Fast Decoding of Explicit almost Optimal ε-balanced q-ary Codes and Fast Approximation of Expanding k-CSPs

Fernando Granha Jeronimo*

June 5, 2022

WORKING DRAFT
Please do not distribute

Good codes over an alphabet of constant size q can approach but not surpass distance $1 - 1/q$. This makes the use of q-ary codes a necessity in some applications, and much work has been devoted to the case of constant alphabet q. In the large distance regime, namely, distance $1 - 1/q - \varepsilon$ for small $\varepsilon > 0$, the Gilbert–Varshamov (GV) bound asserts that rate $\Omega_q(\varepsilon^2)$ is achievable whereas the q-ary MRRW bound gives a rate upper bound of $O_q(\varepsilon^2 \log(1/\varepsilon))$. In this sense, the GV bound is almost optimal in this regime. To the best of our knowledge prior to this work there was no known explicit and efficiently decodable q-ary codes near the GV bound, in this large distance regime, for any constant $q \geq 3$.

We design an $\widetilde{O}_{\varepsilon,q}(N)$ time decoder for explicit (expander based) families of linear codes $\mathcal{C}_{N,q,\varepsilon} \subseteq \mathbb{F}_q^N$ of distance $(1 - 1/q)(1 - \varepsilon)$ and rate $\Omega_q(\varepsilon^{2+o(1)})$, for any desired $\varepsilon > 0$ and any constant prime q, namely, almost optimal in this regime. These codes are ε-balanced,i.e., for every non-zero codeword, the frequency of each symbol lies in the interval $[1/q - \varepsilon, 1/q + \varepsilon]$. A key ingredient of the q-ary decoder is a new near-linear time approximation algorithm for linear equations (k-LIN) over $\mathbb{Z}_q$ on expanding hypergraphs, in particular, those naturally arising in the decoding of these codes.

We also investigate k-CSPs on expanding hypergraphs in more generality. We show that special trade-offs available for k-LIN over $\mathbb{Z}_q$ hold for linear equations over a finite group. To handle general finite groups, we design a new matrix version of weak regularity for expanding hypergraphs. We also obtain a near-linear time approximation algorithm for general expanding k-CSPs over q-ary alphabet. This later algorithm runs in time $\widetilde{O}_{k,q}(m + n)$, where m is the number of constraints and n is the number of variables. This improves the previous best running time of $O(n^{\Theta_{k,q}(1)})$ by a Sum-of-Squares based algorithm of [AJT, 2019] (in the expanding regular case).

We obtain our results by generalizing the framework of [JST, 2021] based on weak regularity decomposition for expanding hypergraphs. This framework was originally designed for binary k-XOR with the goal of providing near-linear time decoder for explicit binary codes, near the GV bound, from the breakthrough work of Ta-Shma [STOC, 2017]. The explicit families of codes over prime $\mathbb{F}_q$ are based on suitable instatiations of the Jalan–Moshkovitz (Abelian) generalization of Ta-Shma's distance amplification procedure.

*This material is based upon work supported by the NSF grant CCF-1900460. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

Contents

1 Introduction

Codes over small alphabet sizes have attracted a lot of effort in coding theory [GRS19]. There is now a vast theory about them, but important mysteries remain. One very natural alphabet is the binary alphabet, which has a myriad of uses and applications. However, it also comes with an important limitation, namely, a family of good binary codes cannot[1] surpass distance $1/2$. By using a q-ary alphabet, a family of good codes can approach distance $1 - 1/q$ but not surpass it. This makes the use of q-ary codes a necessity whenever larger distances are needed. Working towards explicit and efficiently decodable codes with optimal trade-offs between rate and distance has been a challenging but fruitful guiding goal in coding theory.

In the large distance case, namely, distances are of the form $1 - 1/q - \varepsilon$ for small values of $\varepsilon > 0$, the Gilbert–Varshamov (GV) bound [Gil52, Var57] asserts that rate $\Omega_q(\varepsilon^2)$ is achievable whereas the q-ary version of McEliece, Rodemich, Rumsey and Welch (MRRW) [MRRW77] gives an impossibility upper of $O_q(\varepsilon^2 \log(1/\varepsilon))$. This means that the GV bound is nearly optimal in this regime of constant alphabet size q and large distance. To the best of our knowledge, in this regime, (prior to this work) no explicit and efficiently decodable families of q-ary codes near the GV bound were known for any $q \geq 3$.

Two widely used approaches in the construction of q-ary codes for small q are based on code concatenation [For66] and on algebraic geometry (AG) constructions [Sti08, TVN07]. Using code concatenation, it is possible to obtain explicit constructions achieving the sub-optimal Zyablov bound trade-off between rate and distance, which gives a rate of $\Omega_q(\varepsilon^3)$. Some explicit families of AG codes are celebrated for beating the GV bound in some specific parameter regimes, e.g., the seminal work of Tsfasman, Vlădut and Zink[2] [TVZ82] or the (non-linear) construction of Elkies [Elk01]. This surprising phenomenon of explicit AG codes beating random codes cannot happen in a major way in the large distance and constant alphabet regime since the GV bound is nearly optimal. Furthermore, known explicit constructions of linear AG codes are far from the GV bound for large distances and constant q. Another drawback of several explicit families of good AG codes is that known decoders can take much longer than linear time in the blocklength [NW19].

On a more combinatorial side, in a breakthrough work using expander graphs, Ta-Shma [TS17] gave the first explicit construction of binary codes of distance $1/2 - \varepsilon$ and rate $\Omega(\varepsilon^{2+o(1)})$, namely, near the Gilbert–Varshamov bound. A polynomial time decoder for these binary codes was first given in [JQST20] followed by a near-linear time decoder in [JST21]. Subsequently, Jalan and Moshkovitz [JM21] extended Ta-Shma's analysis [TS17] to handle (in particular) codes over larger alphabets[3]. Suitable instantiations of [JM21] imply explicit codes over prime $\mathbb{F}_q$ of distance $1 - 1/q - \varepsilon$ with rate $\Omega_q(\varepsilon^{2+o_q(1)})$, namely, again near the (q-ary) GV bound for constant q.

Motivated by the above situation, we design a near-linear time decoder for explicit families of q-ary codes of distance $(1 - 1/q)(1 - \varepsilon)$ and rate $\Omega(\varepsilon^{2+o_q(1)})$ for any constant prime q, namely, near the GV bound in the large distance regime. More precisely, our main result is as follows (answering a question from [JM21]).

[1] This is a consequence of the Plotkin bound.

[2] More precisely, the TVZ bound [TVZ82] establishes a rate of $r \geq 1 - \delta - 1/(\sqrt{q} - 1)$ with respect to the relative distance δ.

[3] More precisely, [JM21] analyzed the (scalar) Abelian case of Ta-Shma's amplification.

Theorem 1.1 (Main I - Near-linear Time Unique Decoding over $\mathbb{F}_q$)**.** *Let q be a prime. For every $\varepsilon > 0$ sufficiently small, there are explicit linear Ta-Shma codes $\mathcal{C}_{N,q,\varepsilon} \subseteq \mathbb{F}_q^N$ for infinitely many values $N \in \mathbb{N}$ with*

(i) distance at least $(1-1/q)(1-\varepsilon)$ (actually ε-balanced),

(ii) rate $\Omega_q(\varepsilon^{2+\alpha})$ where $\alpha = O(1/(\log_2(1/\varepsilon))^{1/6})$, and

(iii) an $r(q/\varepsilon) \cdot \tilde{O}(N)$ time randomized unique decoding algorithm that decodes within radius $((1-1/q)(1-\varepsilon))/2$,

where $r(x) = \exp(\exp(\mathrm{poly}(x)))$.

In fact, we actually prove the following stronger *list* decoding result.

Theorem 1.2 (Near-linear time List Decoding over $\mathbb{F}_q$)**.** *Let q be a prime. For every $\varepsilon > 0$ sufficiently small, there are explicit binary linear Ta-Shma codes $\mathcal{C}_{N,q,\varepsilon} \subseteq \mathbb{F}_q^N$ for infinitely many values $N \in \mathbb{N}$ with*

(i) distance at least $(1-1/q)(1-\varepsilon)$ (actually ε-balanced),

(ii) rate $\Omega_q(\varepsilon^{2+\alpha})$ where $\alpha = O(1/(\log_2(1/\varepsilon))^{1/6})$, and

(iii) an $r(q/\varepsilon) \cdot \tilde{O}(N)$ time randomized list decoding algorithm that decodes within radius $1 - 1/q - 2^{-\Theta_q((\log_2(1/\varepsilon))^{1/6})}$ and works with high probability,

where $r(x) = \exp(\exp(\mathrm{poly}(x)))$.

We obtain our results by building on and extending the *binary* decoding framework in [JST21]. This framework is based on a generalization of the weak regularity decomposition to (sparse) *expanding* hypergraphs that generalizes the seminal work of Frieze and Kannan [FK96]. The weak regularity decomposition of [JST21] was then used to approximate *expanding* k-XOR instances naturally arising in the decoding of binary Ta-Shma's codes [TS17]. Similarly, constraint satisfaction problems (CSPs) will play a key role in our decoder. Here, we also take the opportunity to investigate *expanding* CSPs more broadly.

An instance of a k-CSP is given by a k-uniform (ordered) constraint hypergraph $W \subseteq [n]^k$, where each vertex is associated with a variable taking values in an alphabet of size q and each edge is associated with a constraint involving the variables of its vertices. While even approximating a CSP is NP-hard in general, suitable notions of expansion of the constraint hypergraph allow for efficient approximation algorithms as in [JST21]. One such notion is *splittability* [AJT19] (*cf.*, Definition 4.3). Roughly speaking, a τ-splittable collection of tuples for some $\tau \in (0,1]$ is the higher-order analogue of the second largest singular value of the normalized adjacency matrix of a graph (the smaller the τ the more expanding is the collection). Approximating expanding k-CSPs is at the core of some decoding algorithms for expander based constructions of codes [DHK+19, AJQ+20, JQST20, JST21, BD22].

As mentioned above, approximating expanding k-CSPs will be again at the core of our extension of [JST21] to more general constraints over larger alphabets. Our new q-ary decoder will need to handle instances of linear equations over the alphabet $\mathbb{Z}_q$, where each equation involves a sum of k variables. This kind of k-CSP is commonly denoted k-LIN

over alphabet $\mathbb{Z}_q$. We will see that the special algebraic structure of these linear constraints will allow to obtain some improved parameter trade-offs, which will be explored in the decoding application. More precisely, the expansion (splittability) parameter τ will have no dependence on alphabet size q and only a polynomial dependence on the arity[4] k, and this allows us to obtain better approximation guarantees. Our second result follows.

Theorem 1.3 (Main II)**.** *Let $\mathcal{I}$ be an instance of MAX k-LIN$_q$ on n variables with alphabet $\mathbb{Z}_q$ and constraints supported on a regular[5] collection of tuples $W \subseteq [n]^k$. If W is τ-splittable with $\tau \leq \tau_0(k,\delta) := \mathrm{poly}(\delta/k)$, then we can compute an assignment satisfying $\mathsf{OPT} - \delta$ in time $r(q/\tau_0) \cdot \widetilde{O}(|W| + n)$, where $r(x) = \exp(\exp(\mathrm{poly}(x)))$.*

We show that this phenomenon of no dependence of the expansion on the alphabet size q and only polynomial dependence on arity k also occurs for linear equations over a general finite groups $\mathfrak{G}$. Similarly, this leads to better approximation guarantees. To actually implement and obtain this advantage, we will design a new matrix version of the weak regularity decomposition for expanding hypergraphs. Our third result follows.

Theorem 1.4 (Main III)**.** *Let $\mathcal{I}$ be an instance of MAX k-LIN$_{\mathfrak{G}}$ on n variables with alphabet a finite group $\mathfrak{G}$ and constraints supported on a regular collection of tuples $W \subseteq [n]^k$. If W is τ-splittable with $\tau \leq \tau_0(k,\delta) := \mathrm{poly}(\delta/k)$, then we can compute an assignment satisfying $\mathsf{OPT} - \delta$ in time $O_{|\mathfrak{G}|,k,\delta}(1) \cdot \mathrm{poly}(|W| + n)$.*

Remark 1.5. *In Theorem 1.4, we did not attempt to make the running time near-linear in the number of constraints and variables, but it is plausible that it can be done.*

We find intriguing this interplay between the type of constraint used in the CSP and the expansion requirement for a given approximation. A natural question is to investigate this interplay for more general constraint types.

In this work, we also investigate how fast we can approximate expanding k-CSPs over q-ary alphabet without making any assumptions on the constraints. We show that k-CSPs can be approximated in near-linear time in the number of constraints and variables, assuming k and q are constants, and provided the constraint hypergraph is sufficiently expanding (splittable Definition 4.3). An important caveat of this general case is that the expansion requirements will now depend on both the alphabet size q and arity k in an exponential way (of the form $q^{-O(k)}$).

Theorem 1.6. *Let $\mathcal{I}$ be an instance of MAX k-CSP on n variables with alphabet $[q]$ and constraints supported on a regular collection of tuples $W \subseteq [n]^k$. If W is τ-splittable with $\tau \leq \tau_0(k,q,\delta) := \mathrm{poly}(\delta/(kq^k))$, then we can compute an assignment satisfying $\mathsf{OPT} - \delta$ in time $r(kq/\delta) \cdot \widetilde{O}(|W| + n)$, where $r(x) = \exp(\exp(\exp(\mathrm{poly}(x))))$.*

We obtain the above result via a reduction to the "binary" weak regularity in [JST21] in a somewhat similar fashion to [FK98]. Even though it is not hard to make this connection, we think it is worth stating it since this result may be more broadly applicable. Moreover, for fixed arity k and alphabet size q, this improves the running time in the expanding regime of the Sum-of-Squares based algorithm in [AJT19] and also the expanding regime[6] of earlier results 2-CSPs [BRS11, GS11, GS12, OGT15].

[4] In the binary case of [JST21], it was also possible to have a polynomial dependence on the arity k.

[5] See Definition 4.1 for the definition of regular. This is an analog to tuples of a graph being d-vertex regular.

[6] We point out these approaches also consider when the expansion is defective (low threshold rank case). Since we are interested in near-linear running time, we need to focus on the expanding case.

For comparison, we recall the expanding regime[7] of [AJT19] below.

Theorem 1.7 (Sum-of-Squares [AJT19])**.** *Let $\mathfrak{I}$ be an instance of MAX k-CSP on n variables with alphabet $[q]$ and constraints supported $W \subseteq [n]^k$. If W is τ-splittable with $\tau \leq \tau_0(k, q, \delta) := \mathrm{poly}(\delta/k) \cdot q^{-k}$, then we can compute an assignment satisfying $\mathsf{OPT} - \delta$ in time $n^{\mathrm{poly}(1/\tau_0)}$.*

Remark 1.8. *In the new theorem above, we do not attempt to optimize the function $r(x)$.*

Related Work: As we mentioned above, our work is an extension of the *binary* framework of [JST21]. This framework was designed for approximating expanding k-XOR and to give a near-liner time decoding algorithm for the explicit binary codes of Ta-Shma [TS17], near the GV bound. The first polynomial time decoder for these codes was given in [JQST20] using the Sum-of-Squares semi-definite programming hierarchy and its running time, albeit polynomial, is very far from near-linear in the blocklength.

AG codes are widely used in the study of explicit constructions over constant q-ary alphabets. Some of these constructions achieve very competitive parameter trade-offs (e.g., rate versus distance) if not the best known in several cases. However, explicit and efficiently decodable codes near GV bound for large distances,i.e., $1 - 1/q - \varepsilon$, and constant alphabet size were not known prior to this work. In fact, the first explicit construction only appeared in the breakthrough work of [TS17] for binary codes using more combinatorial expander based techniques. This absence of explicit construction near the GV bound in this regime means that much is yet to be discovered about this case. We view our near-linear time decoder of prime q-ary codes in this regime as not only reaching previously unattained parameter regimes with an explicit construction, but also offering a more combinatorial perspective among a wealthy of algebraic techniques.

For *non-explicit* families of codes approaching the GV bound, much more is known. Random linear codes achieve this bound, but their decoding is believed to be computationally hard. It is possible to construct more structured ensembles of random codes that allow for efficient decoding in this regime. We have the non-explicit classical Goppa codes. Another important technique is based on Thommesen's [Tho83] technique of concatenation with random inner codes. These Thommesen based ensembles can sometimes approach the GV bound and also allow for efficient decoding [GI04, GKO$^+$17, HRW17, KRRZ$^+$21] and even near-linear time decoding [HRW17, KRRZ$^+$21].

More recently, Blanc and Doron [BD22] used the framework in [JST21] to decode explicit binary codes near the GV bound with improved parameters, where they obtain a polynomial improvement on the $o(1)$ error term of the rate $\Omega(\varepsilon^{2+o(1)})$ (the α in Theorem 1.1) and also put forward some interesting conjectures towards further improving the rate. It is plausible that their improvement also applies here for q-ary alphabets.

In the constant alphabet case, a different parameter regime that has received much attention is the near-capacity regime [GRZ22, GX13, HRW17, KRRZ$^+$21] of list decoding from radius $1 - r - \varepsilon$ with rate r for small values of $\varepsilon > 0$. This regime can only occur when the alphabet size q is a function of ε. Note that our near GV bound regime is the opposite, we have a fixed constant q and we can take ε arbitrarily small (smaller than some function of q).

[7] Using the improved analysis of swap walks by Dikstein and Dinur [DD19].

2 Proof Strategy

We will now describe our contributions in more detail. Our algorithmic results will be based on extensions of the *binary* weak regularity framework of [JST21]. Roughly speaking, this framework being a *"low level"* framework gives fine control over its components leading to a near-linear time decoder for Ta-Shma's codes [TS17] over $\mathbb{F}_2$. This same low level structure means that extensions may require suitable generalizations in several of these components as well technical work to implement them. The extensions to handle codes over prime q-ary alphabet and a matrix version of weak regularity will be no exception.

First, we will recall the weak regularity decomposition of Frieze and Kannan [FK98] in a more analytic form [TTV09]. We will also first consider its *existential* form and later discuss its algorithmic form. Our setup will be as follows. Let $W \subseteq [n]^k$ be a collection of tuples endowed with the uniform probability measure μ_k. Suppose that we have a function $g\colon W \to \mathbb{C}$ that we want to approximate using a simpler approximating function, which will be made precise below. Further suppose that the quality of approximation will be measured with respect to correlations with a class of test functions $\mathcal{F}$. Given some desired approximation error $\delta > 0$, the goal will be to find a *"simple"* approximator $h \approx g$ such that

$$\max_{f \in \mathcal{F}} \left| \langle g - h, f \rangle_{\mu_k} \right| \leq \delta .$$

As an *existential* result, it is well-known that an h of the form $h = \sum_{\ell=1}^{p} c_\ell \cdot f_\ell$ always exists, where c_ℓ's are scalars and the f_ℓ's are functions belonging to $\mathcal{F}$. Furthermore, the number of test functions p is small being at most[8] $O(1/\delta^2)$. This means that h is indeed *"simple"* since it is the sum of a small number of test functions, so h is almost as complex as the test functions it needs to fool.

To motivate the generalizations in the weak regularity framework, we will start the discussion of the important case of linear equations over $\mathbb{Z}_q$ as a motivating example. As mentioned above, approximating k-LIN over $\mathbb{Z}_q$ will be crucial in the near-linear time decoding algorithm for prime q-ary alphabets. For us, an instance $\mathfrak{I}$ of k-LIN is given by a system of linear equations[9]

$$x_{i_1} + \cdots + x_{i_k} \equiv r_w \pmod{q} \qquad \forall\, w = (i_1, \ldots, i_k) \in W, \tag{1}$$

where $(r_w)_{w \in W} \in \mathbb{Z}_q^W$ are given RHS coefficients. We will need to model this problem in a way that is amenable to the weak regularity approach. We will also take advantage of the algebraic structure of the constraints to avoid any dependence of the alphabet size q and to have only a mild dependence on the arity k in the expansion the framework will require from W.

"Global" Approximation of Dirac Delta Functions: An elementary property of Fourier analysis over $\mathbb{Z}_q$ is that the Dirac delta function $x \mapsto \mathbf{1}_{[x=y]}$ admits a simple but extremely handy Fourier decomposition which we now recall. Let $\omega = \exp(2\pi\sqrt{-1}/q)$. Using orthogonality of characters, we have

$$\mathbf{1}_{[x=y]} = \mathop{\mathbb{E}}_{a \in \mathbb{Z}_q} \left[\omega^{a(x-y)} \right] .$$

[8] The ℓ_1-norm of the coefficients is "small", i.e., $\sum_{\ell=1}^{p} |c_\ell|$.

[9] The coefficients of the variables are always taken to be 1 here.

Suppose we have an assignment $b \in \mathbb{Z}_q^n$ to the variables of our system of linear equations $\mathfrak{I}$. Then, the fraction of satisfied constraints, which we denote by $\mathsf{val}(\mathfrak{I}, b)$ and refer as the value of this assignment, can be expressed as

$$\mathsf{val}(\mathfrak{I}, b) := \mathop{\mathbb{E}}_{w=(i_1,\ldots,i_k)\sim\mu_k}\left[\mathbf{1}_{[b_{i_1}+\cdots+b_{i_k}\equiv r_w]}\right] = \mathop{\mathbb{E}}_{w=(i_1,\ldots,i_k)\sim\mu_k}\left[\mathop{\mathbb{E}}_{a\in\mathbb{Z}_q}\left[\omega^{a(b_{i_1}+\cdots+b_{i_k}-r_w)}\right]\right].$$

This suggests defining q functions one for each $a \in \mathbb{Z}_q$ of the form $g_a \colon W \to \mathbb{C}$ as $g_a(w) := \omega^{a\cdot b_w}$, the *"harmonic"* components. We also endow the space $\mathbb{C}^W$ with the inner product defined by the measure μ_k on W. We will need some additional notation. For $b \in \mathbb{Z}_q^n$, we define the function $\chi_{b,a}$ on $[n]$ as $\chi_{b,a}(i) = \omega^{a\cdot b_i}$. We can now rexpress $\mathsf{val}(\mathfrak{I}, b)$ in terms of its harmonic components as

$$\begin{aligned}
\mathsf{val}(\mathfrak{I}, b) &= \mathop{\mathbb{E}}_{w=(i_1,\ldots,i_k)\sim\mu_k}\left[\mathop{\mathbb{E}}_{a\in\mathbb{Z}_q}\left[\omega^{a(b_{i_1}+\cdots+b_{i_k}-r_w)}\right]\right] \\
&= \mathop{\mathbb{E}}_{a\in\mathbb{Z}_q}\left[\mathop{\mathbb{E}}_{w=(i_1,\ldots,i_k)\sim\mu_k}\left[\omega^{-a\cdot r_w}\cdot\omega^{a(b_{i_1}+\cdots+b_{i_k})}\right]\right] \\
&= \mathop{\mathbb{E}}_{a\in\mathbb{Z}_q}\left[\left\langle g_a, \underbrace{\chi_{b,a}\otimes\cdots\otimes\chi_{b,a}}_{k}\right\rangle_{\mu_k}\right] \\
&= \mathop{\mathbb{E}}_{a\in\mathbb{Z}_q}\left[\left\langle g_a, (\chi_{b,a})^{\otimes k}\right\rangle_{\mu_k}\right].
\end{aligned}$$

We can now try to further approximate each g_a using a simpler function h_a that behaves similarly to g_a with respect to functions of the form $f_{b,a} = \chi_{b,a}\otimes\cdots\otimes\chi_{b,a}$ as in the inner product above. We can view functions of form $f_{b,a}$ as tests with respect to which g_a and its simpler approximator have similar correlations. This means that we can model the problem in way amenable to the existential weak regularity framework. For each a, we will consider a (slightly) more general class of test functions $\mathrm{CUT}^{\otimes k}_{\omega,q,a}$ defined as follows

$$\mathrm{CUT}^{\otimes k}_{\omega,q,a} := \{\chi_{b^{(1)},a}\otimes\cdots\otimes\chi_{b^{(k)},a} \mid b^{(1)},\ldots,b^{(k)} \subseteq \mathbb{Z}_q^n\}.$$

A simple yet useful remark is that if we can find a decomposition fooling a larger class of test functions, this would suffice since, in particular, it fools the initial class of test.

