z}} \hat{f}(y)$$

A.1 Estimating $\Psi_{\mathbb{Z}_q,r}(0)$

We have

$$\Psi_{\mathbb{Z}_q,r}(0) = \frac{\sum_{x\in q\mathbb{Z}} e^{-\pi x^2/r^2}}{\sum_{x\in\mathbb{Z}} e^{-\pi x^2/r^2}} = \frac{\sum_{x\in\mathbb{Z}} e^{-\pi (q/r)^2 x^2}}{r\sum_{x\in\mathbb{Z}} e^{-\pi r^2 x^2}}$$

where the last equality holds by applying the Poisson summation formula on the denominator. When $1 < r < q$, both summations in the numerator and denominator can be bounded by constants. Thus $\Psi_{\mathbb{Z}_q,r}(0) = \Theta(1/r)$.

A.2 Estimating $\mathbb{E}[\chi(\Psi_{\mathbb{Z}_q,r})]$

For $e \leftarrow \Psi_{\mathbb{Z}_q,r}$, we then have $\rho = \mathbb{E}[\chi(e)] = \mathbb{E}[\omega_q(e)]$. This means that $r \in [0,\infty]$ is the (unique) solution to the equation

$$\rho = \mathbb{E}[\eta_i] = \frac{\sum_{j\in\mathbb{Z}_q} \omega_q^j \sum_{x\in q\mathbb{Z}+j} e^{-\pi x^2/r^2}}{\sum_{x\in\mathbb{Z}} e^{-\pi x^2/r^2}} \tag{33}$$

$$= \frac{\sum_{x\in\mathbb{Z}} \omega_q^x e^{-\pi x^2/r^2}}{\sum_{x\in\mathbb{Z}} e^{-\pi x^2/r^2}} = \frac{\sum_{x\in\mathbb{Z}} \cos\left(\frac{2\pi x}{q}\right) e^{-\pi x^2/r^2}}{\sum_{x\in\mathbb{Z}} e^{-\pi x^2/r^2}} \tag{34}$$

Let us obtain a lower bound for (34). For $g(x) = \cos\left(\frac{2\pi x}{q}\right) e^{-\pi x^2/r^2}$, $\hat{g}(y) = \frac{r}{2}\left(e^{-\pi r^2\left(y-\frac{1}{q}\right)^2} + e^{-\pi r^2\left(y+\frac{1}{q}\right)^2}\right)$ and for $h(x) = e^{-\pi x^2/r^2}$, $\hat{h}(y) = re^{-\pi r^2 y^2}$. Using these in (34), we obtain

$$\rho = \frac{\sum_{y\in\mathbb{Z}} \hat{g}(y)}{\sum_{y\in\mathbb{Z}} \hat{h}(y)} = \frac{e^{-\pi r^2/q^2}\sum_{y\in\mathbb{Z}} \cosh\left(\frac{2\pi r^2 y}{q}\right) e^{-\pi r^2 y^2}}{\sum_{y\in\mathbb{Z}} e^{-\pi r^2 y^2}} \geq e^{-\pi r^2/q^2}$$

In our results, we see that the sample complexity required for both distinguishing methods using Kikuchi graphs depends inverse polynomially on ρ. Since the width r for the LWE distribution is chosen to be less than q, this dependency will only contribute a constant term.

B A quantum speedup

Some Notation. $\boldsymbol{x}^\dagger$ denotes the conjugate transpose of a vector $\boldsymbol{x}$. $|\boldsymbol{x}\rangle$ denotes the unit vector in the direction of $\boldsymbol{x}$ and $\langle \boldsymbol{x}| := |\boldsymbol{x}\rangle^\dagger$. For a matrix A, $\|A\|_{\max} := \max_{i,j} |A_{ij}|$.

In essence, to solve decisional k-LIN(R), the Kikuchi method performs estimation of the spectral norm of a sparse Hermitian matrix in a space of dimension $N = \binom{n}{l}(|R|-1)^l$. For any known classical method, such as the power iteration, eigenvalue estimation requires matrix multiplications, so it is unlikely for the time complexity to go below $\widetilde{\Omega}(N)$. However, with quantum algorithmic techniques, we can do better.

A Hamiltonian H on t qubits is a Hermitian matrix of dimension 2^t. H is called *s-sparse* if each row and each column of H has at most s non zero entries. Throughout the discussion, we shall assume the normalization constraint that $\|H\|_{\max} \leq 1$. Approximating the spectral norm of a sparse Hamiltonian is well studied in quantum computing. Indeed, for a local Hamiltonian (which is a special case of sparse hamiltonian), this problem is QMA complete. However, if we are provided with a *guiding* state $|\gamma\rangle$, i.e. a state that has good overlap with a unit eigenvector $|v\rangle$ corresponding to the maximum absolute eigenvalue, i.e. $(\langle\gamma|v\rangle)^2$ is large, the problem can be easier. If we are provided with a guiding state having $1/\texttt{poly}(t)$ overlap with the leading eigenvector, then Gharibian and Le Gall [GLG22] show that the problem is BQP-complete.

Schmidhuber et al. [SOKB25] provide a quantum algorithm to solve the Guided Sparse Hamiltonian problem by combining standard quantum algorithmic techniques of Hamiltonian simulation, phase estimation and amplitude estimation. We define the problem and state their result below. For details, the reader may refer to their paper.