Suppose that for some $\delta \in (0,1)$ we can find a δ-approximation $h_a = \sum_{\ell=1}^{p_a} c_{a,\ell}\cdot\chi_{b^{(a,\ell,1)},a}\otimes\cdots\otimes\chi_{b^{(a,\ell,k)},a}$ to g_a with respect to a class of test functions, i.e.,

$$\max_{f\in\mathrm{CUT}^{\otimes k}_{\omega,q,a}}\left|\langle g_a - h_a, f\rangle_{\mu_k}\right| \le \delta.$$

By replacing g_a with h_a in the computation of $\mathsf{val}(\mathfrak{I}, b)$ above, we obtain[10]

$$\mathsf{val}(\mathfrak{I}, b) = \mathop{\mathbb{E}}_{a\in\mathbb{Z}_q}\left[\left\langle g_a, (\chi_{b,a})^{\otimes k}\right\rangle\right] = \mathop{\mathbb{E}}_{a\in\mathbb{Z}_q}\left[\left\langle h_a, (\chi_{b,a})^{\otimes k}\right\rangle\right] \pm \delta.$$

We will explain how to algorithmically find h_a in near-linear time later. Now, we will argue why having access to weak regularity decomposition greatly simplifies our task of approximating $\mathsf{val}(\mathfrak{I}, b)$ and also later while decoding q-ary codes.

[10]For scalars x, y (real or complex) and real $\delta \in \mathbb{R}^+$, we use the notation $x = y \pm \delta$ if $|x - y| \le \delta$.

We can simplify the above equation for $\mathsf{val}(\mathfrak{I}, b)$ even further using the assumed expansion (splittability) of W. A suitable version of the expander mixing lemma allows us to pass from the measure μ_k to the product measure $\mu_1^{\otimes k}$, where μ_1 is the uniform measure on $[n]$. More precisely, we can show that if W is sufficiently expanding (depending on δ), then

$$\begin{aligned}\mathsf{val}(\mathfrak{I}, b) = \mathop{\mathbb{E}}_{a\in\mathbb{Z}_q}\left[\left\langle h_a, (\chi_{b,a})^{\otimes k}\right\rangle_{\mu_k}\right] \pm \delta &= \mathop{\mathbb{E}}_{a\in\mathbb{Z}_q}\left[\left\langle h_a, (\chi_{b,a})^{\otimes k}\right\rangle_{\mu_1^{\otimes k}}\right] \pm 2\delta \\ &= \mathop{\mathbb{E}}_{a\in\mathbb{Z}_q}\left[\sum_{\ell=1}^{p_a} c_{a,\ell}\cdot\prod_{j=1}^{k}\left\langle \chi_{b^{(a,\ell,1)},a}, \chi_{b,a}\right\rangle_{\mu_1}\right] \pm 2\delta\,.\end{aligned}$$

The *low complexity* of the approximator h_a will allow us to simplify the search for an approximately optimal assignment $b \in \mathbb{Z}_q^n$. The expression above reveals that we only need to know the values of

$$\left\{\left\langle \chi_{b^{(a,\ell,j)},a}, \chi_{b,a}\right\rangle_{\mu_1}\right\}_{a\in\mathbb{Z}_q,\ell\in[p_a],j\in[k]}.$$

Luckily, algorithmically, there will be only $O(qk^3/\delta^2)$ such numbers (no dependence on n and only slightly more than the $O(qk/\delta^2)$ from the existential result). Using brute-force search, it is possible to find sufficiently fine and (close to valid) approximations for these numbers.

To make the entire process efficient and near-linear time we still need to say how to find the functions h_a's in near-linear time. As in [JST21], we will reduce the problem of finding a weak regularity decomposition with respect to a class of k-tensors, in this case the class $\mathrm{CUT}^{\otimes k}_{\omega,q,a}$, to multiple applications of the 2-tensor case (in a sparse regime). To execute this process in near-linear, we will again use the expansion of W to conveniently move to easier to handle product measures (as above). This involves finding a constant factor approximation for the following expression

$$\max_{x,y\in\mathbb{Z}_q^n}\left|\sum_{i,j=1}^{n}\mathsf{A}_{i,j}\cdot\omega^{a\cdot x_i}\cdot\omega^{a\cdot y_j}\right|, \tag{2}$$

This kind of optimization is known as the *Grothendieck problem* and, in this case, it is for roots of unity going beyond the ± 1 case of Alon and Naor [AN04]. In [SZY07], So, Zhang and Ye considered a more restricted version of this problem (with positive semi-definite matrices) known as the *little* Grothendieck problem. We will extend their analysis to the Grothendieck problem building on some ingredients present in their proof. In our application, the matrices A will be sparse with $m \approx n$ non-zero entries and to achieve a near-linear time we will need to find an (additive) approximation to the Grothendieck problem in time $\widetilde{O}(m)$ of Eq. (2). This can be done using the fast SDP solver of Arora and Kale [AK07].

We now explain how the above weak regularity decomposition can be used in decoding of the expander based construction of Ta-Shma's codes [TS17]. We will see that the decoding problem can be naturally phrased as a k-LIN instance over $\mathbb{Z}_q$, which is a natural q-ary extension of the k-XOR over $\mathbb{Z}_2$ from [AJQ+20, JQST20, JST21]. First, we briefly describe Ta-Shma's code construction over alphabet $\mathbb{F}_q$, with q prime, as analyzed[11] in [JM21].

[11] In [JM21], they considered the more general (scalar) Abelian case.

The idea is to start with a good base code $\mathcal{C}_0 \subseteq \mathbb{F}_q^n$ and to use a carefully constructed collection of tuples $W \subseteq [n]^k$ to amplify its distance via the direct-sum encoding. For any $z \in \mathbb{F}_q^n$, recall that its direct-sum encoding is a new word denoted $y = \mathrm{dsum}_W(z)$ in $\mathbb{F}_q^W$ and defined as

$$y_{(i_1,\ldots,i_k)} = z_{i_1} + \cdots + z_{i_k} \pmod q \qquad \forall\, (i_1,\ldots,i_k) \in W\,.$$

The direct-sum code $\mathcal{C} = \mathrm{dsum}_W(\mathcal{C}_0)$ is defined as $\mathcal{C} = \{\mathrm{dsum}_W(z) \mid z \in \mathcal{C}_0\}$. Note the similarity of the above equation and the system of linear equations from Eq. (1). In the decoding task, we are given a (possibly) corrupted version of $\widetilde{y}$ of some codeword $y = \mathrm{dsum}_W(z) \in \mathcal{C}$, with $z \in \mathcal{C}_0$. We can view $\widetilde{y}$ as defining the RHS coefficients of an instance of k-LIN, namely, $r_W = \widetilde{y}_w$.

Having an instance of k-LIN over $\mathbb{Z}_q$, we can now use weak regularity as described above. For each $a \in \mathbb{Z}_q$, let g_a be the *harmonic* component associated with RHS vector $\widetilde{y}$ (as above). Similarly, we find a weak regularity approximation h_a for each function g_a.

If the distance $\Delta(\widetilde{y}, \mathrm{dsum}_W(z)) \leq (1 - 1/q)(1 - \beta)$ is not too large, we will be able to deduce that some harmonic function h_a "captures" the structure of the codeword z in the following sense. Set $\mathcal{R} = \{\omega^{a \cdot a'} \mid a' \in \mathbb{Z}_q\}$ and let $f_1, \ldots, f_r \colon [n] \to \mathcal{R}$ be the functions appearing in the decomposition of h_a. For each tuple $(y_1, \ldots, y_r) \in \mathcal{R}^r$, we can consider the set

$$\{x \in [n] \mid f_1(x) = y_1, \ldots, f_r(x) = y_r\}.$$

These sets partition[12] the space $[n]$, and we can show that z is approximately constant in most of these parts. In this sense, the low complexity structure of h_a captures the structure of the codeword z. In this last argument, we use that assumption that q is prime[13].

The case of k-LIN over a finite group will also allow for a weak regularity decomposition in a similar spirit as above, where scalar Fourier characters are replaced by larger dimensional representations and "global" approximation of Dirac delta functions are performed. Extending the weak regularity framework to this case will require considering matrix valued functions. The way we model this case is done in Section 3.4 and it uses very elementary properties of representation theory. This case again exhibits an interesting interplay between the type of constraints and the requirement on expansion. (The reader who is only interested in decoding can safely ignore this extension and focus on the $\mathbb{Z}_q$ case.)

3 Constraint Types and Alphabets

We explore the role of different types of constraints and corresponding alphabets going beyond the binary k-XOR considered in [JST21]. For the special case of linear equations over $\mathbb{Z}_q$ or over an arbitrary finite group $\mathfrak{G}$, we will explore the special structure of the constraints and obtain results with improved parameters.

[12] Possibly with empty parts.

[13] So that all non-trivial roots of unity are primitive roots. It is plausiable that this restriction is not necessary.

3.1 General CSPs via the Binary Regularity

We will prove our first result for approximating a general expanding k-CSPs over a q-ary alphabet in near-linear time. We obtain this result using the *binary* near-linear time weak regularity decomposition from [JST21] in a similar way that Frieze and Kannan modeled k-CSPs [FK98] using regularity. We formalize this (relatively simple) connection since we believe this result may be of independent interest and may find applications elsewhere. Moreover, it also improves the running time of [AJT19] to near-linear time, for fixed k and q, while offering a different approach to approximating general expanding k-CSPs which could be simpler than their Sum-of-Squares based algorithm. We now restate and proceed to prove this result.

Theorem 1.6. *Let $\mathfrak{I}$ be an instance of MAX k-CSP on n variables with alphabet $[q]$ and constraints supported on a regular collection of tuples $W \subseteq [n]^k$. If W is τ-splittable with $\tau \leq \tau_0(k,q,\delta) := \mathrm{poly}(\delta/(kq^k))$, then we can compute an assignment satisfying $\mathsf{OPT} - \delta$ in time $r(kq/\delta) \cdot \widetilde{O}(|W| + n)$, where $r(x) = \exp(\exp(\exp(\mathrm{poly}(x))))$.*

We will find a weak regularity decomposition with respect to $0/1$ valued test functions $\mathcal{F} = \mathrm{CUT}^{\otimes k}$ where

$$\mathrm{CUT}^{\otimes k} := \{\pm \mathbf{1}_{S_1} \otimes \cdots \otimes \mathbf{1}_{S_k} \mid S_1, \ldots, S_k \subseteq [n]\}\,.$$

The near-linear weak regularity decomposition of [JST21], which we recall below, can handle this class of functions.

Theorem 3.1 (Efficient Weak Regularity from [JST21])**.** *Let $W \subseteq [n]^k$ be a τ-splittable collection of tuples. Suppose $\mathcal{F}$ is one of $\mathrm{CUT}^{\otimes k}$, $\mathrm{CUT}_{\pm}^{\otimes k}$. Let $\mathcal{R}$ be the domain of the functions in $\mathcal{F}$, when $k = 1$. Let $g \in \mathcal{R}^{W[1]^k}$ be supported on W with $\|g\|_{\mu_k} \leq 1$. For every $\delta > 0$, if $\tau \leq \delta^2/(k^3 \cdot 2^{20})$, then we can find $h = \sum_{\ell=1}^{p} c_\ell \cdot f_\ell$ with $p = O(k^2/\delta^2)$, $c_1, \ldots, c_p \in \mathbb{R}$ and functions $f_1, \ldots, f_p \in \mathcal{F}$, such that $\|h\|_{\mu_1^{\otimes k}} \leq 2$, $\sum_{\ell=1}^{p} |c_\ell| = O(k/\delta)$ and h is a good approximator to g in the following sense*

$$\max_{f \in \mathcal{F}} \left| \left\langle g - \left(\frac{d}{n}\right)^{k-1} h, f \right\rangle \right| \leq \delta \cdot |W|\,,$$

where the inner product is over the counting measure on $W[1]^k$. Furthermore, h can be found in $\widetilde{O}(2^{2^{\widetilde{O}(k^2/\delta^2)}} \cdot |W|)$ time.

Having access to a weak regularity decomposition as above makes the task of approximating the value of a CSP instance relatively simple, as we now describe. This is a common feature of weak regularity based arguments,e.g., [FK98, OGT15]. Here, we consider both arbitrary arity k and arbitrary alphabet size q.

We will first need some notation. Let $\alpha \in [q]^k$ and define $W_\alpha = \{w \in W \mid P_w(\alpha) = 1\}$ to be the set of tuples whose predicates P_w are satisfied by on the input α. Let $\mathcal{A}(\mathfrak{I}) = \{\alpha \in [q]^k \mid W_\alpha \neq \emptyset\}$ be the set of satisfying inputs of at least one predicate of $\mathfrak{I}$.

We will use the following claim which relates the value of an assignment to the structure of the weak regularity decomposition.

Claim 3.2. *Suppose that for every $\alpha \in [q]^k$, we have a weak regularity decomposition h_α, from Theorem 3.1, of the indicator $\mathbf{1}_{W(\alpha)}$ with error parameter $\delta > 0$ and with respect to the test class $\mathrm{CUT}^{\otimes k}$. Let $b \in [q]^n$ (viewed as an assignment), which induces a partition $T_1 \sqcup \cdots \sqcup T_q$ of $[n]$. Then,*

$$\mathsf{val}(\mathfrak{I}, b) \;=\; \sum_{\alpha \in \mathcal{A}} \sum_{\ell=1}^{p_\alpha} c_{\alpha,\ell} \frac{\left|S_1^{\alpha,\ell} \cap T_{\alpha_1}\right|}{n} \cdots \frac{\left|S_k^{\alpha,\ell} \cap T_{\alpha_k}\right|}{n} \;\pm\; \delta \cdot |\mathcal{A}(\mathfrak{I})| \;.$$

Proof. Let $\mathcal{A} = \mathcal{A}(\mathfrak{I})$. The value of this assignment is

$$\begin{aligned}
\mathsf{val}(\mathfrak{I}, b) \;&=\; \sum_{\alpha \in \mathcal{A}} \left\langle \mathbf{1}_{W_\alpha}, \mathbf{1}_{T_{\alpha_1}} \otimes \cdots \otimes \mathbf{1}_{T_{\alpha_k}} \right\rangle_{\mu_k} \\
&=\; \frac{1}{|W|} \sum_{\alpha \in \mathcal{A}} \left\langle \left(\frac{d}{n}\right)^{k-1} h_\alpha, \mathbf{1}_{T_{\alpha_1}} \otimes \cdots \otimes \mathbf{1}_{T_{\alpha_k}} \right\rangle \;\pm\; \delta \cdot |\mathcal{A}| \\
&=\; \frac{1}{|W|} \sum_{\alpha \in \mathcal{A}} \left\langle \left(\frac{d}{n}\right)^{k-1} \sum_{\ell=1}^{p_\alpha} c_{\alpha,\ell} \cdot \mathbf{1}_{S_1^{\alpha,\ell}} \otimes \cdots \otimes \mathbf{1}_{S_k^{\alpha,\ell}}, \mathbf{1}_{T_{\alpha_1}} \otimes \cdots \otimes \mathbf{1}_{T_{\alpha_k}} \right\rangle \;\pm\; \delta \cdot |\mathcal{A}| \\
&=\; \frac{1}{n^k} \sum_{\alpha \in \mathcal{A}} \sum_{\ell=1}^{p_\alpha} c_{\alpha,\ell} \cdot \left\langle \mathbf{1}_{S_1^{\alpha,\ell}} \otimes \cdots \otimes \mathbf{1}_{S_k^{\alpha,\ell}}, \mathbf{1}_{T_{\alpha_1}} \otimes \cdots \otimes \mathbf{1}_{T_{\alpha_k}} \right\rangle \;\pm\; \delta \cdot |\mathcal{A}| \\
&=\; \sum_{\alpha \in \mathcal{A}} \sum_{\ell=1}^{p_\alpha} c_{\alpha,\ell} \cdot \left\langle \mathbf{1}_{S_1^{\alpha,\ell}}, \mathbf{1}_{T_{\alpha_1}} \right\rangle_{\mu_1} \cdots \left\langle \mathbf{1}_{S_k^{\alpha,\ell}}, \mathbf{1}_{T_{\alpha_k}} \right\rangle_{\mu_1} \;\pm\; \delta \cdot |\mathcal{A}| \\
&=\; \sum_{\alpha \in \mathcal{A}} \sum_{\ell=1}^{p_\alpha} c_{\alpha,\ell} \frac{\left|S_1^{\alpha,\ell} \cap T_{\alpha_1}\right|}{n} \cdots \frac{\left|S_k^{\alpha,\ell} \cap T_{\alpha_k}\right|}{n} \;\pm\; \delta \cdot |\mathcal{A}| \;,
\end{aligned}$$

concluding the proof. ∎

Proof of Theorem 1.6. Let $\mathfrak{I}$ be an instance of a k-CSP over alphabet $[q]$ supported on a collection of tuples $W \subseteq [n]^k$ and with predicates $(P_w \colon [q]^k \to \{0,1\})_{w \in W}$.

For each $\alpha \in \mathcal{A}(\mathfrak{I})$, we apply the weak regularity decomposition of Theorem 3.1 to the function $\mathbf{1}_{W_\alpha}$ with error parameter $\delta > 0$ and test class $\mathcal{F} = \mathrm{CUT}^{\otimes k}$. This gives an approximation $h_\alpha = \sum_{\ell=1}^{p_\alpha} c_{\alpha,\ell} \cdot \mathbf{1}_{S_1^{\alpha,\ell}} \otimes \cdots \otimes \mathbf{1}_{S_k^{\alpha,\ell}}$.

A crucial property is that instead of having to know an assignment $b \in [q]^n$, represented as a partition $T_1 \sqcup \cdots \sqcup T_q = [n]$, it is enough to know the values of the following inner products

$$\left\{ \left\langle \mathbf{1}_{S_j^{\alpha,\ell}}, \mathbf{1}_{T_{\alpha_j}} \right\rangle_{\mu_1} \right\}_{\alpha \in \mathcal{A}(\mathfrak{I}), \ell \in [p_\alpha], j \in [k]}$$

The decomposition is *low complexity*, in the sense that there are only a few of these values. However, we cannot take arbitrary values for these inner products since they may be far from *realizable*, i.e., no true assignment $b \in [q]^n$ can give rise to these values even approximately. From the inner products above, we can extract the following class of functions

$$\mathcal{F}' = \left\{ \mathbf{1}_{S_j^{\alpha,\ell}} \right\}_{\alpha \in \mathcal{A}(\mathfrak{I}), \ell \in [p_\alpha], j \in [k]} ,$$

whose size $r = |\mathcal{F}'| = O(|\mathcal{A}(\mathfrak{I})|\, k^3/\delta^2)$ is independent from n.

Using Claim 3.2, to be able to approximate $\mathsf{val}(\mathfrak{I}, b)$ within error $\delta' > 0$ we need to choose the error of the weak regularity decomposition[14] to be $\delta = \delta'/(2\,|\mathcal{A}(\mathfrak{I})|)$ In this case, we have $r = O(|\mathcal{A}(\mathfrak{I})|^2\, k^3/(\delta')^2) = O(q^{2k}k^3/(\delta')^2)$ and the τ-splittability parameter of W needs to satisfy $\tau \le \mathrm{poly}(\delta'/(kq^k))$.

For convenience, label the functions of $\mathcal{F}'$ as $f_1, \ldots, f_r$. Their range is the (simple) binary set $\mathcal{R} = \{0, 1\}$. We will consider the factor (see Section 4.2) $\mathcal{B}$ defined by the collection $\mathcal{F}'$, which, roughly speaking, is a partition of $[n]$ according to the values of these functions. More precisely, for every tuple $(y_1, \ldots, y_r) \in \mathcal{R}^r$ we have a (possibly empty) part (or atom) of the form

$$\{x \in [n] \;\mid\; f_1(x) = y_1, \ldots, f_r(x) = y_r\}\,.$$

In this case, we have at most $\mathcal{R}^r = 2^r$ atoms in the factor. By definition the functions $\mathcal{F}'$ are constant in each of them. An assignment b gives rise to a distribution on $[q]$ in each atom of the factor. Conversely, any approximate distribution on $[q]$ in each atom approximately corresponds to a realizable assignment b.

Let $L = \sum_{\alpha \in \mathcal{A}(\mathfrak{I}), \ell \in [p_\alpha]} |c_{\alpha,\ell}| \le |\mathcal{A}(\mathfrak{I})|\, O(k/\delta)$. Set $\eta = \delta/(k \cdot L \cdot q)$. We can η-approximate these distributions in ℓ_1-norm on each atom[15]. The number of approximate distributions can be (crudely) bounded as

$$(1/(\eta q))^{\mathcal{R}^r} \le \exp(\exp(\exp(\mathrm{poly}(qk/\delta'))))\,.$$

With this fine enough discretization of the distributions on each atom, when computing the expression

$$\mathsf{val}(\mathfrak{I}, b) = \sum_{\alpha \in \mathcal{A}} \sum_{\ell=1}^{p_\alpha} c_{\alpha,\ell} \frac{\left|S_1^{\alpha,\ell} \cap T_{\alpha_1}\right|}{n} \cdots \frac{\left|S_k^{\alpha,\ell} \cap T_{\alpha_k}\right|}{n} \;\pm\; \delta \cdot |\mathcal{A}|$$

we incur an additional error of $\delta'/2$. By our choice of δ, the total approximation error is at most δ'.

Running Time: By Theorem 5.12, the running time of the weak regularity decomposition is $\widetilde{O}(2^{2^{\widetilde{O}(k^2/\delta^2)}} \cdot |W|)$ per each computation of h_α. Combining the enumeration running and the time to compute these decompositions, we conclude the result. ∎

3.2 Stating the Extended Weak Regularity Framework

We now show how to obtain our main results for linear equations k-LIN over $\mathbb{Z}_q$ in Theorem 1.3 and over a finite group $\mathfrak{G}$ in Theorem 1.4.

In Section 3.3, we will see that to approximate k-LIN over $\mathbb{Z}_q$ it suffices to find a good weak regularity decomposition with respect to the test functions $\mathcal{F} = \mathrm{CUT}^{\otimes k}_{\omega,q,a}$ defined as follows (see Section 4 for a formal definition)

$$\mathrm{CUT}^{\otimes k}_{\omega,q,a} \coloneqq \{\chi_{b_1,a} \otimes \cdots \otimes \chi_{b_k,a} \;\mid\; b_1, \ldots, b_k \subseteq \mathbb{Z}_q^n\}\,.$$

[14] We can assume without loss of generality that $\mathcal{A}(\mathfrak{I}) \neq \emptyset$ since otherwise the value of the CSP is always zero.

[15] If the atom is too smaller than $1/(\eta q)$, then we can consider all the possible exact distribution.

In Section 3.4, we will see that to approximate k-LIN over a finite group, it suffices to find a good weak regularity decomposition with respect to the *matrix* valued test functions $\mathcal{F}$ defined as follows

$$\mathrm{CUT}_{\rho}^{\otimes k} := \{\rho_{b_1} \otimes \cdots \otimes \rho_{b_k} \mid b_1, \ldots, b_k \in \mathfrak{G}^n\}\,.$$

It will be more convenient to enlarge the test class $\mathcal{F}$ to unitary valued functions as follows

$$\mathrm{CUT}_{\mathbb{U}_{s,k,\delta}}^{\otimes k} := \{f_1 \otimes \cdots \otimes f_k \mid f_1, \ldots, f_k \colon [n] \to \mathbb{U}_{s,k,\delta}\}\,,$$

where $\mathbb{U}_{s,k,\delta}$ will be a fine enough discretization of the matrices[16] $M_s(\mathbb{C})$ of operator norm at most 1.

We will extend the framework to additionally handle the classes of functions $\mathrm{CUT}_{\omega,q,a}^{\otimes k}$ and $\mathrm{CUT}_{\mathbb{U}_{s,k,\delta}}^{\otimes k}$. This will be proven in Section 5.3. Let $\mathbb{K}$ be the underlying field which is either $\mathbb{R}$ or $\mathbb{C}$. Our extended framework gives the following efficient algorithmic result.

Theorem 5.12 (Efficient Weak Regularity (Extension of [JST21]))**.** *Let $W \subseteq [n]^k$ be a τ-splittable collection of tuples. Suppose $\mathcal{F}$ is one of $\mathrm{CUT}^{\otimes k}$, $\mathrm{CUT}_{\pm}^{\otimes k}$, $\mathrm{CUT}_{\omega,q,a}^{\otimes k}$, for $q \geq 3$, or $\mathrm{CUT}_{\mathbb{U}_{s,k,\delta}}^{\otimes k}$. Let $\mathcal{R}$ be the domain of the functions in $\mathcal{F}$, when $k = 1$. Let $g \in \mathcal{R}^{W[1]^k}$ be supported on W with $\|g\|_{\mu_k} \leq 1$. For every $\delta > 0$, if $\tau \leq \delta^2/(k^3 \cdot 2^{20})$, then we can find $h = \sum_{\ell=1}^{p} c_\ell \cdot f_\ell$ with $p = O(k^2/\delta^2)$, scalars $c_1, \ldots, c_p \in \mathbb{K}$ and functions $f_1, \ldots, f_p \in \mathcal{F}$, such that $\|h\|_{\mu_1^{\otimes k}} \leq 2$, $\sum_{\ell=1}^{p} |c_\ell| = O(k/\delta)$ and h is a good approximator to g in the following sense*

$$\max_{f \in \mathcal{F}} \left| \left\langle g - \left(\frac{d}{n}\right)^{k-1} h, f \right\rangle \right| \leq \delta \cdot |W|\,,$$

where the inner product is over the counting measure on $W[1]^k$. Furthermore, h can be found in $\widetilde{O}(2^{|\mathcal{R}|^{\widetilde{O}(k^2/\delta^2)}} \cdot |W|)$ time in the scalar valued case and in time $\widetilde{O}_{s,k,\delta}(\mathrm{poly}(|W|))$, otherwise.

3.3 Improved Case: k-LIN over $\mathbb{Z}_q$

The goal of this section is to prove Theorem 1.3 (restated below) assuming the new extended efficient regularity algorithm from Theorem 5.12.

Theorem 1.3 (Main II)**.** *Let $\mathfrak{I}$ be an instance of MAX k-LIN$_q$ on n variables with alphabet $\mathbb{Z}_q$ and constraints supported on a regular[17] collection of tuples $W \subseteq [n]^k$. If W is τ-splittable with $\tau \leq \tau_0(k, \delta) := \mathrm{poly}(\delta/k)$, then we can compute an assignment satisfying $\mathsf{OPT} - \delta$ in time $r(q/\tau_0) \cdot \widetilde{O}(|W| + n)$, where $r(x) = \exp(\exp(\mathrm{poly}(x)))$.*

For k-LIN over alphabet $\mathbb{Z}_q$, we are given a collection of equations (each variable appearing with coefficient one) specified as collection of tuples $W \subseteq [n]^k$ and we are given a collection of corresponding RHS $(r_w)_{w \in W} \in \mathbb{Z}_q^W$. The system of linear equations can be written as follows

$$x_{i_1} + \cdots + x_{i_k} = r_w \pmod q \qquad \forall\, w = (i_1, \ldots, i_k) \in W.$$

Orthogonality of Fourier characters will be crucially used here.

[16] We use $M_s(\mathbb{C})$ for the set of $s \times s$ matrices over $\mathbb{C}$.

[17] See Definition 4.1 for the definition of regular. This is an analog to tuples of a graph being d-vertex regular.

Fact 3.3 (Character Orthogonality)**.** *Let ω be a non-trivial q-th root of unit. Then,*

$$\mathop{\mathbb{E}}_{a\in\mathbb{Z}_q}\left[\omega^a\right] = 0\,.$$

Orthogonality allows for a convenient way of implementing the Dirac delta function on (the alphabet) $\mathbb{Z}_q$.

Fact 3.4. *Fix $y \in \mathbb{Z}_q$. The indicator function $x \mapsto \mathbf{1}_{[x=y]}$ on $\mathbb{Z}_q$ can be expressed as*

$$\mathop{\mathbb{E}}_{a\in\mathbb{Z}_q}\left[\omega^{a(x-y)}\right].$$

We now make precise the argument sketched in the proof strategy of Section 2.

Proof of Theorem 1.3. Let $\mathfrak{I}$ be an instance of k-LIN over $\mathbb{Z}_q$ with constraints supported on $W \subseteq [n]^k$ and RHS values $\{r_w\}_{w\in W}$. For every $a \in \mathbb{Z}_q$, we define $g_a \colon W \to \mathbb{C}$ as the map $w \in W \mapsto \omega^{a\cdot r_w}$, where $\omega = \exp(2\pi\sqrt{-1}/q)$.

Apply the efficient weak regularity decomposition of Theorem 5.12 to each g_a using error parameter $\delta > 0$ and test functions $\mathcal{F} = \mathrm{CUT}^{\otimes k}_{\omega,q,a}$. Note that this requires the splittability (expansion) parameter τ of W to satisfy $\tau \le O(\delta^2/k^3)$. We obtain a function $h_a = \sum_{\ell=1}^{p_a} c_{a,\ell}\cdot\chi_{b^{(a,\ell,1)},a}\otimes\cdots\otimes\chi_{b^{(a,\ell,k)},a}$, where $b^{(a,\ell,1)},\ldots,b^{(a,\ell,k)} \in \mathbb{Z}_q^n$, for every $a \in \mathbb{Z}_q$ and every $\ell \in [p_a]$. Let $b \in \mathbb{Z}_q^n$ be an assignment to the variables of the system of linear equations. The value of this CSP on input b can be computed as

$$\begin{aligned}
\mathsf{val}(\mathfrak{I},b) = \mathop{\mathbb{E}}_{a\in\mathbb{Z}_q}\left[\left\langle g_a,\chi_{b,a}\otimes\cdots\otimes\chi_{b,a}\right\rangle_{\mu_k}\right] &= \mathop{\mathbb{E}}_{a\in\mathbb{Z}_q}\left[\mathop{\mathbb{E}}_{w=(i_1,\ldots,i_k)\sim\mu_k}\left[\omega^{-a\cdot r_w}\omega^{a(b_{i_1}+\cdots+b_{i_k})}\right]\right]\\
&= \mathop{\mathbb{E}}_{w=(i_1,\ldots,i_k)\sim\mu_k}\left[\mathop{\mathbb{E}}_{a\in\mathbb{Z}_q}\left[\omega^{a(b_{i_1}+\cdots+b_{i_k}-r_w)}\right]\right]\\
&= \mathop{\mathbb{E}}_{w=(i_1,\ldots,i_k)\sim\mu_k}\left[\mathbf{1}_{[b_{i_1}+\cdots+b_{i_k}=r_w]}\right].
\end{aligned}$$

Using the weak regularity decomposition h_a of each g_a, we obtain

$$\begin{aligned}
\mathsf{val}(\mathfrak{I},b) &= \mathop{\mathbb{E}}_{a\in\mathbb{Z}_q}\left[\left\langle g_a,\chi_{b,a}\otimes\cdots\otimes\chi_{b,a}\right\rangle_{\mu_k}\right]\\
&= \frac{1}{|W|}\mathop{\mathbb{E}}_{a\in\mathbb{Z}_q}\left[\left\langle\left(\frac{d}{n}\right)^{k-1}h_a,\chi_{b,a}\otimes\cdots\otimes\chi_{b,a}\right\rangle\right] \pm \delta\\
&= \frac{1}{n^k}\mathop{\mathbb{E}}_{a\in\mathbb{Z}_q}\left[\sum_{\ell=1}^{p_a}c_{a,\ell}\cdot\left\langle\chi_{b^{(a,\ell,1)},a}\otimes\cdots\otimes\chi_{b^{(a,\ell,k)},a},\chi_{b,a}\otimes\cdots\otimes\chi_{b,a}\right\rangle\right] \pm \delta\\
&= \mathop{\mathbb{E}}_{a\in\mathbb{Z}_q}\left[\sum_{\ell=1}^{p_a}c_{a,\ell}\cdot\left\langle\chi_{b^{(a,\ell,1)},a},\chi_{b,a}\right\rangle_{\mu_1}\cdots\left\langle\chi_{b^{(a,\ell,k)},a},\chi_{b,a}\right\rangle_{\mu_1}\right] \pm \delta\,,
\end{aligned}$$

concluding the proof.

Now it suffices to approximate the following values

$$\left\{\left\langle\chi_{b^{(a,\ell,j)},a},\chi_{b,a}\right\rangle_{\mu_1}\right\}_{a\in\mathbb{Z}_q,\ell\in[p_a],j\in[k]},$$

so that there is always a true assignment $b \in [q]^n$ which gives these values.

To this end, we first define the following collection $\mathcal{F}'$ of functions

$$\mathcal{F}' = \left\{ \chi_{b^{(a,\ell,j)},a} \right\}_{a \in \mathbb{Z}_q, \ell \in [p_a], j \in [k]}.$$

Note that $r = |\mathcal{F}'| = O(qk^3/\delta^2)$. The functions above have range $\mathcal{R} = \{\omega^{a'} \mid a' \in \mathbb{Z}_q\}$. They form a factor (see Section 4.2) $\mathcal{B}$ with at most $|\mathcal{R}|^r$ atoms. By the definition of a factor, the functions $\mathcal{F}'$ are constant in each one of them, so to compute $\left\langle \chi_{b^{(a,\ell,j)},a}, \chi_{b,a} \right\rangle_{\mu_1}$ it suffices to know the distribution of symbols of b in each atom.

Let $L = \sum_{a \in \mathbb{Z}_q, \ell \in [p_a]} |c_{a,\ell}| = O(qk/\delta)$ and set $\eta = \delta/(k \cdot L \cdot q)$. The total number of η-approximate distributions in ℓ_1-norm on each atom can be (crudely) bounded as

$$(1/\eta q)^{|\mathcal{R}|^r} \leq \exp(\exp(\mathrm{poly}(qk/\delta))).$$

Using these distributions, we can approximate

$$\mathsf{val}(\mathfrak{I}, b) = \mathop{\mathbb{E}}_{a \in \mathbb{Z}_q} \left[\sum_{\ell=1}^{p_a} c_{a,\ell} \cdot \left\langle \chi_{b^{(a,\ell,1)},a}, \chi_{b,a} \right\rangle_{\mu_1} \cdots \left\langle \chi_{b^{(a,\ell,k)},a}, \chi_{b,a} \right\rangle_{\mu_1} \right] \pm \delta,$$

incurring an additional error of δ.

Running Time: To compute the decomposition h_a of g_a takes $\widetilde{O}(2^{|\mathcal{R}|^{\widetilde{O}(k^2/\delta^2)}} \cdot |W|)$ time (see Theorem 5.12). Combining the running time to find all h_a and the running time of enumerating over all η-approximate distributions, we obtain the claimed bound on the running time. ∎

3.4 Improved Case: k-LIN over a Finite Group $\mathfrak{G}$

The goal of this section is to prove Theorem 1.4 (restated below) assuming the new extended efficient regularity algorithm Theorem 5.12.

Theorem 1.4 (Main III). *Let $\mathfrak{I}$ be an instance of MAX k-LIN$_{\mathfrak{G}}$ on n variables with alphabet a finite group $\mathfrak{G}$ and constraints supported on a regular collection of tuples $W \subseteq [n]^k$. If W is τ-splittable with $\tau \leq \tau_0(k,\delta) := \mathrm{poly}(\delta/k)$, then we can compute an assignment satisfying $\mathsf{OPT} - \delta$ in time $O_{|\mathfrak{G}|,k,\delta}(1) \cdot \mathrm{poly}(|W| + n)$.*

We will need to work with matrix valued functions now in the weak regularity framework. We first establish some notation. Let $\mathcal{X}$ be a set endowed with a measure μ. Let $f, g \colon \mathcal{X} \to M_s(\mathbb{C})$. We define an inner product on $M_s(\mathbb{C})^{\mathcal{X}}$ as

$$\langle f, g \rangle := \mathop{\mathbb{E}}_{x \sim \mu} \left[\frac{\mathrm{Tr}(f(x)^\dagger g(x))}{s} \right].$$

Let $\mathfrak{G}$ be a finite group not necessarily Abelian. For k-LIN over alphabet $\mathfrak{G}$, we are given a collection of LHS sum of variables (each appearing with coefficient one) specified as collection of tuples $W \subseteq [n]^k$ and we are given a collection of corresponding RHS $(r_w)_{w \in W} \in \mathfrak{G}^W$. The system of linear equations can be written as follows

$$x_{i_1} \cdots x_{i_k} = r_w \qquad \forall w = (i_1, \ldots, i_k) \in W.$$

Let $\mathrm{Irrep}(\mathfrak{G})$ be a set of non-isomorphic irreducible representations. We define a distribution on this set by assigning probability $\dim(\rho)^2/|\mathfrak{G}|$ to an irreducible ρ. We will need the following simple fact about the sum of characters. This is the generalization of the orthogonality of Fourier characters on Abelian groups.

Fact 3.5 ([SS96]). *Let $\mathfrak{g} \in \mathfrak{G}$. Then*

$$\mathop{\mathbb{E}}_{\rho\sim\mathrm{Irrep}(G)}\left[\frac{\mathrm{Tr}(\rho(\mathfrak{g}))}{\dim(\rho)}\right] = \mathbf{1}_{[\mathfrak{g}=1]}\,,$$

where 1 is the identity element in $\mathfrak{G}$.

Let $b \in \mathfrak{G}^n$ be an assignment. Let $\mathbb{U}_s$ be the unitary group acting on $\mathbb{C}^s$. For a representation[18] $\rho\colon \mathfrak{G} \to \mathbb{U}_s$ of $\mathfrak{G}$, we define $\rho_b\colon [n] \to \mathbb{U}_s$ as $\rho(i) := \rho(b_i)$. We define $g_\rho\colon W \to \mathbb{U}_s$ as $g_\rho(w) = \rho(r_w)$. The value of the CSP can be expressed as

$$\begin{aligned}
\mathop{\mathbb{E}}_{\rho\sim\mathrm{Irrep}(\mathfrak{G})}\left[\left\langle g_\rho, \rho_b\otimes\cdots\otimes\rho_b\right\rangle_{\mu_k}\right] &= \mathop{\mathbb{E}}_{\rho\sim\mathrm{Irrep}(\mathfrak{G})}\left[\mathop{\mathbb{E}}_{w=(i_1,\ldots,i_k)\sim\mu_k}\left[\frac{\mathrm{Tr}(\rho(r_w)^\dagger\rho(b_{i_1})\cdots\rho(b_{i_k}))}{\dim(\rho)}\right]\right]\\
&= \mathop{\mathbb{E}}_{w=(i_1,\ldots,i_k)\sim\mu_k}\left[\mathop{\mathbb{E}}_{\rho\sim\mathrm{Irrep}(\mathfrak{G})}\left[\frac{\mathrm{Tr}(\rho(r_w^{-1}b_{i_1}\cdots b_{i_k}))}{\dim(\rho)}\right]\right]\\
&= \mathop{\mathbb{E}}_{w=(i_1,\ldots,i_k)\sim\mu_k}\left[\mathbf{1}_{\left[b_{i_1}\cdots b_{i_k}=r_w\right]}\right] = \mathsf{val}(\mathfrak{I},b)\,.
\end{aligned}$$

Proof of Theorem 1.4. Suppose we can find a weak regularity decomposition for each g_ρ of the form

$$\sum_{\ell=1}^{p_\rho} c_{\rho,\ell}\cdot\rho_{b^{\ell,1}}\otimes\cdots\otimes\rho_{b^{\ell,k}},$$

where $b^{\ell,1},\ldots,b^{\ell,k}\in\mathfrak{G}^n$ for every $\ell\in[p_\rho]$. Using these decompositions, the value becomes

$$\begin{aligned}
\mathsf{val}(\mathfrak{I},b) &= \mathop{\mathbb{E}}_{\rho\sim\mathrm{Irrep}(\mathfrak{G})}\left[\left\langle g_\rho,\rho_b\otimes\cdots\otimes\rho_b\right\rangle_{\mu_k}\right]\\
&= \frac{1}{|W|}\mathop{\mathbb{E}}_{\rho\sim\mathrm{Irrep}(\mathfrak{G})}\left[\left\langle\left(\frac{d}{n}\right)^{k-1}h_\rho,\rho_b\otimes\cdots\otimes\rho_b\right\rangle\right]\ \pm\ \delta\\
&= \frac{1}{n^k}\mathop{\mathbb{E}}_{\rho\sim\mathrm{Irrep}(\mathfrak{G})}\left[\sum_{\ell=1}^{p_\rho}c_{\rho,\ell}\cdot\left\langle\rho_{b^{\ell,1}}\otimes\cdots\otimes\rho_{b^{\ell,k}},\rho_b\otimes\cdots\otimes\rho_b\right\rangle\right]\ \pm\ \delta\\
&= \mathop{\mathbb{E}}_{\rho\sim\mathrm{Irrep}(\mathfrak{G})}\left[\sum_{\ell=1}^{p_\rho}c_{\rho,\ell}\cdot\mathop{\mathbb{E}}_{\mu_1^{\otimes k}}\left[\frac{\mathrm{Tr}\left(\left(\rho_{b^{\ell,1}}\otimes\cdots\otimes\rho_{b^{\ell,k}}\right)^\dagger\left(\rho_b\otimes\cdots\otimes\rho_b\right)\right)}{\dim(\rho)}\right]\right]\ \pm\ \delta\,.
\end{aligned}$$

The collection of functions $\{\rho_{b^{\ell,j}}\}_{\rho\in\mathrm{Irrep}(\mathfrak{G}),\ell\in[p_\rho],j\in[k]}$ gives rise to a factor $\mathcal{B}$ corresponding to a partition of $[n]$. The product factor $\mathcal{B}^{\otimes k}=\prod_{j=1}^k\mathcal{B}$ defines a partition of $[n]^k$. Note that each function $\rho_{b^{\ell,1}}\otimes\cdots\otimes\rho_{b^{i,k}}$ is $\mathcal{B}^{\otimes k}$-measurable since each $\rho_{b^{\ell,j}}$ is constant in $\mathcal{B}$.

[18] A unitary representation of a finite group G is a homomorphism from $\rho\colon G\to\mathbb{U}_s$. See [SS96, Sag13].

Instead of having to specify ρ_b with $b \in \mathfrak{G}^n$, we can brute force over all approximate distributions of values of $\mathfrak{G}$ on each part of the factor $\mathcal{B}$. We claim that this brute-force can be done in time $\widetilde{O}_{|\mathfrak{G}|,\delta,k}(n)$. The number of parts of $\mathcal{B}$ is at most $|\mathfrak{G}|^{O(k|\mathfrak{G}|/\delta^2)}$ and can be computed in time $\widetilde{O}_{|\mathfrak{G}|,\delta,k}(n)$. For each part, enumerating over all approximate distributions can be done in $O_{\mathfrak{G},\delta}(1)$ time. Let $\overline{\rho}_b \in M_{\dim(\rho)}(\mathbb{C})^{[n]}$ be an approximate function arising from this enumeration. Naively, using each $\overline{\rho}_b$, we can approximate the value (last line of equation above) in time $\widetilde{O}_{|\mathfrak{G}|,\delta,k}(n^k)$ by considering the expectation $\mu_1^{\otimes k}$. More carefully, we can exploit the fact that we have a product distribution $\mu_1 \otimes \mu_1^{\otimes k-1}$ to take the expectation with the first copy of μ_1 and then repeat this process to $\mu_1^{\otimes k-1}$. This leads to $\widetilde{O}_{|\mathfrak{G}|,\delta,k}(n)$ time.

We actually compute the decomposition using functions in $\mathrm{CUT}^{\otimes k}_{\mathbb{U}_{s,k,\delta}}$. Instead of $|\mathfrak{G}|$ in the computations above, we will have $|\mathbb{U}_{s,k,\delta}| = O_{s,k,\delta}(1)$. ∎

4 Some Definitions and Notation

We now introduce some notation. The asymptotic notation $\widetilde{O}(r(n))$ hides polylogarithmic factors in $r(n)$. We borrow some notation and definitions from [JST21].

4.1 Splittable Tuples

We now formally define the notion of (ordered) hypergraph expansion that we use. The *splittability* property for a collection of tuples

$W \subseteq [n]^k$. For $1 \le a \le b \le k$, we define $W[a,b] \subseteq [n]^{(b-a+1)}$ as

$$W[a,b] := \{(i_a, i_{a+1}, \ldots, i_b) \mid (i_1, i_2, \ldots, i_k) \in W\},$$

and use $W[a]$ to stand for $W[a,a]$. We will work with d-regular tuples in the following sense.

Definition 4.1 (Regular tuple collection)**.** *We say that $W \subseteq [n]^k$ is d-regular if for every $1 \le a \le b \le k$, we have*

- $|W[a,b]| = d^{b-a} \cdot n$,
- $W[a] = [n]$.

A collection W being d-regular is analogous to a graph being d-regular.

Example 4.2. *The collection W of all length-$(k-1)$ walks on a d-regular connected graph $G = ([n], E)$ is a d-regular collection of tuples.*

Let $\mathbb{K}$ be a field that is either $\mathbb{R}$ or $\mathbb{C}$. The space of functions $\mathbb{K}^{W[a,b]}$ is endowed with an inner product associated to the uniform measure $\mu_{[a,b]}$ on $W[a,b]$. We use the shorthand μ_b for $\mu_{[1,b]}$.

Definition 4.3 (Splitable tuple collection)**.** *Let $\tau > 0$. We say that a collection $W \subseteq [n]^k$ is* τ-splittable *if it is d-regular and either $k = 1$ or for every $1 \le a \le t < b \le k$ we have*

- *the* split *operator* $\mathsf{S}_{W[a,t],W[t+1,b]} \in \mathbb{K}^{W[a,t]\times W[t+1,b]}$ *defined as*

$$\left(\mathsf{S}_{W[a,t],W[t+1,b]}\right)_{(i_a,\ldots,i_t),(i_{t+1},\ldots,i_b)} := \frac{\mathbf{1}\left[(i_a,\ldots,i_t,i_{t+1},\ldots i_b) \in W[a,b]\right]}{d^{k-t}}$$

satisfy $\sigma_2(\mathsf{S}_{W[a,t],W[t+1,b]}) \leq \tau$*, where* σ_2 *denotes the second largest singular value.*

Example 4.4. *The collection* W *of all length-*$(k-1)$ *walks on a* d*-regular a graph* $G = ([n], E)$ *whose normalized adjacency matrix has second largest singular value at most* τ *is a collection of* τ*-splittable tuples as shown in [AJQ+20].*

Example 4.5. *The collection* W *of tuples arising (from a slight modification) of the direct sum construction of Ta-Shma [TS17] is a* τ*-splittable as shown in [JQST20].*

4.2 Factors

It will be convenient to use the language of factors, to search the decompositions identified by regularity lemmas, for relevant codewords. This concept (from ergodic theory) takes a rather simple form in our finite settings: it is just a partition of base set $\mathcal{X}$, with an associated operation of averaging functions defined on $\mathcal{X}$, separately over each piece.

Definition 4.6 (Factors and measurable functions)**.** *Let* $\mathcal{X}$ *be a finite set. Let* $\mathcal{Y}$ *be a Hilbert space over* $\mathbb{K}$ *endowed with an inner product* $\langle\cdot,\cdot\rangle_{\mathcal{Y}}$*. We further assume that* $\mathcal{Y}$ *is a ring*[19] *so that in particular we can multiply the elements of* $\mathcal{Y}$*. Let* $\mathcal{R} \subseteq \mathcal{Y}$ *be a (possibly discrete) subset. A* factor $\mathcal{B}$ *is a partition of the set* $\mathcal{X}$*, and the subsets of the partition are referred to as* atoms of the factor. A function $f : \mathcal{X} \to \mathcal{Y}$ is said to measurable with respect to $\mathcal{B}$ ($\mathcal{B}$-measurable) if f is constant on each atom of $\mathcal{B}$.

Definition 4.7 (Conditional averages)**.** *If* $f : \mathcal{X} \to \mathcal{Y}$ *is a function,* μ *is a measure on the space* $\mathcal{X}$*, and* $\mathcal{B}$ *is a factor, then we define the* conditional average function $\mathbb{E}[f|\mathcal{B}]$ *as*

$$\mathbb{E}\left[f|\mathcal{B}\right](x) := \mathop{\mathbb{E}}_{y\sim\mu|\mathcal{B}(x)}\left[f(y)\right],$$

where $\mathcal{B}(x)$ *denotes the atom containing* x*. Note that the function* $\mathbb{E}[f|\mathcal{B}]$ *is measurable with respect to* $\mathcal{B}$*.*

We will need the following simple observation regarding conditional averages.