B.1 Guided Sparse Hamiltonian Problem

The input Hamiltonian H is assumed to be given in the form of quantum oracle access by providing oracles O_H and O_F. O_H is called the *adjacency matrix* oracle, which on input (i, j) (for $i, j \in [2^t]$) gives the entry H_{ij} of H up to rounding. O_F is the *adjacency list* oracle, which on input (i, r) (for $i \in [2^t], r \in [s]$) either returns the index of the rth non-zero entry in the ith row, or indicates that there are less than r non-zero entries in i^{th} row.

The advantage of having oracle access to H in this manner is that we do not have to store the Hamiltonian H that may possibly be of large dimension; rather we can access it "on the fly" when required.

Definition B.1. *(Definition 62 in [SOKB25]) In the GSHS problem, we are given the following as input :*

1. *An s-sparse hamiltonian H on t qubits with $\|H\|_{\max} \leq 1$ as specified with the oracles O_H and O_F.*
2. *A quantum circuit using G gates which performs the mapping : $|0^t\rangle |0^A\rangle \mapsto |\Psi\rangle |0^A\rangle$.*
3. *Parameters $\lambda \in \mathbb{R}, \alpha \in (0, 1), \gamma \in (0, 1]$.*

To solve it, one must output :

- `YES`*, if $\|\Pi_{\geq\lambda}(H) |\Psi\rangle\| \geq \gamma$ (which also implies $\|H\| \geq \lambda$)*
- `NO`*, if $\|H\| \leq (1 - \alpha)\lambda$*

Theorem B.1. *(Theorem 63 in [SOKB25]) The Guided Sparse Hamiltonian problem can be solved with high probability by a quantum algorithm that uses:*

- $(Q = \widetilde{O}(s/\gamma\alpha\lambda)$ *queries to the oracles*
- $\widetilde{O}\big(G/\gamma + \mathrm{polylog}(Q)/\gamma + Qt\big)$ *gates*
- $(A + O(t) + \widetilde{O}(\log Q)$ *qubits*

For a k-LIN(R) instance $\mathcal{I}$ of dimension n and m samples, the associated Hermitian matrix $\mathcal{K}_l(\mathcal{I})$ has the following parameters associated with it:

1. $2^t = N = \binom{n}{l}(q^l$, so $t = O(l \log n)$
2. Its sparsity $s = \Delta = \frac{m\theta^{(n,l,k)}}{\binom{n}{l}(q^l} = O(\log N) = O(l \log n)$ for the separation in spectral norm for random vs planted case (from Propositions 5.9, 5.15)
3. $\lambda > (1 + O(1))\rho\Delta$, say $\lambda = 3/2\rho\Delta$ (lower bound of spectral norm for planted case)
4. $(1 - \alpha)\rho\Delta = (1 - O(1))\rho\Delta$, say $\alpha = 1/2$ (upper bound of spectral norm for random case)

Additionally, the oracles O_H and O_F can be efficiently computed using Construction 4.1.

From point 3. of Definition B.1, γ^2 is the parameter which corresponds to the overlap of the guiding state $|\Psi\rangle$. If γ is less than inverse polynomial, then from points 1 and 2 of Theorem B.1, $1/\gamma$ is the dominating factor for the time (i.e. gate) complexity.

Note that if $|\Psi\rangle$ is a random unit vector, then its overlap (with high probability) to the cutoff eigenspace (corr. to $\Pi_{\geq\lambda}$ that in the worst case is of small dimension) will be inversely proportional to the dimension with high probability, i.e.

$$\gamma^2 = \Omega\left(\frac{1}{\binom{n}{l}(|R| - 1)^l}\right)$$

in our setting. This already yields a quantum algorithm with complexity $\frac{\mathrm{poly}(n)}{\gamma} = \mathrm{poly}(n) \cdot \sqrt{\binom{n}{l}(|R| - 1)^l}$, a (nearly) quadratic speedup over the classical algorithm.

The other key observation of [SOKB25] is the ability to prepare a better guiding state, one which has overlap of $\Omega\left(\frac{1}{\sqrt{\binom{n}{l}(|R|-1)^l}}\right)$ with the cutoff-eigenspace.

B.2 Intuition for Guiding State

Again, we give a brief intuition behind the better guiding state; the details of the proofs will be analogous to Appendix C of [SOKB25].

Suppose for an instance $\mathcal{I} = \{(\boldsymbol{a}_i, b_i)\}_{i\in[m]}$ we construct l/k copies of the following state:

$$|\gamma(\mathcal{I})\rangle := \frac{1}{\sqrt{m}} \sum_i \chi(b_i) \, |\boldsymbol{a}_i\rangle \in \mathbb{C}^{R_k^n}$$

This state correlates well (better than random) with the planted assignment $|\boldsymbol{s}\rangle = \frac{1}{\sqrt{n}} \sum_i \chi(s_i) \, |i\rangle$:

$$|\langle \boldsymbol{s}^{\otimes \ell} | \gamma(\mathcal{I})^{\otimes \ell/k} \rangle|^2 = \left(\langle \boldsymbol{s}^{\otimes k} | \gamma(\mathcal{I}) \rangle \right)^{2\ell/k} \approx \left(\frac{\rho \cdot m}{n^k q^k} \right)^{\ell/k}$$

If $m \approx \Theta(((|R|-1)n)^{k/2})$ then the overlap $\sim ((|R|-1)n)^{-l/2}$, which is a quadratic improvement from that of a random guiding state. The actual guiding state is approximately a symmetrized and a normalized version of $|\gamma(\mathcal{I})\rangle$. Also, the other technical challenge is to show a high overlap with the cutoff eigenspace of $\mathcal{K}_l(\mathcal{I})$ rather than with $\boldsymbol{s}$. These issues are handled in detail in Appendix C of [SOKB25].