Proposition 4.8. *Let* $h : \mathcal{X} \to \mathcal{Y}$ *be a* $\mathcal{B}$*-measurable function, and let* $f : \mathcal{X} \to \mathcal{Y}$ *be any function. Then, for any measure* μ *over* $\mathcal{X}$*, we have*

$$\langle h, f\rangle_\mu = \langle h, \mathbb{E}\left[f|\mathcal{B}\right]\rangle_\mu .$$

Proof. By definition of the $\mathcal{B}$-measurability, h is constant on each atom, and thus we can write $h(x)$ as $h(\mathcal{B}(x))$.

$$\begin{aligned}\langle h, f\rangle_\mu = \mathop{\mathbb{E}}_{x\sim\mu}\left[h(x)\cdot f(x)\right] &= \mathop{\mathbb{E}}_{x\sim\mu}\mathop{\mathbb{E}}_{y\sim\mu|\mathcal{B}(x)}\left[h(y)\cdot f(y)\right]\\ &= \mathop{\mathbb{E}}_{x\sim\mu}\left[h(\mathcal{B}(x))\cdot\mathop{\mathbb{E}}_{y\sim\mu|\mathcal{B}(x)}\left[f(y)\right]\right]\\ &= \mathop{\mathbb{E}}_{x\sim\mu}\left[h(x)\cdot\mathbb{E}\left[f|\mathcal{B}\right](x)\right] = \langle h, \mathbb{E}\left[f|B\right]\rangle_\mu . \qquad \blacksquare\end{aligned}$$

[19] Under our assumptions $\mathcal{Y}$ is an algebra over $\mathbb{K}$.

The factors we will consider will be defined by a finite collection of functions appearing in a regularity decomposition.

Definition 4.9 (Function factors)**.** *Let $\mathcal{X}$ and $\mathcal{R}$ be finite sets, and let $\mathcal{F}_0 = \{f_1, \ldots, f_r : \mathcal{X} \to \mathcal{R}\}$ be a finite collection of functions. We consider the factor $\mathcal{B}_{\mathcal{F}_0}$ defined by the functions in $\mathcal{F}_0$, as the factor with atoms $\{x \mid f_1(x) = a_1, \ldots, f_r(x) = a_r\}$ for all $(a_1, \ldots, a_r) \in \mathcal{R}^r$.*

Remark 4.10. *Note that when the above functions are indicators for sets i.e., each $f_j = \mathbf{1}_{S_j}$ for some $S_j \subseteq \mathcal{X}$, then the function factor $\mathcal{B}_{\mathcal{F}_0}$ is the same as the σ-algebra generated by these sets*[20]*. Also, given the functions $f_1, \ldots, f_r$ as above, the function factor $\mathcal{B}_{\mathcal{F}_0}$ can be computed in time $O(|\mathcal{X}| \cdot |\mathcal{R}|^r)$.*

4.3 Functions and Measures

We describe below some classes of functions, and spaces with associated measures, arising in our proof. The measures we consider are either uniform on the relevant space, or are products of measures on its component spaces.

Function classes. Let $S \subseteq [n]$. We define $\chi_S \colon [n] \to \{\pm 1\}$ as $\chi_S(i) := (-1)^{\mathbf{1}_{i \in S}}$ (we observe that as defined χ_S is not a character[21]).

Let q be a positive integer and let $\omega = \exp(2\pi\sqrt{-1}/q)$ be a primitive qth root of the unity. We define $\mathbb{S}^1_q = \{\omega^a \mid a \in \mathbb{Z}_q\}$. Given $a \in \mathbb{Z}_q$ and $b \in \mathbb{Z}_q^n$, we define $\chi_{b,a} \colon [n] \to \mathbb{S}^1_q$ as $\chi_{b,a}(i) := \omega^{a \cdot b_i}$. Let $\mathbb{U}_s$ be the unitary group acting on $\mathbb{C}^s$.

We will need the following collections of functions.

Definition 4.11 (Generalized CUT functions)**.** *We define the set of* $0/1$ CUT *cut functions as*

$$\mathrm{CUT}^{\otimes k} := \{\pm \mathbf{1}_{S_1} \otimes \cdots \otimes \mathbf{1}_{S_k} \mid S_1, \ldots, S_k \subseteq [n]\}.$$

We define the set of ± 1 CUT *functions as*

$$\mathrm{CUT}^{\otimes k}_{\pm} := \{\pm \chi_{S_1} \otimes \cdots \otimes \chi_{S_k} \mid S_1, \ldots, S_k \subseteq [n]\}.$$

For $a \in \mathbb{Z}_q$, we define the set of roots of unity CUT *functions as*

$$\mathrm{CUT}^{\otimes k}_{\omega,q,a} := \{\chi_{b^{(1)},a} \otimes \cdots \otimes \chi_{b^{(k)},a} \mid b^{(1)}, \ldots, b^{(k)} \subseteq \mathbb{Z}_q^n\}.$$

We define the set of unitary CUT *functions as*

$$\mathrm{CUT}^{\otimes k}_{\mathbb{U}_s} := \{f_1 \otimes \cdots \otimes f_k \mid f_1, \ldots, f_k \colon [n] \to \mathbb{U}_s\}.$$

Let $\mathfrak{G}$ be a group and $\rho \colon G \to \mathbb{U}_s$ be a unitary representation. Let $b \in \mathfrak{G}^n$. We defined $\rho_b \colon [n] \to \mathrm{im}(\rho)$ as $\rho_b(i) := \rho(b_i)$. We define the set of ρ CUT *functions as*

$$\mathrm{CUT}^{\otimes k}_{\rho} := \{\rho_{b^{(1)}} \otimes \cdots \otimes \rho_{b^{(k)}} \mid b^{(1)}, \ldots, b^{(k)} \in \mathfrak{G}^n\}.$$

[20] For a finite $\mathcal{X}$, the σ-algebra generated by $S_1, \ldots, S_p \subseteq \mathcal{X}$ is the smallest subset of the power set of $\mathcal{X}$ containing $\mathcal{X}$, $S_1, \ldots, S_p$ that is closed under union, intersection and complement. This finite version will be enough for us in this work (see [Bil95] for the general definition).

[21] Strictly speaking χ_S is not a character but by identifying the elements of $[n]$ with those of a canonical basis of $\mathbb{F}_2^n$ it becomes a character for $\mathbb{F}_2^n$.

Let $f\colon \mathcal{X} \to \mathcal{Y}$. If $\mathcal{Y}$ the set of matrices $M_s(\mathbb{K})$, then we will use the operator norm $\|f\|_\infty = \max_{x\in\mathcal{X}} \|f(x)\|_{\mathrm{op}}$. Note that If $\mathcal{Y} = \mathbb{K}$ (equivalently $s = 1$), then we have the usual definition $\|f\|_\infty \coloneqq \max_{x\in\mathcal{X}} |f(x)|$.

Some of our results hold for more general classes of functions.

Definition 4.12 (t-split functions)**.** *Suppose W is a regular collection of k-tuples. For $t \in \{0,\dots,k-1\}$, we define a generic class of tensor product functions $\mathcal{F}_t$ as*

$$\mathcal{F}_t \subseteq \left\{ \pm f_1 \otimes \cdots \otimes f_t \otimes f_{t+1} \mid f_j \in \mathcal{R}^{W[1]} \text{ for } j \le t, f_{t+1} \in \mathcal{R}^{W[t+1,k]}, \left\|f_j\right\|_\infty \le 1 \text{ for } j \le t+1 \right\}.$$

To avoid technical issues, we assume that each $\mathcal{F}_t$ is finite.

Fixing some $\mathcal{F} \subseteq \mathcal{R}^{\mathcal{X}}$, we define the set of functions that are linear combinations of function from $\mathcal{F}$ with coefficients of bounded support size and bounded ℓ_1-norm as follows

$$\mathcal{H}(R_0, R_1, \mathcal{F}) \coloneqq \left\{ \sum_{\ell=1}^{p} c_\ell \cdot f_\ell \mid p \le R_0, \sum |c_\ell| \le R_1, f_\ell \in \mathcal{F} \right\}.$$

Measures and inner products. Recall that $\mu_1 \coloneqq \mu_{[1,1]}$ is the uniform measure on $W[1]$ (equivalently uniform measure on $W[i]$ since W is regular) and $\mu_{[t+1,k]}$ is the uniform measure on $W[t+1,k]$. We define the following measure ν_t as

$$\nu_t \coloneqq (\mu_1)^{\otimes t} \otimes \left(\mu_{[t+1,k]}\right).$$

Note that ν_0 is the equal to μ_k and ν_{k-1} is equal to $\mu_1^{\otimes k}$. We will need to consider inner products of functions according to various measures defined above, which we will denote as $\langle \cdot,\cdot \rangle_\mu$ for the measure μ. When a measure is not indicated, we take the inner product $\langle f, g\rangle$ to be according to the counting measure on the domains of the functions f and g.

Constraint Satisfaction Problems

An instance $\mathfrak{I}$ of a k-CSP on n variables over alphabet $[q]$ is given by a collection of tuples $W \subseteq [n]^k$ endowed with a probability measure μ_k together with a collection of predicates $(P_w\colon [q]^k \to \{0,1\})_{w\in W}$ on these tuples. For an assignment $b \in [q]^n$, its values $\mathsf{val}(\mathfrak{I}, b)$ are defined as

$$\mathsf{val}(\mathfrak{I}, b) \coloneqq \mathop{\mathbb{E}}_{w=(i_1,\cdots,i_k)\sim_{\mu_k} W} \left[P_w(b_{i_1},\dots,b_{i_k})\right].$$

The optimum value $\mathsf{OPT}(\mathfrak{I})$ of $\mathfrak{I}$ is given by $\mathsf{OPT}(\mathfrak{I}) = \max_{b\in[q]^n} \mathsf{val}(\mathfrak{I}, b)$.

5 Weak Regularity

We will generalize the weak regularity framework [JST21] for more general classes of functions. For the *existential* weak regularity, most results extend quite naturally to accommodate some general classes except for a new version of their splittable mixing lemma for

matrix valued functions. We make the *algorithmic* component more modular to accommodate these new classes and we will also need to implement specific correlation oracles for them. We will need to generalize and extend the framework in several directions, but when this is not necessary we will borrow from [JST21].

5.1 Abstract Weak Regularity Lemma

We now show a weak regularity decomposition lemma for functions that works in some generality and does not require splittability. This will be a simple extension of the abstract decomposition of [JST21] to also handle the field of complex numbers. We now fix some notation for this section. Let $\mathcal{X}$ be a finite set endowed with a probability measure μ. Let $\mathbb{K}$ be a field that is either $\mathbb{R}$ or $\mathbb{C}$. Let $\mathcal{Y}$ be a Hilbert space over $\mathbb{K}$ endowed with an inner product $\langle \cdot, \cdot \rangle_{\mathcal{Y}}$. Let $\mathcal{R} \subseteq \mathcal{Y}$ be a finite set, and let $\mathcal{Y}^{\mathcal{X}}$ be a Hilbert space endowed with inner product $\langle f, g \rangle_\mu = \mathbb{E}_{x \sim \mu} \left[\langle f(x), g(x) \rangle_{\mathcal{Y}} \right]$ and associated norm $\|\cdot\|_\mu = \sqrt{\langle \cdot, \cdot \rangle_\mu}$. Let $\mathcal{F} \subseteq \{ f \colon \mathcal{X} \to \mathcal{Y} \mid \|f\|_\mu \leq 1 \}$ be a finite collection of functions.

In a nutshell, given any $g \in \mathcal{R}^{\mathcal{X}}$, the abstract weak regularity lemma will allow us to find an approximator h, with respect to the semi-norm[22] $g - h \mapsto \max_{f \in \mathcal{F}} |\langle g - h, f \rangle|$, which is a linear combinations of a certain *small* number of functions from $\mathcal{F}$ (where this number depends only on the approximation accuracy and the norm $\|g\|_\mu$). This means that g and h have approximately the same correlations with functions from $\mathcal{F}$. We will produce h in an iterative procedure, where at each step an oracle of the following kind (*cf.*, Definition 5.1) is invoked.

Definition 5.1 (Correlation Oracle)**.** *Let $1 \geq \delta \geq \delta' > 0$ be accuracy parameters and $B > 0$. We say that $\mathcal{O}_{\mu,B}$ is a (δ, δ')-correlation oracle for $\mathcal{F}$ if given $h \in \mathcal{R}^{\mathcal{X}}$ with $\|h\|_\mu^2 = O(B)$ if there exists $f \in \mathcal{F}$ with $|\langle h, f \rangle| \geq \delta$, then $\mathcal{O}_{\mu,B}$ returns some $f' \in \mathcal{F}$ with $|\langle h, f' \rangle| \geq \delta'$.*

More precisely, our abstract weak regularity decomposition is as follows.

Lemma 5.2 (Abstract Weak Regularity (Extension of [JST21]))**.** *Let $\mathcal{O}_{\mu,B}$ be a (δ, δ')-correlation oracle for $\mathcal{F}$ with $\delta \geq \delta' > 0$. Let $g \colon \mathcal{X} \to \mathcal{R}$ satisfy $\|g\|_\mu^2 \leq B$. Then, we can find $h = \sum_{\ell=1}^{p} c_\ell \cdot f_\ell \in \mathcal{H}(B/(\delta')^2, B/\delta', \mathcal{F})$ with $f_\ell \in \mathcal{F}$, $c_\ell \in [\delta'/(1 + \delta'/\sqrt{B})^p, \delta']$ and $\|h\|_\mu^2 \leq B$ such that*

$$\max_{f \in \mathcal{F}} \quad \left| \langle g - h, f \rangle_\mu \right| \ \leq \ \delta.$$

Furthermore, if $\mathcal{O}_{\mu,B}$ runs in time $\mathcal{T}_{\mathcal{O}_{\mu,B}}$, then h can be computed in

$$\widetilde{O}\left(\mathrm{poly}(B, 1/\delta') \cdot (\mathcal{T}_{\mathcal{O}_{\mu,B}} + |\mathrm{Supp}(\mu)|) \right)$$

time, where $\mathrm{Supp}(\mu)$ is the support of μ. The function h is constructed in Algorithm 5.3 as the final function in a sequence of approximating functions $h^{(\ell)} \in \mathcal{H}(B/(\delta')^2, B/\delta', \mathcal{F})$.

The proof is based on the following algorithm.

[22] See [Rud91, Chapter 1] for a definition of semi-norm.

Algorithm 5.3 (Regularity Decomposition Algorithm).
Input $g\colon \mathcal{X} \to \mathcal{R}$
Output $h = \sum_{\ell=1}^{p} c_\ell \cdot f_\ell$

- *Let* Π *be the projector onto the convex ball* $\{g' \in \mathcal{R}^{\mathcal{X}} \mid \|g'\|_\mu^2 \le B\}$.
- *Let* $\ell = 0$ *and* $h^{(\ell)} = 0$
- *While* $\max_{f\in\mathcal{F}} \left|\left\langle g - h^{(\ell)}, f\right\rangle_\mu\right| \ge \delta$*:*
  - $\ell = \ell + 1$
  - *Let* $f_\ell \in \mathcal{F}$ *be such that* $\left|\left\langle g - h^{(\ell-1)}, f_\ell\right\rangle_\mu\right| \ge \delta'$ (Correlation Oracle $\mathcal{O}_{\mu,B}$ Step)
  - *Let* $\rho \cdot \exp(i\theta) = \left\langle g - h^{(\ell-1)}, f_\ell\right\rangle_\mu$ *be its polar form*
  - *Let* $c_\ell = \delta' \cdot \exp(-i\theta)$
  - $h^{(\ell)} = \Pi(h^{(\ell-1)} + c_\ell \cdot f_\ell)$
- *Let* $p = \ell$
- *return* $h = \sum_{\ell=1}^{p} c_\ell \cdot f_\ell$

We will need the following general fact about projections[23] onto a convex body.

Fact 5.4 (Implicit in Lemma 3.1 of [Bub15])**.** *Let* $\mathcal{Y}$ *be a compact convex body in a finite dimensional Hilbert space* $\mathcal{V}$ *equipped with inner product* $\langle\cdot,\cdot\rangle_\nu$ *and associated norm* $\|\cdot\|_\nu$*. Let* $\Pi_{\mathcal{Y}}$ *be projector onto* $\mathcal{Y}$*. Then, for* $y \in \mathcal{Y}$ *and* $x \in \mathcal{V}$*, we have*

$$\|y - x\|_\nu^2 \;\ge\; \|y - \Pi_{\mathcal{Y}}(x)\|_\nu^2 \;+\; \|\Pi_{\mathcal{Y}}(x) - x\|_\nu^2 .$$

Proof of Lemma 5.2. We will show that the norm of $\left\|g - h^{(\ell)}\right\|_\mu$ strictly decreases as the algorithm progresses. Computing we obtain

$$\begin{aligned}
\left\|g - h^{(\ell)}\right\|_\mu^2 &= \left\|g - \Pi(h^{(\ell-1)} + c_\ell \cdot f_\ell)\right\|_\mu^2 \\
&\le \left\|g - (h^{(\ell-1)} + c_\ell \cdot f_\ell)\right\|_\mu^2 - \left\|(h^{(\ell-1)} + c_\ell \cdot f_\ell) - \Pi(h^{(\ell-1)} + c_\ell \cdot f_\ell)\right\|_\mu^2 \quad \text{(By Fact 5.4)} \\
&\le \left\|g - (h^{(\ell-1)} + c_\ell \cdot f_\ell)\right\|_\mu^2 \\
&= \left\|g - h^{(\ell-1)}\right\|_\mu^2 + |c_\ell|^2 \cdot \|f_\ell\|_\mu^2 - 2\Re\left(c_\ell \cdot \left\langle g - h^{(\ell-1)}, f_\ell\right\rangle_\mu\right) \\
&= \left\|g - h^{(\ell-1)}\right\|_\mu^2 + |c_\ell|^2 \cdot \|f_\ell\|_\mu^2 - 2\delta\left|\left\langle g - h^{(\ell-1)}, f_\ell\right\rangle_\mu\right| \\
&\le \left\|g - h^{(\ell-1)}\right\|_\mu^2 - (\delta')^2
\end{aligned}$$

[23] See [Bub15, Chapter 3] for a defintion of projector.

where the inequality follows from $c_\ell = \delta' \cdot \exp(-i\theta)$ with $\left\langle g - h^{(\ell-1)}, f_\ell \right\rangle_\mu = \rho \cdot \exp(i\theta)$ its polar form, the bound $\|f_\ell\|_\mu \leq 1$ and

$$\left| \left\langle g - h^{(\ell-1)}, f_\ell \right\rangle_\mu \right| \geq \delta'.$$

Since $\|g\|_\mu^2 \leq B$ and $\left\|g - h^{(\ell)}\right\|_\mu^2$ decreases by at least $(\delta')^2$ in each iteration, we conclude that the algorithm halts in at most $p \leq B/(\delta')^2$ steps.

From this point, the proof continues as in [JST21]. By construction each c_ℓ is initialized to δ' and can not increase (it can only decrease due to projections). Thus, we obtain $\sum_{\ell=1}^{p} |c_\ell| \leq p \cdot \delta' \leq B/\delta'$. Also by construction at termination $\|h\|_\mu^2 \leq B$. It remains to show that $c_\ell \geq \delta'/(1+\delta'/\sqrt{B})^p$. Note that the projection $\Pi(h^{(\ell-1)} + c_\ell \cdot f_\ell)$ at each iteration either does nothing to the coefficients c_ℓ's or scales them by a factor of at most $(1+\delta'/\sqrt{B})$ since $\left\|h^{(\ell-1)}\right\|_\mu + \|c_\ell \cdot f_\ell\|_\mu \leq \sqrt{B}(1+\delta'/\sqrt{B})$. This readily implies the claimed lower bound on the coefficients c_ℓ's at termination. Moreover, we have $h^{(\ell)} \in \mathcal{H}(B/(\delta')^2, B/\delta', \mathcal{F})$ also by construction.

Running Time: The decomposition algorithm calls the correlation oracle at most $p+1$ times. Since the coefficients c_ℓ always lie in $[\delta'/(1+\delta'/\sqrt{B})^p, \delta'] \subseteq [\delta'/\exp(p\delta'/\sqrt{B}), \delta']$, the bit complexity is $C = O(p\delta'/\sqrt{B})$ and computing the projection (which amounts to computing $h^{(\ell)}/\left\|h^{(\ell)}\right\|_\mu$ if $\left\|h^{(\ell)}\right\|_\mu^2 > B$) takes at most $\widetilde{O}(p^2 \cdot \mathrm{poly}(C) \cdot |\mathrm{Supp}(\mu)|)$. Then the total running time is at most

$$\widetilde{O}(p(\mathcal{T}_{\mathcal{O}_{\mu,B}} + p^2 \cdot \mathrm{poly}(C) \cdot |\mathrm{Supp}(\mu)|)) = \widetilde{O}\left(\mathrm{poly}(B, 1/\delta') \cdot (\mathcal{T}_{\mathcal{O}_{\mu,B}} + |\mathrm{Supp}(\mu)|)\right),$$

concluding the proof. ∎

Remark 5.5. *If we are only interested in an existential version of Lemma 5.2, we can always use a trivial existential (δ, δ)-correlation oracle. However, to obtain weak regularity decompositions efficiently in our settings, we will later use efficient (δ, δ')-correlation oracle with $\delta' = \Omega(\delta)$.*

Splittable Matrix Mixing Lemma

A splittable collection of tuples gives rise to several expanding split operators (see Definition 4.3). This allows us to show that a splittable collection satisfies some higher-order analogues of the well known expander mixing lemmas for graphs (*cf.*,[HLW06][Section 2.4]) as we make precise next. The extension to complex numbers from [JST21] is immediate.

We now prove a new matrix version of the splittable mixing lemma which is a high-order generalization of the matrix mixing lemma [CMR13]. The non-commutativity will require some extra care in this higher-dimensional version.

Lemma 5.6 (Splittable Matrix Mixing Lemma)**.** *Let $\mathcal{Y} = M_\ell(\mathbb{K})$, i.e., the images of our functions are $\ell \times \ell$ matrices over $\mathbb{K}$. Suppose $W \subseteq [n]^k$ is a τ-splittable collection of tuples. For every $t \in \{0, \ldots, k-2\}$ and every $f, f' \in \mathcal{F}_{t+1}$, we have*

$$\left| \langle f, f' \rangle_{\nu_{t+1}} - \langle f, f' \rangle_{\nu_t} \right| \leq \tau.$$

Proof. Recall that by our assumption on $\mathcal{F}_{t+1}$, the functions map to matrices of operator norm at most 1 (see definition in Section 4).

Let $f = f_1 \otimes \cdots \otimes f_t \otimes f_{t+1} \otimes f_{t+2}$ and $f' = f'_1 \otimes \cdots \otimes f'_t \otimes f'_{t+1} \otimes f'_{t+2}$. To simplify the notation we can write $f = f_a \otimes f_b$ and $f' = f'_a \otimes f'_b$, where $f_a = f_1 \otimes \cdots \otimes f_t$, $f_b = f_{t+1} \otimes f_{t+2}$, $f'_a = f'_1 \otimes \cdots \otimes f'_t$ and $f'_b = f'_{t+1} \otimes f'_{t+2}$. For fixed $\mathfrak{s} \in W[1]^t$, $\mathfrak{t} \in W[1]$ and $\mathfrak{u} \in W[1]^{k-t-1}$, using the cyclic property of the trace, we have[24]

$$\begin{aligned}
\operatorname{Tr}(f(\mathfrak{stu})^\dagger f'(\mathfrak{stu})) &= \operatorname{Tr}(f_b(\mathfrak{tu})^\dagger f_a(\mathfrak{s})^\dagger f'_a(\mathfrak{s}) f'_b(\mathfrak{tu})) \\
&= \operatorname{Tr}(f_a(\mathfrak{s})^\dagger f'_a(\mathfrak{s}) f'_b(\mathfrak{tu}) f_b(\mathfrak{tu})^\dagger) \\
&= \operatorname{Tr}(f_a(\mathfrak{s})^\dagger f'_a(\mathfrak{s}) f'_{t+1}(\mathfrak{t}) f'_{t+2}(\mathfrak{u}) f_{t+2}(\mathfrak{u})^\dagger f_{t+1}(\mathfrak{t})^\dagger) \\
&= \operatorname{Tr}(f_{t+1}(\mathfrak{t})^\dagger f_a(\mathfrak{s})^\dagger f'_a(\mathfrak{s}) f'_{t+1}(\mathfrak{t}) f'_{t+2}(\mathfrak{u}) f_{t+2}(\mathfrak{u})^\dagger)\,.
\end{aligned}$$

Let $f_0 = \mathbb{E}_{\mathfrak{s}\sim\mu_1^{\otimes t}} f_a(\mathfrak{s})^\dagger f'_a(\mathfrak{s})$. Note that

$$\begin{aligned}
\ell \cdot \langle f, f' \rangle_{\nu_{t+1}} &= \mathop{\mathbb{E}}_{\mathfrak{s}\sim\mu_1^{\otimes t}} \mathop{\mathbb{E}}_{(\mathfrak{t},\mathfrak{u})\sim\mu_1\otimes\mu_{[t+2,k]}} \operatorname{Tr}(f(\mathfrak{stu})^\dagger f'(\mathfrak{stu})) \\
&= \mathop{\mathbb{E}}_{\mathfrak{s}\sim\mu_1^{\otimes t}} \mathop{\mathbb{E}}_{(\mathfrak{t},\mathfrak{u})\sim\mu_1\otimes\mu_{[t+2,k]}} \operatorname{Tr}(f_{t+1}(\mathfrak{t})^\dagger f_a(\mathfrak{s})^\dagger f'_a(\mathfrak{s}) f'_{t+1}(\mathfrak{t}) f'_{t+2}(\mathfrak{u}) f_{t+2}(\mathfrak{u})^\dagger) \\
&= \mathop{\mathbb{E}}_{(\mathfrak{t},\mathfrak{u})\sim\mu_1\otimes\mu_{[t+2,k]}} \operatorname{Tr}\left(f_{t+1}(\mathfrak{t})^\dagger \mathop{\mathbb{E}}_{\mathfrak{s}\sim\mu_1^{\otimes t}} \left[f_a(\mathfrak{s})^\dagger f'_a(\mathfrak{s}) \right] f'_{t+1}(\mathfrak{t}) f'_{t+2}(\mathfrak{u}) f_{t+2}(\mathfrak{u})^\dagger \right) \\
&= \mathop{\mathbb{E}}_{(\mathfrak{t},\mathfrak{u})\sim\mu_1\otimes\mu_{[t+2,k]}} \operatorname{Tr}\left(f_{t+1}(\mathfrak{t})^\dagger f_0 f'_{t+1}(\mathfrak{t}) f'_{t+2}(\mathfrak{u}) f_{t+2}(\mathfrak{u})^\dagger \right)\,,
\end{aligned}$$

and similarly $\ell \cdot \langle f, f' \rangle_{\nu_t} = \mathbb{E}_{(\mathfrak{t},\mathfrak{u})\sim\mu_{[t+1,k]}} \operatorname{Tr}\left(f_{t+1}(\mathfrak{t})^\dagger f_0 f'_{t+1}(\mathfrak{t}) f'_{t+2}(\mathfrak{u}) f_{t+2}(\mathfrak{u})^\dagger \right)$.

Let $f''_{t+1} = f^\dagger_{t+1} f_0 f'_{t+1}$ and $f''_{t+2} = f_{t+2} f^\dagger_{t+2}$. Now to bound $\left| \langle f, f' \rangle_{\nu_{t+1}} - \langle f, f' \rangle_{\nu_t} \right|$, it suffices to bound the following operator norm

$$\left\| \mathop{\mathbb{E}}_{\mu_1\otimes\mu_{[t+2,k]}} f''_{t+1} \otimes f''_{t+2} - \mathop{\mathbb{E}}_{\mu_{[t+1,k]}} f''_{t+1} \otimes f''_{t+2} \right\|_{\mathrm{op}}\,.$$

Note that

$$\mathop{\mathbb{E}}_{\mu_1\otimes\mu_{[t+2,k]}} f''_{t+1} \otimes f''_{t+2} - \mathop{\mathbb{E}}_{\mu_{[t+1,k]}} f''_{t+1} \otimes f''_{t+2} = \left\langle f''_{t+1}, \left(\left(\frac{\mathsf{J}_{\mathrm{rec}}}{|W[t+2,k]|} - \mathsf{S}_{W[t+1],W[t+2,k]} \right) \otimes \mathsf{I}_\ell \right) f''_{t+2} \right\rangle_{\mu_1}\,,$$

where $\mathsf{J}_{\mathrm{rec}}$ is the (rectangular) $|W[t+1]| \times |W[t+2,k]|$ all ones matrix. Using the τ-splittability assumption, we have the following bound on the largest singular value

$$\sigma\left(\frac{\mathsf{J}_{\mathrm{rec}}}{|W[t+2,k]|} - \mathsf{S}_{W[t+1],W[t+2,k]} \right) \le \sigma_2\left(\mathsf{S}_{W[t+1],W[t+2,k]} \right) \le \tau.$$

Then, we have

$$\left\| \mathop{\mathbb{E}}_{\mu_1\otimes\mu_{[t+2,k]}} f''_{t+1} \otimes f''_{t+2} - \mathop{\mathbb{E}}_{\mu_{[t+1,k]}} f''_{t+1} \otimes f''_{t+2} \right\|_{\mathrm{op}} \le \tau,\,.$$

concluding the proof. ∎

[24] To clarify the notation, $\mathfrak{stu}$ and $\mathfrak{tu}$ are used to denote tuple concatenation.

Remark 5.7. *When $\ell = 1$, the above lemma (naturally) becomes the original scalar case from [JST21].*

We can iterate the preceding lemma to obtain the following.

Lemma 5.8 (Splittable Mixing Lemma Iterated [JST21])**.** *Suppose $W \subseteq [n]^k$ is a τ-splittable collection of tuples. For every $f = f_1 \otimes \cdots \otimes f_k \in \mathcal{F}_{k-1}$, we have*

$$\left| \mathop{\mathbb{E}}_{\nu_0} f \; - \; \mathop{\mathbb{E}}_{\nu_{k-1}} f \right| \; \leq \; (k-1) \cdot \tau.$$

In Section 5.3, we will need two corollaries of the splittable mixing lemma.

Claim 5.9 ([JST21])**.** *Let $W \subseteq [n]^k$ be a τ-splittable collection of tuples. Let $t \in \{0, \ldots, k-2\}$ and $h_{t+1} \in \mathcal{H}(R_0, R_1, \mathcal{F}_{t+1})$. For every $f \in \mathcal{F}_{t+1}$, we have*

$$\left| \langle h_{t+1}, f \rangle_{\nu_{t+1}} \; - \; \langle h_{t+1}, f \rangle_{\nu_t} \right| \; \leq \; \tau \cdot R_1.$$

Claim 5.10 ([JST21])**.** *Let $W \subseteq [n]^k$ be a τ-splittable collection of tuples. Let $t \in \{0, \ldots, k-2\}$ and $h_{t+1} \in \mathcal{H}(R_0, R_1, \mathcal{F}_{t+1})$. Then*

$$\left| \|h_{t+1}\|^2_{\nu_{t+1}} \; - \; \|h_{t+1}\|^2_{\nu_t} \right| \; \leq \; \tau \cdot R_1^2.$$

5.2 Existential Weak Regularity Decomposition

Using the abstract weak regularity lemma, Lemma 5.2, together splittable mixing lemmas of Section 5.1, we can obtain (non-constructive) existential weak regularity decompositions for splittable structures.

Lemma 5.11 (Existential Weak Regularity for Splittable Tuples)**.** *Let $W \subseteq [n]^k$ be a τ-splittable structure. Let $g \in \mathcal{R}^{W[1]^k}$ be supported on W with $\|g\|_{\mu_k} \leq 1$. Let $\mathcal{F} = \mathcal{F}_{k-1}$ (cf., Definition 4.12) be arbitrary. For every $\delta > 0$, if $\tau \leq O(\delta^2/(k-1))$, then there exists $h \in \mathcal{R}^{W[1]^k}$ supported on $O(1/\delta^2)$ functions in $\mathcal{F}$ such that*

$$\max_{f \in \mathcal{F}} \; |\langle g - h, f \rangle| \; \leq \; \delta \cdot |W|,$$

where the inner product is over the counting measure on $W[1]^k$.

5.3 Efficient Weak Regularity Decomposition

The goal of this section is to provide an efficient version of weak regularity that is sufficiently general to accommodate a decomposition with respect to $\mathcal{F}$ as one of the function classes: $\mathrm{CUT}^{\otimes k}$, $\mathrm{CUT}^{\otimes k}_{\pm}$, $\mathrm{CUT}^{\otimes k}_{\omega,q,a}$ and $\mathrm{CUT}^{\otimes k}_{\mathbb{U}_{s,\delta,k}}$, where the last two classes are new to this work and the former two classes were present in [JST21]. The main result of this section is the following efficient version of the weak regularity decomposition.

Theorem 5.12 (Efficient Weak Regularity (Extension of [JST21])). *Let $W \subseteq [n]^k$ be a τ-splittable collection of tuples. Suppose $\mathcal{F}$ is one of $\mathrm{CUT}^{\otimes k}$, $\mathrm{CUT}_{\pm}^{\otimes k}$, $\mathrm{CUT}_{\omega,q,a}^{\otimes k}$, for $q \geq 3$, or $\mathrm{CUT}_{\mathbb{U}_{s,k,\delta}}^{\otimes k}$. Let $\mathcal{R}$ be the domain of the functions in $\mathcal{F}$, when $k = 1$. Let $g \in \mathcal{R}^{W[1]^k}$ be supported on W with $\|g\|_{\mu_k} \leq 1$. For every $\delta > 0$, if $\tau \leq \delta^2/(k^3 \cdot 2^{20})$, then we can find $h = \sum_{\ell=1}^{p} c_\ell \cdot f_\ell$ with $p = O(k^2/\delta^2)$, scalars $c_1, \ldots, c_p \in \mathbb{K}$ and functions $f_1, \ldots, f_p \in \mathcal{F}$, such that $\|h\|_{\mu_1^{\otimes k}} \leq 2$, $\sum_{\ell=1}^{p} |c_\ell| = O(k/\delta)$ and h is a good approximator to g in the following sense*

$$\max_{f \in \mathcal{F}} \left| \left\langle g - \left(\frac{d}{n}\right)^{k-1} h, f \right\rangle \right| \leq \delta \cdot |W| \,,$$

where the inner product is over the counting measure on $W[1]^k$. Furthermore, h can be found in $\widetilde{O}(2^{|\mathcal{R}|^{\widetilde{O}(k^2/\delta^2)}} \cdot |W|)$ time in the scalar valued case and in time $\widetilde{O}_{s,k,\delta}(\mathrm{poly}(|W|))$, otherwise.

We now proceed to prove our main result in this section, namely Theorem 5.12. First, we establish some extra notation now. Let W be a d-regular collection of tuples. Most of our derivations, which are existential, hold for a generic $\mathcal{F}_t$ (*cf.*, Definition 4.12). The work [JST21] only derives near-linear time algorithmic results when $\mathcal{F}_t$ is either the CUT functions

$$\mathcal{F}_t^{0/1} := \left\{ \pm \mathbf{1}_{S_1} \otimes \cdots \otimes \mathbf{1}_{S_t} \otimes \mathbf{1}_T \mid S_j \subseteq W[1], T \subseteq W[t+1,k] \right\},$$

or "signed" CUT functions

$$\mathcal{F}_t^{\pm 1} := \left\{ \pm \chi_{S_1} \otimes \cdots \otimes \chi_{S_t} \otimes \chi_T \mid S_j \subseteq W[1], T \subseteq W[t+1,k] \right\},$$

where above we recall that for $S \subseteq [n]$, we have $\chi_S(i) = (-1)^{\mathbf{1}_{i \in S}}$ for $i \in [n]$. Observe that the condition $S_j \subseteq W[1]$ is equivalent to $S_j \subseteq W[i]$ since W is d-regular. In this work, we obtain new near-linear time weak regularity decomposition result for

$$\mathcal{F}_t^{\omega,q,a} := \left\{ \chi_{S_1,a} \otimes \cdots \otimes \chi_{S_t,a} \otimes \chi_{T,a} \mid b_j \in \mathbb{Z}_q^{W[1]}, T \in \mathbb{Z}_q^{W[t+1,k]} \right\},$$

and polynomial time algorithmic results for

$$\mathcal{F}_t^{\mathbb{U}_{s,\delta,k}} := \left\{ f_{S_1} \otimes \cdots \otimes f_{S_t} \otimes f_T \mid f_{S_j} \in \mathbb{U}_{s,\delta,k}^{W[1]}, f_T \in \mathbb{U}_{s,\delta,k}^{W[t+1,k]} \right\},$$

where $\mathbb{U}_{s,\delta,k}$ will be a suitable net over matrices in $M_s(\mathbb{C})$ of operator norm at most 1 depending on s, δ and k.

For quick reference, we collect the notation needed in our algorithmic weak regularity decomposition in the following table.

$\mathcal{F}_t := \left\{ \pm f_1 \otimes \cdots \otimes f_t \otimes f_{t+1} \mid f_j \subseteq \mathcal{R}^{W[1]} \text{ for } i \leq t, f_{t+1} \subseteq \mathcal{R}^{W[t+1,k]}, \|f_j\|_\infty \leq 1 \right\}$
$\mathcal{H}(R_0, R_1, \mathcal{F}) := \left\{ \sum_{\ell=1}^{p} c_\ell \cdot f_\ell \mid p \leq R_0, \sum |c_\ell| \leq R_1, f_\ell \in \mathcal{F} \right\}$
μ_1 is the uniform distribution on $W[1]$ and $\mu_{[t+1,k]}$ is the uniform distribution on $W[t+1,k]$
$\nu_t := (\mu_1)^{\otimes t} \otimes \left(\mu_{[t+1,k]}\right)$

Our main result of this section, namely, the near-linear time weak regularity decomposition Theorem 5.12, can be readily deduced from Lemma 5.13 below.

Lemma 5.13 (Efficient Weak Regularity Induction (Extension of [JST21])). *Let $W \subseteq [n]^k$ be a τ-splittable d-regular collection of tuples. Let $g \in \mathcal{F}_0$ and $t \in \{0, \dots, k-1\}$ with $\|g\|_{\mu_k} \leq 1$. For every $\delta > 0$, if $\tau \leq \delta^2/(k \cdot 2^{18})$, then there exists $h_t \in \mathcal{H}(O(1/\delta^2), 2^8(1+1/k)^t/\delta, \mathcal{F}_t)$ with $\|h_t\|_{\nu_t}^2 \leq (1+1/k)^t$ such that*

$$\max_{f \in \mathcal{F}_t} \left| \left\langle g - \left(\frac{d}{n}\right)^t h_t, f \right\rangle_{\nu_t} \right| \leq 2 \cdot \left(\frac{d}{n}\right)^t \cdot t \cdot \delta.$$

Furthermore, the function h_t can be found in $\widetilde{O}((2t)^{t|\mathcal{R}|^{O(1/\delta^2)}} \cdot |W|)$ time in the scalar valued case, and in time $\widetilde{O}_{s,k,\delta}(\mathrm{poly}(|W|))$ in the matrix valued case.

We restate Theorem 5.12 below and then prove it assuming Lemma 5.13.

Theorem 5.12 (Efficient Weak Regularity (Extension of [JST21])). *Let $W \subseteq [n]^k$ be a τ-splittable collection of tuples. Suppose $\mathcal{F}$ is one of $\mathrm{CUT}^{\otimes k}$, $\mathrm{CUT}_{\pm}^{\otimes k}$, $\mathrm{CUT}_{\omega,q,a}^{\otimes k}$, for $q \geq 3$, or $\mathrm{CUT}_{\mathbb{U}_{s,k,\delta}}^{\otimes k}$. Let $\mathcal{R}$ be the domain of the functions in $\mathcal{F}$, when $k = 1$. Let $g \in \mathcal{R}^{W[1]^k}$ be supported on W with $\|g\|_{\mu_k} \leq 1$. For every $\delta > 0$, if $\tau \leq \delta^2/(k^3 \cdot 2^{20})$, then we can find $h = \sum_{\ell=1}^p c_\ell \cdot f_\ell$ with $p = O(k^2/\delta^2)$, scalars $c_1, \dots, c_p \in \mathbb{K}$ and functions $f_1, \dots, f_p \in \mathcal{F}$, such that $\|h\|_{\mu_1^{\otimes k}} \leq 2$, $\sum_{\ell=1}^p |c_\ell| = O(k/\delta)$ and h is a good approximator to g in the following sense*

$$\max_{f \in \mathcal{F}} \left| \left\langle g - \left(\frac{d}{n}\right)^{k-1} h, f \right\rangle \right| \leq \delta \cdot |W|,$$

where the inner product is over the counting measure on $W[1]^k$. Furthermore, h can be found in $\widetilde{O}(2^{|\mathcal{R}|^{\widetilde{O}(k^2/\delta^2)}} \cdot |W|)$ time in the scalar valued case and in time $\widetilde{O}_{s,k,\delta}(\mathrm{poly}(|W|))$, otherwise.

Proof. Set $\mathcal{F}_t$ according to the choice of $\mathcal{F}$. We apply Lemma 5.13 with $t = k-1$, accuracy δ as $\delta/(2k)$ and input function g. This gives $h_t = \sum_{\ell=1}^p c'_\ell \cdot f_\ell \in \mathcal{H}(O(k^2/\delta^2), O(k/\delta), \mathcal{F}_t)$ such that

$$\max_{f \in \mathcal{F}_t} \left| \left\langle g - \left(\frac{d}{n}\right)^t h_t, f \right\rangle_{\nu_t} \right| \leq 2 \cdot \left(\frac{d}{n}\right)^t \cdot t \cdot \delta. \tag{3}$$

Note that $\nu_t = \nu_{k-1} = \mu_1^{\otimes k}$ is the uniform measure on $W[1]^k$. Since W is d-regular, $|W| = |W[1]|^k \cdot (d/n)^{k-1}$. Set $h = \cdot h_t$. Then the guarantee in Eq. (3) becomes

$$\max_{f \in \mathcal{F}} \left| \left\langle g - \left(\frac{d}{n}\right)^{k-1} h, f \right\rangle \right| \leq \delta \cdot |W|,$$

where the inner product is under the counting measure. By Lemma 5.13, we have $\|h_t\|_{\nu_t}^2 \leq (1+1/k)^t \leq e$, so $\|h_t\|_{\nu_t} \leq 2$. Then $\|h\|_{\mu_1^{\otimes k}} \leq 2$. The running time follows from Lemma 5.13 completing the proof. ∎

We now prove Lemma 5.13 above assuming the following algorithmic result which we prove later. The proof of Lemma 5.13 is almost the same as the corresponding one in [JST21] with the exception that we will need to use some absolute values to handle the case when the underlying field is $\mathbb{C}$.

Lemma 5.14 (Algorithmic Weak Regularity Step (Extension of [JST21]))**.** *Let $\delta > 0$ and $t \in \{0, \dots, k-2\}$. Let $h_t \in \mathcal{H}(O(B/\delta^2), O(B/\delta), \mathcal{F}_t)$ with $\|h_t\|_{\nu_t}^2 \leq B$. Then there exists $h_{t+1} \in \mathcal{H}(O(B/\delta^2), 2^8 B/\delta, \mathcal{F}_{t+1})$ with $\|h_{t+1}\|_{\nu_t}^2 \leq B$ such that*

$$\max_{f \in \mathcal{F}_{t+1}} \left| \langle h_t - h_{t+1}, f \rangle_{\nu_t} \right| \leq \delta.$$

Furthermore, each h_{t+1} can be found in time $\widetilde{O}((2t)^{t|\mathcal{R}|^{O(1/\delta^2)}} \cdot |W|)$ in the scalar valued case, and in time $\widetilde{O}_{s,k,\delta}(\mathrm{poly}(|W|))$ in the matrix valued case.

Proof of Lemma 5.13. We will prove the lemma with the following simple equivalent conclusion

$$\left| \left\langle g - \left(\frac{d}{n}\right)^t h_t, f \right\rangle_{\nu_t} \right| \leq 2 \cdot \left(\frac{d}{n}\right)^t \cdot t \cdot \delta \qquad \Leftrightarrow \qquad \left| \left\langle \left(\frac{n}{d}\right)^t g - h_t, f \right\rangle_{\nu_t} \right| \leq 2 \cdot t \cdot \delta,$$

which we will prove holds for every $f \in \mathcal{F}_t$. The base case $t = 0$ follows immediately by setting $h_0 = g$. Let $t \in \{0, \dots, k-2\}$. Since $h_t \in \mathcal{H}(O(1/\delta^2), 2^8(1+1/k)^t/\delta, \mathcal{F}_t)$, invoking Lemma 5.14 with accuracy parameter δ and input function h_t, we obtain $h_{t+1} \in \mathcal{H}(O(1/\delta^2), 2^8(1+1/k)^{t+1}/\delta, \mathcal{F}_{t+1})$ satisfying

$$\max_{f \in \mathcal{F}_{t+1}} \left| \langle h_t - h_{t+1}, f \rangle_{\nu_t} \right| \leq \delta. \tag{4}$$

Let $f \in \mathcal{F}_{t+1}$. We will show that h_{t+1} satisfies the conclusion of the lemma. Expanding we have

$$\begin{aligned} \left\langle \left(\frac{n}{d}\right)^{t+1} g - h_{t+1}, f \right\rangle_{\nu_{t+1}} &= \underbrace{\left\langle \left(\frac{n}{d}\right)^t g - h_t, f \right\rangle_{\nu_t}}_{(i)} + \left(\frac{n}{d}\right)^t \cdot \underbrace{\left(\frac{n}{d} \langle g, f \rangle_{\nu_{t+1}} - \langle g, f \rangle_{\nu_t} \right)}_{(ii)} \\ &\quad + \underbrace{\langle h_t - h_{t+1}, f \rangle_{\nu_t}}_{(iii)} + \underbrace{\langle h_{t+1}, f \rangle_{\nu_t} - \langle h_{t+1}, f \rangle_{\nu_{t+1}}}_{(iv)}. \end{aligned}$$

We will bound each of the terms in RHS above.

Term (i): Suppose $f = f_1 \otimes \cdots \otimes f_{t+1} \otimes f_{t+2} \in \mathcal{F}_{t+1}$. Let $f' = f_1 \otimes \cdots \otimes f_t \otimes f'_{t+1}$, where $f'_{t+1} = (f_{t+1} \otimes f_{t+2})|_{W[t+2,k]}$, so that $f' \in \mathcal{F}_t$. Using the induction hypothesis, we have

$$\left| \left\langle \left(\frac{n}{d}\right)^t g - h_t, f \right\rangle_{\nu_t} \right| = \left| \left\langle \left(\frac{n}{d}\right)^t g - h_t, f' \right\rangle_{\nu_t} \right| \leq 2 \cdot t \cdot \delta.$$

Term (ii): Since $g \in \mathcal{F}_0$, it is supported on W and so we have

$$\begin{aligned} \langle g, f \rangle_{\nu_t} &= \frac{1}{|W[1]|^t \, |W[t+1,k]|} \sum_{\mathfrak{s} \in W} \langle g(\mathfrak{s}), f(\mathfrak{s}) \rangle_{\mathcal{Y}} \\ &= \frac{n}{d} \cdot \frac{1}{|W[1]|^{t+1} \, |W[t+2,k]|} \sum_{\mathfrak{s} \in W} \langle g(\mathfrak{s}), f(\mathfrak{s}) \rangle_{\mathcal{Y}} = \frac{n}{d} \cdot \langle g, f \rangle_{\nu_{t+1}}. \end{aligned}$$

where the second equality follows from $|W[t+1,k]| = d \cdot |W[t+2,k]|$ by the d-regular assumption.

Term (iii): By Eq. (4), we have $\left|\langle h_t - h_{t+1}, f\rangle_{\nu_t}\right| \leq \delta$.

Term (iv): For notional convenience, set $R_1 = 2^8(1+1/k)^{t+1}/\delta$. Since $h_{t+1} \in \mathcal{H}(\infty, R_1, \mathcal{F}_{t+1})$ and the splittability parameter τ satisfies $\tau \leq \delta^2/(k \cdot 2^{18})$, from Claim 5.9 we obtain

$$\left|\langle h_{t+1}, f\rangle_{\nu_t} - \langle h_{t+1}, f\rangle_{\nu_{t+1}}\right| \leq \tau \cdot R_1 \leq \delta.$$

Putting everything together yields

$$\left|\left\langle \left(\frac{n}{d}\right)^{t+1} g - h_t, f\right\rangle_{\nu_{t+1}}\right| \leq \underbrace{2 \cdot t \cdot \delta}_{(i)} + \left(\frac{n}{d}\right)^t \cdot \underbrace{0}_{(ii)} + \underbrace{\delta}_{(iii)} + \underbrace{\delta}_{(iv)} \leq 2 \cdot (t+1) \cdot \delta,$$

concluding the claimed inequality.

Now we use the bound $\|h_{t+1}\|_{\nu_t}^2 \leq \|h_t\|_{\nu_t}^2$ from Lemma 5.14 together with the splittability assumption $\tau \leq \delta^2/(k \cdot 2^{18})$ to bound the norm $\|h_{t+1}\|_{\nu_{t+1}}^2$ under the new measure ν_{t+1}. Under these assumptions and using Claim 5.10 we get

$$\begin{aligned}\left|\|h_{t+1}\|_{\nu_{t+1}}^2 - \|h_{t+1}\|_{\nu_t}^2\right| \leq \tau \cdot R_1^2 &\leq \frac{\delta^2}{k \cdot 2^{18}} \cdot \frac{2^{16}(1+1/k)^{2(t+1)}}{\delta^2} \\ &\leq \frac{(1+1/k)^t}{k}.\end{aligned}$$

where we used the bounds on τ, R_1 and $(1+1/k)^{(t+2)} \leq 4$ for $0 \leq t \leq k-2$. From the previous inequality and the induction hypothesis $\|h_t\|_{\nu_t}^2 \leq (1+1/k)^t$, we finally get $\|h_{t+1}\|_{\nu_{t+1}}^2 \leq (1+1/k)^{t+1}$ as desired. ∎

We now show a near-linear time weak regularity decomposition for special functions of the form $h_t \in \mathcal{H}(O(1/\delta^2), O(1/\delta), \mathcal{F}_t)$ that admit a tensor product structure. The goal is to design a correlation oracle that exploits the special tensor product structure of the function $(h_t - h_{t+1}^{(\ell)})$, where $h_{t+1}^{(\ell)}$ is the ℓth approximator of h_t in the abstract weak regularity algorithm (*cf.*, Algorithm 5.3).

Lemma 5.14 (Algorithmic Weak Regularity Step (Extension of [JST21]))**.** *Let $\delta > 0$ and $t \in \{0, \ldots, k-2\}$. Let $h_t \in \mathcal{H}(O(B/\delta^2), O(B/\delta), \mathcal{F}_t)$ with $\|h_t\|_{\nu_t}^2 \leq B$. Then there exists $h_{t+1} \in \mathcal{H}(O(B/\delta^2), 2^8 B/\delta, \mathcal{F}_{t+1})$ with $\|h_{t+1}\|_{\nu_t}^2 \leq B$ such that*

$$\max_{f \in \mathcal{F}_{t+1}} \left|\langle h_t - h_{t+1}, f\rangle_{\nu_t}\right| \leq \delta.$$

Furthermore, each h_{t+1} can be found in time $\widetilde{O}((2t)^{t|\mathcal{R}|^{O(1/\delta^2)}} \cdot |W|)$ in the scalar valued case, and in time $\widetilde{O}_{s,k,\delta}(\mathrm{poly}(|W|))$ in the matrix valued case.

Similarly to [JST21], our correlation oracle for higher-order tensors will make calls to a correlation oracle for matrices (i.e., 2-tensors). The matrix oracles are presented in Section 6.1. We now provide the statement of each concrete oracle the framework can now handle. The first one for binary valued functions was given in [JST21].

Theorem 6.1 (Alon–Naor Correlation Oracle [JST21]). *Let $\mathcal{F}$ be either $\mathrm{CUT}^{\otimes 2}$ or $\mathrm{CUT}_{\pm}^{\otimes 2}$ and μ be the uniform measure supported on at most m elements of $[n'] \times [n']$. There exists an algorithmic $(\delta, \alpha_{\mathbf{AN}} \cdot \delta)$-correlation oracle $\mathcal{O}_{\mu,B}$ running in time $\mathcal{T}_{\mathcal{O}_{\mu,B}} = \tilde{O}\left(\mathrm{poly}(B/\delta) \cdot (m + n')\right)$, where $\alpha_{\mathbf{AN}} \geq 1/2^4$ is an approximation ratio constant.*

We now present two new correlation oracles for the framework. The first new correlation oracle is designed to handle the case involving roots of unity. It will be a near-linear time algorithm.

Theorem 6.4 (So–Zhang–Ye Correlation Oracle). *Let $\mathcal{F}$ be $\mathrm{CUT}^{\otimes 2}_{\omega,q,a}$ for some integer $q \geq 3$ and μ be the uniform measure supported on at most m elements of $[n'] \times [n']$. There exists an algorithmic $(\delta, \alpha_{\mathbf{SZY}} \cdot \delta)$-correlation oracle $\mathcal{O}_{\mu,B}$ running in time $\mathcal{T}_{\mathcal{O}_{\mu,B}} = \tilde{O}\left(\mathrm{poly}(B/\delta) \cdot (m + n')\right)$, where $\alpha_{\mathbf{SZY}} \geq 1/10$ is an approximation ratio constant.*

The second new correlation oracle is designed to handle the case involving representations. To make the analysis simpler in this more technical case, we only design a polynomial time oracle.

Theorem 6.13 (Naor–Regev–Vidick Correlation Oracle). *Let $\mathcal{F}$ be $\mathrm{CUT}^{\otimes 2}_{\mathbb{U}_{s,\delta,k}}$ and μ be the uniform measure supported on at most m elements of $[n'] \times [n']$. There exists an algorithmic $(\delta, \alpha_{\mathbf{NRV}} \cdot \delta)$-correlation oracle $\mathcal{O}_{\mu,B}$ running in time $\mathcal{T}_{\mathcal{O}_{\mu,B}} = \tilde{O}_{\delta,s,B}\left(\mathrm{poly}(m + n')\right)$, where $\alpha_{\mathbf{NRV}} \geq 1/4$ is the approximation ratio constant.*

Proof. We will apply the abstract weak regularity lemma, *cf.*,Lemma 5.2, with $\mathcal{F} = \mathcal{F}_{t+1}$, δ, $\delta' = \delta/2^8$ and $\mu = \nu_t$. This will result in a function from $\mathcal{H}(O(B/\delta^2), 2^8 B/\delta, \mathcal{F}_{t+1})$.

Correlation oracle task: To make this application take near-linear time, we need to specify a correlation oracle $\mathcal{O}_{\nu_t} = \mathcal{O}_{\nu_t, O(1)}$ and now we take advantage of the special tensor structure in our setting as done in [JST21]. We want an oracle that given

$$h_t = \sum_{\ell=1}^{p} c_\ell \cdot g_\ell, \quad g_\ell \in \mathcal{F}_t, \qquad g_\ell = g_{\ell,1} \otimes \cdots \otimes g_{\ell,t} \otimes \underbrace{g_{\ell,t+1}}_{\in \mathcal{R}^{W[t+1,k]}} \text{ and}$$

$$h_{t+1} = \sum_{\ell=1}^{p} c'_\ell \cdot g'_\ell, \quad g'_\ell \in \mathcal{F}_{t+1}, \quad g'_\ell = g'_{\ell,1} \otimes \cdots \otimes g'_{\ell,t} \otimes \underbrace{g'_{\ell,t+1}}_{\in \mathcal{R}^{W[1]}} \otimes \underbrace{g'_{\ell,t+2}}_{\in \mathcal{R}^{W[t+2,k]}},$$

if there exists

$$f = f_1 \otimes \cdots \otimes f_t \otimes \underbrace{f_{t+1}}_{\in \mathcal{R}^{W[1]}} \otimes \underbrace{f_{t+2}}_{\in \mathcal{R}^{W[t+2,k]}} \in \mathcal{F}_{t+1}$$

satisfying

$$\left| \langle h_t - h_{t+1}, f \rangle_{\nu_t} \right| \geq \delta,$$

for some $f \in \mathcal{F}_{t+1}$, finds $f' \in \mathcal{F}_{t+1}$ in near-linear time such that

$$\left| \langle h_t - h_{t+1}, f' \rangle_{\nu_t} \right| \geq \delta' = \frac{\delta}{2^8}.$$

Here, h_{t+1} is the current approximator of h_t in the abstract weak regularity algorithm and, by Lemma 5.2, $h_{t+1} \in \mathcal{H}(O(1/\delta^2), 2^8(1 + 1/k)^{t+1}/\delta, \mathcal{F}_{t+1})$.

For scalar valued functions, expanding $\langle h_t - h_{t+1}, f \rangle_{\nu_t}$ we get

$$\begin{aligned}\langle h_t - h_{t+1}, f \rangle_{\nu_t} \;=\; & \sum_{\ell=1}^{p} c_\ell \underbrace{\prod_{j=1}^{t} \langle g_{\ell,j}, f_j \rangle_{\mu_1}}_{\gamma_\ell} \cdot \langle g_{\ell,t+1}, f_{t+1} \otimes f_{t+2} \rangle_{\mu_{[t+1,k]}} \;- \\ & \sum_{\ell=1}^{p} c'_\ell \underbrace{\prod_{j=1}^{t} \left\langle g'_{\ell,j}, f_j \right\rangle_{\mu_1}}_{\gamma'_\ell} \cdot \langle g'_{\ell,t+1} \otimes g'_{\ell,t+2}, f_{t+1} \otimes f_{t+2} \rangle_{\mu_{[t+1,k]}} ,\end{aligned}$$

where we define γ_ℓ and γ'_ℓ as above.

For matrix valued functions, when expanding $\langle h_t - h_{t+1}, f \rangle_{\nu_t}$, we need to consider non-commutativity. Using the cyclic property of the trace (similarly to the proof of Lemma 5.6), we deduce that

$$\begin{aligned}\langle h_t - h_{t+1}, f \rangle_{\nu_t} \;=\; & \sum_{\ell=1}^{p} c_\ell \cdot \left\langle \underbrace{\mathop{\mathbb{E}}_{\mu_1^{\otimes k}} \left[\left(\otimes_{j=1}^{t} f_j \right)^\dagger \left(\otimes_{j=1}^{t} g_{\ell,j} \right) \right]}_{\gamma_\ell} g_{\ell,t+1}, f_{t+1} \otimes f_{t+2} \right\rangle_{\mu_{[t+1,k]}} \;- \\ & \sum_{\ell=1}^{p} c'_\ell \cdot \left\langle \underbrace{\mathop{\mathbb{E}}_{\mu_1^{\otimes k}} \left[\left(\otimes_{j=1}^{t} f_j \right)^\dagger \left(\otimes_{j=1}^{t} g'_{\ell,j} \right) \right]}_{\gamma'_\ell} g'_{\ell,t+1} \otimes g'_{\ell,t+2}, f_{t+1} \otimes f_{t+2} \right\rangle_{\mu_{[t+1,k]}} ,\end{aligned}$$

where we define γ_ℓ and γ'_ℓ as above.

Net Construction: Define Γ as the collection of sets $\{(\gamma_\ell, \gamma'_\ell)\}_{\ell \in [p]}$ as we consider all valid choices for $f_1, \ldots, f_t$. In the scalar case, we apply Lemma 5.15 (below) with parameter η to obtain an $(t \cdot \eta)$-net $\widetilde{\Gamma}$ for Γ. In the matrix valued case, we generate a sufficiently fine net for Γ so that we only incur an error proportional to δ. This can be done in time $\widetilde{O}_{s,\delta,k}(n)$ in an analogous ways as described in Section 3.4. Since we are not interested in the exact parameter trade-offs in this case, we omit the details.

Invoking the matrix correlation oracle: For each $\{(\gamma_\ell, \gamma'_\ell)\}_{\ell \in [p]} \in \widetilde{\Gamma}$, let

$$\mathsf{A} \coloneqq \sum_{\ell} \left(c_\ell \cdot \gamma_\ell \cdot g_{\ell,t+1} + c'_\ell \cdot \gamma'_\ell \cdot g'_{\ell,t+1} \otimes g'_{\ell,t+2} \right) .$$

We conveniently view A as a *sparse* matrix of dimension $|W[t+1]| \times |W[t+2,k]|$ with at most $|W[t+1,k]|$ non-zeros entries. Define $\varphi_{\mathsf{A}}(f_{t+1}, f_{t+2}) \coloneqq \left| \langle \mathsf{A}, f_{t+1} \otimes f_{t+2} \rangle_{\mu_{[t+1,k]}} \right|$. Define

$$\mathsf{OPT}(\mathsf{A}) \coloneqq \max_{f_{t+1}, f_{t+2}} \varphi_{\mathsf{A}}(f_{t+1}, f_{t+2}), \tag{5}$$

where f_{t+1}, f_{t+2} range over valid functions (again according to the kind of $\mathcal{F}_{t+1}$ we have).

Using the appropriate concrete correlation oracle, we can find functions $\tilde{f}_{t+1}, \tilde{f}_{t+2}$ such that $\varphi_{\widetilde{\mathsf{A}}}(\tilde{f}_{t+1}, \tilde{f}_{t+2}) \geq \alpha \cdot \delta/4$ since we are under the assumption that $\mathsf{OPT}(\mathsf{A}) \geq \delta$. We defer its details to Section 5.5.

Running Time: First, observe that with our choices of parameters the total number of configurations $\left|\widetilde{\Gamma}\right|$ is at most

$$\left|\widetilde{\Gamma}\right| \;\leq\; (1/\eta)^{t|\mathcal{R}|^{O(p)}} \;\leq\; (O(t/\delta^2))^{t|\mathcal{R}|^{O(p)}} \;\leq\; (2t)^{t|\mathcal{R}|^{O(1/\delta^2)}},$$

so that the correlation oracle $\mathcal{O}_{\nu_t}$ takes time at most

$$\left|\widetilde{\Gamma}\right| \cdot \mathcal{T}_{\mathsf{A}} \;\leq\; (2t)^{t|\mathcal{R}|^{O(1/\delta^2)}} \cdot \widetilde{O}(\mathrm{poly}(1/\delta) \cdot |W[t+1,k]|) \;=\; \widetilde{O}((2t)^{t|\mathcal{R}|^{O(1/\delta^2)}} \cdot |W[t+1,k]|).$$

Using the running time of the oracle $\mathcal{O}_{\nu_t}$, the total running time of the weak regularity decomposition follows from Lemma 5.2 which concludes the proof. ∎

5.4 Realizability Brute Force

We now describe how to generate approximately valid values of inner products in the following setting. We need to keep track of functions taking values in a more general finite range $\mathcal{R}$.

Lemma 5.15. *Let* $\{g_{\ell,1} \otimes \cdots \otimes g_{\ell,t}\}_{\ell\in[p]}$ *and* $\{g'_{\ell,1} \otimes \cdots \otimes g'_{\ell,t}\}_{\ell\in[p]}$*, where each* $g_{\ell,j}, g'_{\ell,j} \in \mathcal{R}^{W[1]}$*. Define*

$$\Gamma := \left\{ \left\{ \left(\prod_{j=1}^{t} \langle g_{\ell,j}, f_j \rangle_{\mu_1} , \prod_{j=1}^{t} \left\langle g'_{\ell,j}, f_j \right\rangle_{\mu_1} \right) \right\}_{\ell\in[p]} \mid f_1, \ldots, f_t \in \mathcal{R}^{W[1]} \right\},$$

to be the set of realizable pairs of inner products. Suppose $\mathcal{R}$ *is finite. For any* $\eta \in (0, 1/2)$*, we can find an* $(t\cdot\eta)$*-net* $\widetilde{\Gamma}$ *for* Γ *in* ℓ_∞*-norm with* $\left|\widetilde{\Gamma}\right| \leq (1/\eta)^{t|\mathcal{R}|^{O(p)}}$.

Proof. Let $\mathcal{B}_j$ be the factor generated by $\{g_{\ell,j}, g'_{\ell,j}\}_{\ell\in[p]}$. By Proposition 4.8, we have that $\langle g_{\ell,j}, f_j \rangle_{\mu_1} = \langle g_{\ell,j}, \mathbb{E}[f_j|\mathcal{B}_j] \rangle_{\mu_1}$ and $\left\langle g'_{\ell,j}, f_j \right\rangle_{\mu_1} = \left\langle g'_{\ell,j}, \mathbb{E}[f_j|\mathcal{B}_j] \right\rangle_{\mu_1}$. Note that for each atom σ in $\mathcal{B}_j$, the values taken by f_j on σ give rise to a distribution on $\mathcal{R}$. Suppose we have an η-approximation in ℓ_1-norm to each such distribution on each atom σ of $\mathcal{B}_j$. We denote by $\mathcal{D}_{\mathcal{B}_j(x)}$ this η-approximate distribution on atom of containing x. Let $\overline{f}_j(x) = \mathbb{E}[f_j|\mathcal{B}_j]_{\mu_1}(x)$ and $\widetilde{f}_j(x) = \mathbb{E}_{x'\sim\mathcal{D}_{\mathcal{B}_j(x)}}[f_j(x')]$. By Hölder's inequality, we have $\left\|\overline{f}_j - \widetilde{f}_j\right\|_\infty \leq \eta$. Applying Hölder's inequality once more, we obtain

$$\left| \left\langle g_{\ell,j}, \overline{f}_j \right\rangle_{\mu_1} - \left\langle g_{\ell,j}, \widetilde{f}_j \right\rangle_{\mu_1} \right|, \left| \left\langle g'_{\ell,j}, \overline{f}_j \right\rangle_{\mu_1} - \left\langle g'_{\ell,j}, \widetilde{f}_j \right\rangle_{\mu_1} \right| \leq \eta.$$

There are at most $(1/(\eta\,|\mathcal{R}|))^{O(|\mathcal{R}|)}$ η-approximate (in ℓ_1-norm) distribution on each atom[25]. Thus, we have at most $(1/(\eta\,|\mathcal{R}|))^{O(|\mathcal{R}||\mathcal{B}_j|)}$ choices for each $j \in [t]$. Therefore, we can bound $\left|\widetilde{\Gamma}\right|$ as

$$\left|\widetilde{\Gamma}\right| \leq (1/(\eta\,|\mathcal{R}|))^{O(\sum_{j=1}^{t}|\mathcal{R}||\mathcal{B}_j|)} \leq (1/\eta)^{t|\mathcal{R}|^{O(p)}}.$$

By triangle inequality, we have

$$\left| \prod_{j=1}^{t} \langle g_{\ell,j}, f_j \rangle_{\mu_1} - \prod_{j=1}^{t} \left\langle g_{\ell,j}, \widetilde{f}_j \right\rangle_{\mu_1} \right|, \left| \prod_{j=1}^{t} \left\langle g'_{\ell,j}, f_j \right\rangle_{\mu_1} - \prod_{j=1}^{t} \left\langle g'_{\ell,j}, \widetilde{f}_j \right\rangle_{\mu_1} \right| \leq t\cdot\eta,$$

concluding the proof. ∎

[25] If the atom is smaller than $(1/(\eta\,|\mathcal{R}|))$, then we consider all exact distributions to make sure that they are realizable.

5.5 Invoking Concrete Matrix Correlation Oracles

In this section, we show how to invoke the concrete correlation oracle.

Invoking the matrix correlation oracle:

We start by considering the scalar valued case. The specific concrete correlation oracle will vary according to the class of functions and will need include the oracle Theorem 6.4 for roots of unity. The fine error computations will be essentially the same as in [JST21]. In the computation of $\mathsf{OPT}(\mathsf{A})$, we have incurred so far an additive error of at most

$$4 \cdot t \cdot \eta \cdot \sum_{\ell} (|c_\ell| + |c'_\ell|).$$

Let $\widetilde{\mathsf{A}}$ be obtained from A by zeroing out all entries of absolute value smaller than $\delta/8$. Note that $\mathsf{OPT}(\widetilde{\mathsf{A}}) \geq \mathsf{OPT}(\mathsf{A}) - \delta/8$ and the absolute value of the entries of $\widetilde{\mathsf{A}}$ lie $[\delta/8, O(1/\delta)]$. For each entry of A, we compute a rational approximation $\pm P/Q$ where $Q = \Theta(1/\delta)$ and $P \in [1, O(1/\delta)]$ (to the real and imaginary parts if $\mathbb{K} = \mathbb{C}$) obtaining $\widetilde{\mathsf{A}}'$ such that

$$\mathsf{OPT}(\widetilde{\mathsf{A}}') \;\geq\; \mathsf{OPT}(\widetilde{\mathsf{A}}) - \delta/8 \;\geq\; \mathsf{OPT}(\widetilde{\mathsf{A}}) \;\geq\; \mathsf{OPT}(\mathsf{A}) - \delta/4.$$

Let α be approximation guarantee of the correlation oracle for $\mathcal{F}$, namely, from either Theorem 6.1 or Theorem 6.4 depending on $\mathcal{F}$. Using this concrete correlation oracle with accuracy parameter $\delta/4$ and input matrix $\widetilde{\mathsf{A}}'$, we obtain in $\mathcal{T}_{\mathsf{A}} := \widetilde{O}(\mathrm{poly}(1/\delta) \cdot |W[t+1,k]|)$ time in the scalar case and time $\mathcal{T}_{\mathsf{A}} := \widetilde{O}_{\delta,s}(\mathrm{poly}(|W[t+1,k]|))$ otherwise, with an extra additive error of $\delta/4$ and a multiplicative guarantee of α, a 2-tensor $\tilde{f}_{t+1} \otimes \tilde{f}_{t+2}$ satisfying

$$\varphi_{\widetilde{\mathsf{A}}}(\tilde{f}_{t+1}, \tilde{f}_{t+2}) \;\geq\; \alpha \cdot \left(\mathsf{OPT}(\mathsf{A}) \;-\; 2 \cdot \frac{\delta}{4} \;-\; 4 \cdot t \cdot \eta \cdot \sum_{\ell} (|c_\ell| + |c'_\ell|) \right).$$

Since $h_t \in \mathcal{H}(O(1/\delta^2), 2^8 \cdot (1+1/k)^t/\delta, \mathcal{F}_t)$ and $h_{t+1} \in \mathcal{H}(O(1/\delta^2), 2^8 \cdot (1+1/k)^{t+1}/\delta, \mathcal{F}_{t+1})$, we have $\sum_\ell (|c_\ell| + |c'_\ell|) \leq 2^{10}/\delta$ and $p = O(1/\delta^2)$. By choosing $\eta \leq O(\delta^2/t)$ appropriately, we can bound

$$4 \cdot t \cdot \eta \cdot \sum_{\ell} (|c_\ell| + |c'_\ell|) \;\leq\; 4 \cdot t \cdot \frac{2^{10}}{\delta} \cdot \eta \;\leq\; \frac{\delta}{4}.$$

Hence, $\varphi_{\widetilde{\mathsf{A}}}(\tilde{f}_{t+1}, \tilde{f}_{t+2}) \geq \alpha \cdot \delta/4$ since we are under the assumption that $\mathsf{OPT}(\mathsf{A}) \geq \delta$.

For the matrix valued case, since we are only interested in polynomial running time. We can invoke Theorem 6.13 and again find suitable (now matrix valued) $\tilde{f}_{t+1}$ and $\tilde{f}_{t+2}$. $\varphi_{\widetilde{\mathsf{A}}}(\tilde{f}_{t+1}, \tilde{f}_{t+2}) \geq \alpha \cdot \delta/4$ since we are under the assumption that $\mathsf{OPT}(\mathsf{A}) \geq \delta$.

6 Concrete Correlation Oracles

We now describe the concrete correlation oracles for matrices (2-tensors). We start by recalling the near-linear time correlation oracles for $\mathrm{CUT}^{\otimes 2}$ and $\mathrm{CUT}^{\otimes 2}_{\pm}$ of [JST21] in Section 6.1. Suitable generalizations of these oracles will be useful in the near-linear time oracle $\mathrm{CUT}^{\otimes 2}_{\omega,q,a}$ in Section 6.2. Next, in Section 6.3, we describe a (relaxed) polynomial time oracle for representations.

6.1 Grothendieck Problem over Boolean Variables

We now recall the near-linear time correlation oracle, Theorem 6.1 below, for $\mathrm{CUT}^{\otimes 2}$ and $\mathrm{CUT}^{\otimes 2}_{\pm}$ from [JST21]. They combine the constant factor approximation algorithms of Alon–Naor [AN04] for $\|\mathsf{A}\|_{\infty\to 1}$ and $\|\mathsf{A}\|_{\square}$ based on semi-definite programming (SDP) with the faster SDP solvers for sparse matrices such as those by Arora and Kale [AK07].

Theorem 6.1 (Alon–Naor Correlation Oracle [JST21])**.** *Let $\mathcal{F}$ be either $\mathrm{CUT}^{\otimes 2}$ or $\mathrm{CUT}^{\otimes 2}_{\pm}$ and μ be the uniform measure supported on at most m elements of $[n'] \times [n']$. There exists an algorithmic $(\delta, \alpha_{\mathbf{AN}} \cdot \delta)$-correlation oracle $\mathcal{O}_{\mu,B}$ running in time $\mathcal{T}_{\mathcal{O}_{\mu,B}} = \tilde{O}\left(\mathrm{poly}(B/\delta) \cdot (m + n')\right)$, where $\alpha_{\mathbf{AN}} \geq 1/2^4$ is an approximation ratio constant.*

Theorem 6.1 is a simple consequence of the following theorem.

Theorem 6.2 ([JST21])**.** *Let $\mathsf{A} \in \mathbb{R}^{n\times n}$ be a matrix of integers with at most m non-zero entries. Let $\delta \in (0, 2^{-5}]$ be an accuracy parameter. Suppose that*

$$\mathsf{OPT} := \max_{x_i, y_i \in \{\pm 1\}} \sum_{i,j=1}^{n} \mathsf{A}_{i,j} x_i y_j \;\geq\; \delta \cdot m.$$

Then, with high probability,i.e., $o_n(1)$, we can find, in $\tilde{O}\left(\mathrm{poly}(\|\mathsf{A}\|_{\infty}/\delta) \cdot (m+n)\right)$ time, vectors $\tilde{x}, \tilde{y} \in \{\pm 1\}^n$ such that

$$\sum_{i,j=1}^{n} \mathsf{A}_{i,j} \tilde{x}_i \tilde{y}_j \;\geq\; \frac{1}{4} \cdot \mathsf{OPT},$$

and find sets $\tilde{S}, \tilde{T} \subseteq [n]$ such that

$$\left| \sum_{i\in\tilde{S}, j\in\tilde{T}} \mathsf{A}_{i,j} \right| \;\geq\; \frac{1}{2^4} \cdot \|\mathsf{A}\|_{\square},$$

where $\|\mathsf{A}\|_{\square}$ is the cut norm of A.

We will need a sparse SDP solver capable of working over the real or the complex field (the latter for $\mathrm{CUT}^{\otimes 2}_{\omega,q,a}$ in Section 6.2). We will use the sparse solver from [AK07] due to its simple structure and ready generalization to the complex field (see [Kal07] for comments about the complex setting). It will be convenient to have the SDP solver wrapper below. Since the way of obtaining this wrapper from [AK07] is analogous to [JST21], we omit the details.

Lemma 6.3 (Sparse SDP Solver Wrapper based on [AK07] with similar statement to [LP20])**.** *Let $\mathbb{K}$ be either $\mathbb{R}$ or $\mathbb{C}$. Let $\mathsf{C} \in \mathbb{K}^{n\times n}$ be a matrix with at most m non-zero entries that is symmetric if $\mathbb{K} = \mathbb{R}$ and Hermitian if $\mathbb{K} = \mathbb{C}$. For every accuracy $\gamma > 0$, with high probability we can find in time $\widetilde{O}((m+n)/\mathrm{poly}(\gamma))$ vectors $u_1, \ldots, u_n \in \mathbb{K}^n$ in the unit ball (i.e., $\|u_i\| \leq 1$) such that the matrix $\widetilde{\mathsf{X}}_{i,j} := \langle u_i, u_j \rangle$ satisfies*

$$\mathrm{Tr}\left(\mathsf{C} \cdot \widetilde{\mathsf{X}}\right) \;\geq\; \max_{\mathsf{X}\succeq 0, X_{i,i}\leq 1} \mathrm{Tr}\left(\mathsf{C} \cdot \mathsf{X}\right) - \gamma \sum_{i,j} \left|\mathsf{C}_{i,j}\right|,$$

where the maximum is over real PSD matrices if $\mathbb{K} = \mathbb{R}$ or complex PSD matrices if $\mathbb{K} = \mathbb{C}$

6.2 Grothendieck Problem over Primitive Roots of Unity

To obtain the savings from the *alphabet reduction* for k-LIN over q-ary alphabet $\mathbb{Z}_q$ (as discussed in Section 3.3), we will need a version of the Grothendieck problem in which variables taking values in ± 1 are replaced by variables taking roots of unity values. Recall that the original Grothendieck problem over ± 1 variables can be equivalently phrased as

$$\max_{u,v \in \mathbb{Z}_2^n} \sum_{i,j} \mathsf{A}_{i,j} \cdot \chi(u_i) \cdot \chi(v_j) \,,$$

where A is a real matrix and χ is the (unique) non-trivial character of $\mathbb{Z}_2$.The Grothendieck problem for roots is defined as follows

$$\max_{u,v \in \mathbb{Z}_q^n} \left| \sum_{i,j} \mathsf{A}_{i,j} \cdot \chi(u_i) \cdot \chi(v_j) \right| \,, \tag{6}$$

where A is a complex matrix and χ is any non-trivial character of $\mathbb{Z}_q$ (we will need to consider all non-trivial characters).

In [SZY07], So, Zhang and Ye consider the Grothendieck problem for Hermitian positive semidefinite (PSD) matrix A (also known as the little Grothendieck problem). We will need to consider general Hermitian matrices not necessarily PSD. Using similar consideration from Alon and Naor [AN04] and borrowing from part of the analysis present in [SZY07], we will be able to accomplish this (simple) extension. Then, it will allows us to obtain a similar identity to Fact 6.12 from Alon and Naor [AN04]. With this identity and similar considerations to those in Section 6.1, we will deduce the following near-linear time correlation oracle for roots of unity.

Theorem 6.4 (So–Zhang–Ye Correlation Oracle)**.** *Let $\mathcal{F}$ be $\mathrm{CUT}^{\otimes 2}_{\omega,q,a}$ for some integer $q \geq 3$ and μ be the uniform measure supported on at most m elements of $[n'] \times [n']$. There exists an algorithmic $(\delta, \alpha_{\mathbf{SZY}} \cdot \delta)$-correlation oracle $\mathcal{O}_{\mu,B}$ running in time $\mathcal{T}_{\mathcal{O}_{\mu,B}} = \tilde{O}\left(\mathrm{poly}(B/\delta) \cdot (m + n')\right)$, where $\alpha_{\mathbf{SZY}} \geq 1/10$ is an approximation ratio constant.*

From the (Rietz) rounding method of So, Zhang and Ye [SZY07], we will need the following functions. For $u, r \in \mathbb{C}^d$, (we tweak their choice using some real $\eta > 0$)

$$r_u(g) \coloneqq \langle u, g \rangle - \eta \cdot \frac{2\sqrt{\pi}}{q \sin(\pi/q)} \mathrm{round}(\langle u, g \rangle) \,, \tag{7}$$

where the rounding function $\mathrm{round} \colon \mathbb{C} \to \mathbb{C}$ is defined as

$$\mathrm{round}(z) \coloneqq \begin{cases} 1 & \text{if } \arg(z) \in [-\pi/q, \pi/q) \\ \omega & \text{if } \arg(z) \in [\pi/q, 3\pi/q) \\ \vdots & \vdots \\ \omega^{q-1} & \text{if } \arg(z) \in [(2q-3)\pi/q, (2q-1)\pi/q) \end{cases} . \tag{8}$$

They show the following identity when g is a complex Gaussian random vector.

Lemma 6.5 (Implicit in [SZY07])**.**

$$\mathbb{E}_g \left[r_u(g)^* r_v(g) \right] = (1 - 2\eta) \langle u, v \rangle + \eta^2 \frac{4\pi}{(q \sin(\pi/q))^2} \mathbb{E}_g \left[\mathrm{round}(\langle u, g \rangle)^* \mathrm{round}(\langle v, g \rangle) \right] .$$

Corollary 6.6. *For a unit vector u, we have*

$$\mathbb{E}_g\left[r_u(g)^* r_u(g)\right] = (1-2\eta) + \eta^2 \frac{4\pi}{(q\sin(\pi/q))^2}\,.$$

Lemma 6.7 (Equivalent form of Lemma 6.5)**.**

$$\mathbb{E}_g\left[\mathrm{round}(\langle u,g\rangle)^* \mathrm{round}(\langle v,g\rangle)\right] = \frac{(q\sin(\pi/q))^2}{4\pi}\frac{(2\eta-1)}{\eta^2}\left(\langle u,v\rangle + \frac{\eta^2}{(2\eta-1)}\mathbb{E}_g\left[r_u(g)^* r_v(g)\right]\right).$$

For a similar strategy of Alon–Naor to work for the full roots of unity Grothendieck problem, we need

$$\eta^2/(2\eta-1)\left[(1-2\eta) + \eta^2\frac{4\pi}{(q\sin(\pi/q))^2}\right] < 1\,,$$

or equivalently

$$\gamma_{\eta,q} := -\eta^2 + \frac{\eta^4}{(2\eta-1)}\frac{4\pi}{(q\sin(\pi/q))^2} < 1\,.$$

For simplicity, choosing $\eta = 1/\sqrt{2}$ makes the value above strictly smaller than 1 for every $q \geq 3$ and this will be enough for a constant factor approximation. We proceed to formalize this claim.

Analogously to the real case of Section 6.1 where we defined a symmetric matrix C from the real A. To approximate Eq. (6) we will define two Hermitian matrices $\mathsf{C}_\Re$ and $\mathsf{C}_\Im$ from the complex A (to capture the real and the imaginary part) as follows

$$\mathsf{C}_\Re = \frac{1}{2}\begin{pmatrix} 0 & A \\ A^\dagger & 0\end{pmatrix} \text{ and } \mathsf{C}_\Im = \frac{1}{2}\begin{pmatrix} 0 & -iA \\ iA^\dagger & 0\end{pmatrix}.$$

To make this precise, we will need the following observations.

$$\begin{pmatrix} x^\dagger & y^\dagger\end{pmatrix}\begin{pmatrix} 0 & A \\ A^\dagger & 0\end{pmatrix}\begin{pmatrix} x \\ y\end{pmatrix} = x^\dagger A y + y^\dagger A^\dagger x = 2\Re(x^\dagger A y)$$

$$\begin{pmatrix} x^\dagger & y^\dagger\end{pmatrix}\begin{pmatrix} 0 & -iA \\ iA^\dagger & 0\end{pmatrix}\begin{pmatrix} x \\ y\end{pmatrix} = x^\dagger(-iA)y + y^\dagger(iA^\dagger)x = 2\Re(x^\dagger(-iA)y) = 2\Im(x^\dagger A y)\,.$$

Note that $\left|\sum_{i,j} A_{i,j}\cdot x_i \cdot y_j\right|^2 = \Re(x^\dagger A y)^2 + \Im(x^\dagger A y)^2$. Thus, we have

$$\max\left\{\left|\Re(x^\dagger A y)\right|, \left|\Im(x^\dagger A y)\right|\right\} \geq \frac{1}{\sqrt{2}}\left|\sum_{i,j} A_{i,j}\cdot x_i \cdot x_j\right|.$$

Then, we optimize using $\mathsf{C} = \mathsf{C}_\Re$ and $\mathsf{C} = \mathsf{C}_\Im$ and taking the maximum, we still get a constant $1/\sqrt{2}$ factor approximation, which is enough for our applications.

It will be convenient to phrase the rounding schemes for Grothendieck problems based on the Rietz method using the following terminology of a rounding *scheme*. This language also accommodate the Alon–Naor rounding scheme [AN04].

Scheme 6.8 (Rietz Rounding Scheme). *Suppose there exist constants $\alpha, \beta, \gamma > 0$, a Rietz function $r\colon \mathbb{K}^d \times \mathbb{K}^d \to \mathbb{K}$ and a rounding function* round$\colon \mathbb{K} \to \mathbb{K}$ *such that for every unit vectors in ℓ_2-norm $u, w \in \mathbb{K}^d$, we have*

$$\mathbb{E}\left[\text{round}\langle u, g\rangle^* \text{round}\langle w, g\rangle\right] \;=\; \alpha\left(\langle u, w\rangle \;+\; \beta \cdot \mathbb{E}\left[r_u(g)^* r_w(g)\right]\right), \tag{9}$$

and $\beta \cdot \mathbb{E}\left[|r_u(g)|^2\right] \leq \gamma < 1$, where the expectations are taken with respect to a random Gaussian vector $g \sim N(0, \mathsf{I}_d)$ in $\mathbb{K}^d$.

Lemma 6.9 (SZY as a Rietz Rounding Scheme). *Let $q \geq 3$ be an integer. Let $\eta > 1/2$ and define*

$$\alpha = \frac{(q\sin(\pi/q))^2}{4\pi}\frac{(2\eta-1)}{\eta^2}, \quad \beta = \frac{\eta^2}{(2\eta-1)} \quad \text{and} \quad \gamma = -\eta^2 + \frac{\eta^4}{(2\eta-1)}\frac{4\pi}{(q\sin(\pi/q))^2}.$$

The Rietz function Eq. (7) and the rounding function Eq. (8) form an Rietz rounding scheme with $\Sigma = \mathbb{S}_q^1$ parameters α, β and γ as above provided $\gamma < 1$.

We can now provide the fast algorithm for Rietz schemes in the sparse regime as follows.

Theorem 6.10 (Rietz Scheme Fast Algorithm). *Suppose that we have a Rietz scheme with parameters $\alpha, \beta, \gamma > 0$ and alphabet Σ. Let $\mathsf{A} \in \mathbb{K}^{n\times n}$ be a matrix with at most m non-zero entries. Let $\delta \in (0, 2^{-5}]$ be an accuracy parameter. Suppose that*

$$\mathsf{OPT} := \max_{x_i, y_i \in \Sigma} \left|\sum_{i,j=1}^{n} \mathsf{A}_{i,j} x_i y_j\right| \;\geq\; \delta \cdot m.$$

Then, with high probability,i.e., $o_n(1)$, we can find, in $\tilde{O}\left(\text{poly}(\|\mathsf{A}\|_\infty/\delta)\cdot(m+n)\right)$ time, vectors $\tilde{x}, \tilde{y} \in \Sigma^n$ such that

$$\left|\sum_{i,j=1}^{n} \mathsf{A}_{i,j}\tilde{x}_i\tilde{y}_j\right| \;\geq\; \frac{1}{\sqrt{2}} \cdot \alpha \cdot (1-\gamma-\delta) \cdot \mathsf{OPT}.$$

Furthermore, if $\mathbb{K} = \mathbb{R}$, then the above factor $1/\sqrt{2}$ can be replaced by 1.

Proof of Theorem 6.10. We now combine the approximation algorithms for scalar Grothendieck problems satisfying the assumptions of Rietz rounding scheme with a near-linear time sparse SDP solver. We need to argue that this indeed leads to the claimed approximation guarantees while being computable in near-linear time overall. Using the input matrix A, we will consider

$$\mathsf{C} = \mathsf{C}_{\Re} = \frac{1}{2}\begin{pmatrix} 0 & \mathsf{A} \\ \mathsf{A}^\dagger & 0 \end{pmatrix}.$$

The case of $\mathsf{C} = \mathsf{C}_{\Im}$ will be analogous. The SDP relaxation for the scalar Grothendieck problem becomes

$$\begin{aligned} \max \quad & \text{Tr}(\mathsf{C}\cdot\mathsf{X}) && =: \mathsf{SDP}^* \\ \text{s.t.} \quad & \mathsf{X}_{i,i} \;\leq\; 1 && \forall i \in [2n] \\ & \mathsf{X} \;\succeq\; 0, \end{aligned}$$

except for the constraints $X_{i,i} \leq 1$ which they instead take to be $X_{i,i} = 1$. This technical difference will play a (small) role in the rounding of this SDP since the Rietz method analysis relies on Gram vectors of X being on the unit sphere. Moreover, we will be solving this SDP within only a weak additive approximation guarantee[26]. Although these technical differences need to be handled, this will be simple to do.

Applying the solver of Lemma 6.3 with accuracy parameter $\gamma = \delta^2 / \|\mathsf{A}\|_\infty$ to the above SDP, we obtain in $\widetilde{O}(\mathrm{poly}(\|\mathsf{A}\|_\infty / \delta) \cdot (m+n))$ time vectors $u_1, \ldots, u_{2n} \in \mathbb{K}^{2n}$ in the unit ball so that the matrix $\widetilde{\mathsf{X}}_{i,j} := \langle u_i, u_j \rangle$ satisfy

$$\mathrm{Tr}\left(\mathsf{C} \cdot \widetilde{\mathsf{X}}\right) \;\geq\; \max_{\mathsf{X} \succeq 0, X_{i,i} \leq 1} \mathrm{Tr}\left(\mathsf{C} \cdot \mathsf{X}\right) \;-\; \delta^2 \cdot m.$$

By assumption, we have $\mathsf{SDP}^* := \max_{\mathsf{X} \succeq 0, X_{i,i} \leq 1} \mathrm{Tr}\left(\mathsf{C} \cdot \mathsf{X}\right) \geq \mathsf{OPT} \geq \delta \cdot m$, in which case the above guarantee becomes

$$\mathrm{Tr}\left(\mathsf{C} \cdot \widetilde{\mathsf{X}}\right) \;\geq\; (1-\delta) \cdot \mathsf{SDP}^*.$$

To obtain diagonal entries equal to 1 in our SDP solution we simply consider the new SDP solution $\widetilde{\mathsf{X}}' = \widetilde{\mathsf{X}} + \Lambda$, where Λ is the diagonal matrix defined as $\Lambda_{i,i} := 1 - \widetilde{\mathsf{X}}_{i,i}$. Gram vectors $u'_1, \ldots, u'_{2n}$ of $\widetilde{\mathsf{X}}'$ can be obtained in near-linear time from $u_1, \ldots, u_{2n}$ and Λ by setting

$$u'_i := u_i \oplus \sqrt{\Lambda_{i,i}} \cdot e_i \;\in\; \mathbb{K}^{2m} \oplus \mathbb{K}^{2m},$$

where $e_i \in \mathbb{K}^{2m}$ has a one at the ith position and zero everywhere else. Observe that for our particular C, we have

$$\mathrm{Tr}\left(\mathsf{C} \cdot \widetilde{\mathsf{X}}'\right) \;=\; \mathrm{Tr}\left(\mathsf{C} \cdot \widetilde{\mathsf{X}}\right).$$

This rounding consists in sampling a Gaussian vector $g \sim N(0, \mathsf{I}_d)$ and setting $\widetilde{x}_i := \mathrm{round}\langle u'_i, g \rangle$ and $\widetilde{y}_{i+n} := \mathrm{round}\langle u'_{i+n}, g \rangle$ for $i \in [n]$. To analyze the approximation guarantee, the following identity is used.

Using Eq. (9) of the Rietz rounding scheme, the expected value of the rounding, i.e.,

$$\mathbb{E}\left[\sum_{i,j} \mathsf{A}_{i,j} \mathrm{round}\langle u'_i, g \rangle^* \mathrm{round}\left\langle u'_{j+n}, g \right\rangle\right],$$

can be expressed as

$$\alpha \cdot \left(\sum_{i,j} \mathsf{A}_{i,j} \left\langle u'_i, u'_{j+n} \right\rangle + \beta \cdot \sum_{i,j} \mathsf{A}_{i,j} \mathbb{E}\left[r_{u'_i}(g)^* r_{u'_{j+n}}(g)\right]\right).$$

Note that $\beta \cdot \mathbb{E}\left[|r_u(g)|^2\right] \leq \gamma < 1$ by the Rietz scheme property. Then, in our setting we obtain

$$\begin{aligned}
\mathbb{E}\left[\sum_{i,j} \mathsf{A}_{i,j} \mathrm{round}\langle u'_i, g \rangle^* \mathrm{round}\left\langle u'_{j+n}, g \right\rangle\right] \;&\geq\; \alpha \cdot ((1-\delta) \cdot \mathsf{SDP}^* - \gamma \cdot \mathsf{SDP}^*) \\
&\geq\; \alpha \cdot (1 - \delta - \gamma) \cdot \mathsf{SDP}^*,
\end{aligned}$$

[26] This may not be sufficient to obtain $X_{i,i} \approx 1$ by an extremality argument

as claimed. By considering the largest between the real and imginary programs (related to $C_{\Re}$ and $C_{\Im}$, resp.), we lose an addition $1/\sqrt{2}$ factor in the objective value.

Using standard techniques, this guarantee on the expected value of the rounded solution can be used to give with high probability a final guarantee of $(1/\sqrt{2})\alpha \cdot (1 - 2\delta - \gamma) \cdot \mathsf{OPT}$ (namely, by repeating this rounding scheme $O_\alpha(\log(1/\delta) \cdot \log(n))$ times). ∎

We now recall and prove the correlation oracle for roots of unity.

Theorem 6.4 (So–Zhang–Ye Correlation Oracle)**.** *Let $\mathcal{F}$ be $\mathrm{CUT}^{\otimes 2}_{\omega,q,a}$ for some integer $q \geq 3$ and μ be the uniform measure supported on at most m elements of $[n'] \times [n']$. There exists an algorithmic $(\delta, \alpha_{\mathbf{SZY}} \cdot \delta)$-correlation oracle $\mathcal{O}_{\mu,B}$ running in time $\mathcal{T}_{\mathcal{O}_{\mu,B}} = \tilde{O}\left(\mathrm{poly}(B/\delta) \cdot (m + n')\right)$, where $\alpha_{\mathbf{SZY}} \geq 1/10$ is an approximation ratio constant.*

Proof. We invoke the algorithm Theorem 6.10 with $\delta = 2^5$ using the Rietz rounding scheme of Lemma 6.9 with $\eta = 1/\sqrt{2}$. Note that $q \sin(\pi/q)$ is a growing function for $q \geq 2$ and its values lie in the interval $[2.598, \pi]$ for $q \geq 3$. With our choice for η, we obtain $(1 - \gamma_{\eta,q}) \geq 0.376$ (indeed we have $\gamma_{\eta,q} < 1$ as needed) and $\alpha \geq 0.44$. Using these bounds and the choice of δ, the approximation factor guarantee of Theorem 6.10 is at least $1/10$. ∎

Alon–Naor in the Language of a Rietz Rounding Scheme

We note that the Alon–Naor correlation oracle from [JST21] (recalled in Section 6.1) can be stated as in the above language and solved using Theorem 6.10 above yielding Lemma 6.11.

Lemma 6.11 (Alon–Naor as a Rietz Rounding Scheme)**.** *Let*

$$\alpha = \frac{2}{\pi}, \quad \beta = 1 \quad \textit{and} \quad \gamma = \frac{\pi}{2} - 1\,.$$

The Rietz function

$$r_u(g) := \langle u, g \rangle - \sqrt{(\pi/2)}\,\mathrm{sgn}\,\langle u, g \rangle\,,$$

and the rounding function

$$\mathrm{round}(\langle u, g \rangle) := \mathrm{sgn}(\langle u, g \rangle)\,,$$

form an Rietz rounding scheme with $\Sigma = \{\pm 1\}$ and parameters α, β and γ.

The above lemma readily follows from the follows fact of Alon and Naor [AN04].

Fact 6.12 (Alon–Naor [AN04], *cf.*,Eq. 5)**.** *Let $u, w \in \mathbb{R}^d$ be unit vectors in ℓ_2-norm. Then*

$$\frac{\pi}{2} \cdot \mathbb{E}\left[\mathrm{sgn}\,\langle u, g \rangle\,\mathrm{sgn}\,\langle w, g \rangle\right] = \langle u, w \rangle + \mathbb{E}\left[\left(\langle u, g \rangle - \sqrt{\frac{\pi}{2}}\,\mathrm{sgn}\,\langle u, g \rangle\right)\left(\langle w, g \rangle - \sqrt{\frac{\pi}{2}}\,\mathrm{sgn}\,\langle w, g \rangle\right)\right],$$

where the expectations are taken with respect to a random Gaussian vector $g \sim N(0, \mathsf{I}_d)$.

6.3 Grothendieck Problem over Representations

We will now describe a correlation oracle for representations of a finite group $\mathfrak{G}$. First, we give some definitions. Let $\rho\colon \mathfrak{G} \to \mathbb{U}$ be an s-dimensional representation. Let $W \subseteq [n] \times [n']$ and let $f\colon W \to M_s(\mathbb{C})$. This function f will play a similar role to the matrix A

from Section 6.1. Ideally, we would like to find a polynomial time constant factor approximation to the following problem

$$\max_{b,b' \in \mathfrak{G}^n} \left| \mathop{\mathbb{E}}_{(i_1,i_2)\in W} \left[\mathrm{Tr}(f(w)^\dagger \rho(b_{i_1}) \rho(b'_{i_2})) \right] \right| . \tag{10}$$

This means that ideally we would like to solve this version of the Grothendieck problem for $\mathrm{CUT}_\rho^{\otimes 2}$. Note that in the weak regularity framework if we want to find a decomposition with respect to a class of functions $\mathcal{F}$, we can instead find a decomposition larger class of function $\mathcal{F}'$ s containing $\mathcal{F}$. Instead of finding a decomposition with respect to $\mathrm{CUT}_\rho^{\otimes 2}$, we will find a decomposition with respect to $\mathrm{CUT}_{\mathbb{U}_s}^{\otimes 2}$. This will make the Grothendieck problem simpler since we are relaxing the maximization in Eq. (10) to

$$\sup_{U,V\colon [n]\to \mathbb{U}_s} \left| \mathop{\mathbb{E}}_{(i_1,i_2)\in W} \left[\mathrm{Tr}(f(w)^\dagger U(i_1) V(i_2)) \right] \right| . \tag{11}$$

While the image of ρ is a discrete subgroup of $\mathbb{U}_s$, the unitary group $\mathbb{U}_s$ is continuous. Since the weak regularity framework relies on functions taking values in a finite codomain, we will need to discretize.

We will use the following correlation oracle.

Theorem 6.13 (Naor–Regev–Vidick Correlation Oracle)**.** *Let $\mathcal{F}$ be $\mathrm{CUT}_{\mathbb{U}_{s,\delta,k}}^{\otimes 2}$ and μ be the uniform measure supported on at most m elements of $[n'] \times [n']$. There exists an algorithmic $(\delta, \alpha_{\mathbf{NRV}} \cdot \delta)$-correlation oracle $\mathcal{O}_{\mu,B}$ running in time $\mathcal{T}_{\mathcal{O}_{\mu,B}} = \tilde{O}_{\delta,s,B}(\mathrm{poly}(m+n'))$, where $\alpha_{\mathbf{NRV}} \geq 1/4$ is the approximation ratio constant.*

We can now reduce our problem to [NRV13].

Corollary 6.14 (Corollary from [NRV13])**.** *There is an algorithm running in time $\mathrm{poly}(|W|, s, \delta, k)$ that given $f\colon W \to M_s(\mathbb{C})$ with $\|f\|_\infty = O_{s,k,\delta}(1)$ finds $\widetilde{U}, \widetilde{V}\colon [n] \to \mathbb{U}_{s,k,\delta}$ such that*

$$\left| \mathop{\mathbb{E}}_{(i_1,i_2)\in W} \left[\mathrm{Tr}(f(w)^\dagger \widetilde{U}(i_1) \widetilde{V}(i_2)) \right] \right| \geq \frac{1}{4} \sup_{U,V\colon [n]\to \mathbb{U}_s} \left| \mathop{\mathbb{E}}_{(i_1,i_2)\in W} \left[\mathrm{Tr}(f(w)^\dagger U(i_1) V(i_2)) \right] \right| + \mathcal{E}$$

where $\mathcal{E} = O_{s,k,\delta}(1) \cdot \sum_{(i_1,i_2)\in W} \mathrm{Tr}(f(i_1,i_2)^\dagger f(i_1,i_2))$.

We now recall their setting where a 4-tensor $T \in \mathbb{C}^{[n]^{\otimes 4}}$ is given as input and the object function is

$$\mathsf{OPT} = \sup_{U,V\in \mathbb{U}_n} \left| \sum_{i,j,k,\ell} T_{i,j,\ell,k} U_{i,j} V_{k,\ell} \right| .$$

They give a $\mathrm{poly}(n, \|T\|_\infty)$ bound approximation algorithm that finds unitaries $\widetilde{U}, \widetilde{V} \in \mathbb{U}_n$ such that

$$\sup_{U,V\in \mathbb{U}_n} \left| \sum_{i,j,k,\ell} T_{i,j,\ell,k} \widetilde{U}_{i,j} \widetilde{V}_{k,\ell} \right| \geq \frac{1}{4} \mathsf{OPT} .$$

The precise constant is not important for our application so we just take it to be $1/4$ for simplicity.

Proof. We show how to reduce our problem to the more general 4-tensor setting of [NRV13].

First, define the following 4-tensor

$$T_{(i_1,j_1),(i_1,j_2),(i_2,k_1),(i_2,k_2)} = \begin{cases} f(i_1,i_2)^\dagger_{k_2,j_1} & \text{if } (i_1,i_2) \in W \text{ and } j_2 = k_1 \\ 0 & \text{otherwise} \end{cases},$$

and undefined entries are set to 0. We have that

$$\begin{aligned} &\sum_{i,j,k,\ell} T_{i,j,\ell,k} U_{i,j} V_{k,\ell} \\ &= \sum_{i=(i_1,j_1),j=(i_1,j_2),k=(i_2,k_1),\ell=(i_2,k_2)} \mathbf{1}\left[(i_1,i_2) \in W\right] \mathbf{1}\left[j_2 = k_1\right] V_{(i_2,k_1),(i_2,k_2)} f(i_1,i_2)^\dagger_{k_2,j_1} U_{(i_1,j_1),(i_1,j_2)} \\ &= \sum_{(i_1,i_2)\in W} \mathrm{Tr}\left(V(i_2) f(i_1,i_2)^\dagger U(i_1)\right), \end{aligned}$$

$U(i_1)$ indexes the i_1th main $s \times s$ block of U and similarly $V(i_2)$ indexes the i_2th main $s \times s$ block of V. Note that the maximization over U and V ranges over all unitaries and, in particular, those that have at least n disjoint main diagonal $s \times s$ blocks (off-diagonal blocks are zero). This means that $U(i_1), V(i_2)$ can range over a set of matrices containing $\mathbb{U}_s$ and have operator norm at most 1. We now discretize the space of matrices $M_s(\mathbb{C})$ of operator norm at most 1 a fine enough net $\mathbb{U}_{s,k,\delta}$ depending on s, δ and k, so that when we replaced $U(i_1), V(i_2)$ by their closest elements in the net $\widetilde{U}(i_1), \widetilde{V}(i_2)$ we still have an additive approximation as needed to conclude this corollary. ∎

Remark 6.15. *Bandeira et al. [BKS16] have an approximation algorithm for a similar looking optimization of Eq. (11). However, they only analyze the* little *Grothendieck version of the problem, so we cannot directly use their result here. It is plausible that it can be generalized to Hermitian matrices.*

7 Fast Decoding Prime q-ary Codes near the GV Bound

We show how to decode explicit codes over $\mathbb{F}_q$ for any prime q close to GV bound in the large distance regime,i.e., $1 - 1/q - \varepsilon$ for small $\varepsilon > 0$. We now proceed towards proving our main result (restated below).

Theorem 1.1 (Main I - Near-linear Time Unique Decoding over $\mathbb{F}_q$)**.** *Let q be a prime. For every $\varepsilon > 0$ sufficiently small, there are explicit linear Ta-Shma codes $\mathcal{C}_{N,q,\varepsilon} \subseteq \mathbb{F}_q^N$ for infinitely many values $N \in \mathbb{N}$ with*

(i) distance at least $(1 - 1/q)(1 - \varepsilon)$ (actually ε-balanced),

(ii) rate $\Omega_q(\varepsilon^{2+\alpha})$ where $\alpha = O(1/(\log_2(1/\varepsilon))^{1/6})$, and

(iii) an $r(q/\varepsilon) \cdot \tilde{O}(N)$ time randomized unique decoding algorithm that decodes within radius $((1 - 1/q)(1 - \varepsilon))/2$,

where $r(x) = \exp(\exp(\mathrm{poly}(x)))$.

We first recall some coding theory terminology in Section 7.1. In Section 7.2, we present the near-linear time decoding algorithm modulo a suitable base code. In Section 7.3, we put everything together with a suitable base code to obtain Theorem 1.1.

7.1 Preliminaries on Codes

We briefly recall some standard code terminology. Let q be a prime. Given $z, z' \in \mathbb{F}_q^n$, recall that the relative Hamming distance between z and z' is $\Delta(z, z') := |\{i \mid z_i \neq z'_i\}|/n$. A linear code is any subspace $\mathcal{C} \subseteq \mathbb{F}_q^n$. The distance of $\mathcal{C}$ is defined as $\Delta(\mathcal{C}) := \min_{z \neq z'} \Delta(z, z')$ where $z, z' \in \mathcal{C}$. The rate of $\mathcal{C}$ is $\log_q(|\mathcal{C}|)/n$, or equivalently $\dim(\mathcal{C})/n$ (if $\mathcal{C}$ is linear).

We will need the following standard notion of bias for q-ary alphabet (using its $(\mathbb{Z}_q, +)$ algebraic structure).

Definition 7.1 (Bias)**.** *The* bias *of a word $z \in \mathbb{F}_q^n$ is defined as* $\mathrm{bias}(z) := \max_{a \in \mathbb{F}_q^n \setminus \{0\}} \left| \mathbb{E}_{i \in [n]} \chi_a(z_i) \right|$. *The bias of a code $\mathcal{C}$ is the maximum bias of any non-zero codeword in $\mathcal{C}$.*

Definition 7.2 (ε-balanced Code)**.** *A code $\mathcal{C}$ is* ε-balanced *if* $\mathrm{bias}(z + z') \leq \varepsilon$ *for every pair of distinct $z, z' \in \mathcal{C}$.*

Direct Sum Lifts

Starting from a base code $\mathcal{C} \subseteq \mathbb{F}_q^n$, we amplify its distance by considering the *direct sum lifting* operation based on a collection $W \subseteq [n]^k$. The direct sum lifting maps each codeword of $\mathcal{C}$ to a new word in $\mathbb{F}_q^{|W|}$ by taking sum in $\mathbb{F}_q$ of its entries on each element of W.

Definition 7.3 (Direct Sum Lifting)**.** *Let $W \subseteq [n]^k$. For $z \in \mathbb{F}_q^n$, we define the* direct sum lifting *as* $\mathrm{dsum}_W(z) = y$ *such that $y_{(i_1,\ldots,i_k)} = \sum_{j=1}^k z_{i_j}$ for all $(i_1, \ldots, i_k) \in W$. The direct sum lifting of a code $\mathcal{C} \subseteq \mathbb{F}_q^n$ is*

$$\mathrm{dsum}_W(\mathcal{C}) = \{\mathrm{dsum}_W(z) \mid z \in \mathcal{C}\}.$$

We will omit W from this notation when it is clear from context.

We will use terminology of *bias reducer* (generalizing the notion of *parity samplers* from [TS17] to larger alphabets).

Definition 7.4 (Bias Reducer)**.** *A collection $W \subseteq [n]^k$ is called an* $(\varepsilon_0, \varepsilon)$-bias reducer *if for all $z \in \mathbb{F}_q^n$ with* $\mathrm{bias}(z) \leq \varepsilon_0$*, we have* $\mathrm{bias}(\mathrm{dsum}_W(z)) \leq \varepsilon$.

7.2 Near-linear Time Prime q-ary Decoding

We now develop list-decoding algorithms for direct-sum codes, using the regularity lemmas obtained in the Section 5. We will prove the following theorem which will be the q-ary generalization, with q prime, of the binary decoder in [JST21]. Having access to regular decomposition (currently requiring randomness to compute), the remaining steps of the decoder will now be deterministic. We currently require q to be prime and the reason for this will be clearer as we develop the algorithm. In a few words, when q is prime all the non-trivial characters of $\mathbb{Z}_q$ are of the form $a \mapsto \omega^a$, where ω is a primitive qth root of unity and this will be important to us[27]. It would be interesting to remove this restriction.

[27] It will allow us to control the distribution of symbols inside atoms in a factor (for most of them) so that that the majority occurs with overwhelming probability.

Theorem 7.5. *Let $\mathcal{C}_0 \subset \mathbb{F}_q^n$, with q prime, be a code with $\mathrm{bias}(\mathcal{C}_0) \leq \varepsilon_0$, which is unique-decodable to distance δ_0 in time $\mathcal{T}_0$. Let $W \subseteq [n]^k$ be a d-regular, τ-splittable collection of tuples, and let $\mathcal{C} = \mathrm{dsum}_W(\mathcal{C}_0)$ be the corresponding direct-sum lifting of $\mathcal{C}_0$ with $\mathrm{bias}(\mathcal{C}) \leq \varepsilon$. Let β be such that*

$$\beta \;\geq\; \max\left\{ \left(2^{20} \cdot \tau \cdot k^3\right)^{1/2},\, 4(1-(C_q \cdot \delta_0/2)^2)^{k/2} \right\},$$

where $C_q := 1 - \cos(\pi/q)$. Then, there exists a randomized algorithm, which given $\tilde{y} \in \mathbb{F}_q^W$, with high probability recovers the list $\mathcal{L}_\beta(\tilde{y}) := \{y \in \mathcal{C} \mid \Delta(\tilde{y}, y) \leq (1 - 1/q)(1-\beta)\}$, in time $\tilde{O}(C_{\beta,k,\varepsilon_0,q} \cdot (|W| + \mathcal{T}_0))$, where $C_{k,\beta,\varepsilon_0,q} = 2^{q^{O(k^3/\beta^2)}}$.

To obtain the decoding algorithm, we first define a function $g_a : [n]^k \to \mathbb{S}_q^1$, for each $a \in \mathbb{F}_q^\times$, supported on W as

$$g_a(i_1, \ldots, i_k) \;:=\; \begin{cases} \omega^{a \cdot \tilde{y}_{(i_1,\ldots,i_k)}} & \text{if } (i_1,\ldots,i_k) \in W \\ 0 & \text{otherwise} \end{cases}$$

For each $z \in \mathbb{F}_q^n$ and $a \in \mathbb{F}_q$, recall that the function $\chi_{z,a} : [n] \to \{\omega^{a \cdot b} | b \in \mathbb{F}_q\}$ is defined as $\chi_{z,a}(i) = \omega^{a \cdot z_i}$. We now relate distance with average bias as follows.

Claim 7.6. *Let $z \in \mathbb{F}_q^n$, and let the functions g_a and $\chi_{z,a}$ be as above. Then,*

$$\Delta(\tilde{y}, \mathrm{dsum}_W(z)) \leq \left(1 - \frac{1}{q}\right)(1-\beta) \quad \Leftrightarrow$$
$$\mathop{\mathbb{E}}_{a \in \mathbb{F}_q^\times}\left[\left\langle g_a, \chi_{z,a}^{\otimes k}\right\rangle_{\mu_k}\right] = \left(\frac{n}{d}\right)^{k-1} \cdot \mathop{\mathbb{E}}_{a \in \mathbb{F}_q^\times}\left[\left\langle g_a, \chi_{z,a}^{\otimes k}\right\rangle_{\mu_1^{\otimes k}}\right] \geq \beta\,.$$

Proof. We have

$$\begin{aligned}
\Delta(\tilde{y}, \mathrm{dsum}_W(z)) \;&=\; \mathop{\mathbb{E}}_{(i_1,\ldots,i_k)\sim W}\left[\mathbf{1}_{\left[\tilde{y}_{(i_1,\ldots,i_k)} \neq z_{i_1}+\cdots+z_{i_k}\right]}\right] \\
&=\; \mathop{\mathbb{E}}_{(i_1,\ldots,i_k)\sim \mu_k}\left[1 - \mathop{\mathbb{E}}_{a \in \mathbb{F}_q}\left[g_a(i_1,\ldots,i_k) \cdot \prod_{t \in [k]} \chi_{z,a}(i_t)\right]\right] \\
&=\; 1 - \frac{1}{q} - \frac{q-1}{q} \cdot \mathop{\mathbb{E}}_{a \in \mathbb{F}_q^\times}\left[\left\langle g_a, \chi_{z,a}^{\otimes k}\right\rangle_{\mu_k}\right] \\
&=\; \left(1 - \frac{1}{q}\right)\left(1 - \mathop{\mathbb{E}}_{a \in \mathbb{F}_q^\times}\left[\left\langle g_a, \chi_{z,a}^{\otimes k}\right\rangle_{\mu_k}\right]\right).
\end{aligned}$$

Finally, using the fact that g_a is only supported on W, and $|W| = d^{k-1} \cdot n$ by d-regularity, we have $\langle g_a, f\rangle_{\mu_k} = (n/d)^{k-1} \cdot \langle g_a, f\rangle_{\mu_1^{\otimes k}}$ for any function $f : [n]^k \to \mathbb{K}$. ∎

Note that each element of the list $\mathcal{L}_\beta(\tilde{y})$ must be equal to $\mathrm{dsum}_W(z)$ for some $z \in \mathcal{C}_0$. Thus, to search for all such z, we will consider the decomposition h_a of the function g_a, given by Theorem 5.12 with respect to the class of functions $\mathcal{F} = \mathrm{CUT}_{\omega,q,a}^{\otimes k}$. Since the functions $\chi_{z,a}^{\otimes k}$ belong to $\mathcal{F}$, it will suffice to only consider the inner product $\langle h_a, \chi_{z,a}^{\otimes k}\rangle_{\mu_1^{\otimes k}}$.

Also, since the approximating function h_a is determined by a small number of functions, say $\left\{f_{a,1}, \ldots, f_{a,r} : [n] \to \mathbb{S}_q^1\right\}$, it will suffice to (essentially) consider only the functions measurable in the factor $\mathcal{B}_a$ determined by $f_{a,1}, \ldots, f_{a,r}$. Recall that the factor $\mathcal{B}_a$ is simply a partition of $[n]$ in q^r pieces according to the values of $f_{a,1}, \ldots, f_{a,r}$. Also, since any $\mathcal{B}_a$-measurable function is constant on each piece, it is completely specified by $|\mathcal{B}_a|$ values in $\mathbb{K}$. We will only consider $\mathcal{B}_a$-measurable functions taking values in $\mathbb{F}_q$. The decoding procedure is described in the following algorithm.

Algorithm 7.7 (List Decoding Algorithm).
Input $\tilde{y} \in \mathbb{F}_q^W$
Output *List* $\mathcal{L} \subseteq \mathcal{C}$

- *Obtain the approximator h_a given by Theorem 5.12 for $\mathcal{F} = \mathrm{CUT}_{\omega,q,a}^{\otimes k}$, $\delta = \beta$, and the function $g_a : [n]^k \to \mathbb{S}_q^1$ defined as*

$$g_a(i_1, \ldots, i_k) := \begin{cases} \omega^{a \cdot \tilde{y}_{(i_1,\ldots,i_k)}} & \text{if } (i_1, \ldots, i_k) \in W \\ 0 & \text{otherwise} \end{cases}$$

- *Let h_a be of the form $h_a = \sum_{j=1}^{p} c_{a,j} \cdot f_{a,j_1} \otimes \cdots \otimes f_{a,j_k}$, with each $f_{a,j_t} : [n] \to \mathbb{S}_q^1$. Let $\mathcal{B}_a$ be the factor determined by the functions $\left\{f_{a,j_t}\right\}_{j \in [p], t \in [k]}$.*

- *Let $\mathcal{L} = \emptyset$. For each $a \in \mathbb{F}_q^\times$ and $\mathcal{B}_a$-measurable function $\widetilde{z}$ given by a value in $\mathbb{F}_q$ for every atom of $\mathcal{B}_a$:*

  – *If there exists $z \in \mathcal{C}_0$ such that*

$$\Delta(\tilde{z}, z) \leq \delta_0 \quad \text{and} \quad \Delta(\tilde{y}, \mathrm{dsum}_W(z)) \leq \left(1 - \frac{1}{q}\right)(1 - \beta),$$

  then $\mathcal{L} \leftarrow \mathcal{L} \cup \{\mathrm{dsum}_W(z)\}$.

- *Return $\mathcal{L}$.*

Note that by our choice of the β in Theorem 7.5, we have that $\tau \leq \beta^2/(2^{20}k^3)$. Thus, we can indeed apply Theorem 5.12 to obtain the functions h_a as required by the algorithm. To show that the algorithm can recover the list, we will show that for each z such that $\mathrm{dsum}_W(z) \in \mathcal{L}_\beta$, the distribution of symbols of z on each of the parts of the factor $\mathcal{B}_a$ is approximately a delta distribution on a single element in $\mathbb{F}_q$.

Next we show that when $z \in \mathbb{F}_q^n$ is such that $\mathbb{E}_{a \in \mathbb{F}_q^\times}[\langle g_a, \chi_{z,a}^{\otimes k} \rangle]$ is large, then the norm of the conditional expectation $\mathbb{E}[\chi_{z,a}|\mathcal{B}_a]$ is also large for some $a \in \mathbb{F}_q^\times$. The procedure will find a $\tilde{z}$ close to z now in a deterministic way. When we have a $z \in \mathcal{C}_0$ with such a property, we can use $\tilde{z}$ to recover z using the unique decoding algorithm for $\mathcal{C}_0$.

We show that average bias in the "lifted" space becomes a much stronger bias in the "base" space with respect to some $a \in \mathbb{F}_q^\times$. More precisely, we have the following.

Lemma 7.8. *Let $z \in \mathbb{F}_q^n$ be such that*

$$\mathop{\mathbb{E}}_{a \in \mathbb{F}_q^\times}\left[\left\langle g_a, \chi_{z,a}^{\otimes k} \right\rangle_{\mu_k}\right] = \left(\frac{n}{d}\right)^{k-1} \cdot \mathop{\mathbb{E}}_{a \in \mathbb{F}_q^\times}\left[\left\langle g_a, \chi_{z,a}^{\otimes k} \right\rangle_{\mu_1^{\otimes k}}\right] \geq \beta.$$

Then, we have $\max_{a\in\mathbb{F}_q^\times} \|\mathbb{E}[\chi_{z,a}|\mathcal{B}_a]\|_{\mu_1}^2 \geq (\beta/4)^{2/k}$.

Proof. Let h_a be the approximating function obtained by applying Theorem 5.12 to g_a with approximation error $\delta = \beta$. Note that we have $\|h_a\|_{\mu_1^{\otimes k}} \leq 2$, and for any $f \in \mathrm{CUT}_{\omega,q,a}^{\otimes k}$,

$$\left| \left(\frac{n}{d}\right)^{k-1} \cdot \left\langle g_a - \left(\frac{d}{n}\right)^{k-1} \cdot h_a \, , f \right\rangle_{\mu_1^{\otimes k}} \right| \;\leq\; \delta \,.$$

Using $f = \chi_{z,a}^{\otimes k}$ and $\delta = \beta/2$, we get

$$\left| \mathop{\mathbb{E}}_{a\in\mathbb{F}_q^\times} \left[\left\langle h_a, \chi_{z,a}^{\otimes k} \right\rangle_{\mu_1^{\otimes k}} \right] \right| \;\geq\; \beta - \delta \;\geq\; \frac{\beta}{2} \,.$$

Using Proposition 4.8, and the fact that $\mathcal{B}_a$ is defined so that all functions in the decomposition of h_a are (by definition) $\mathcal{B}_a$-measurable, we have

$$\begin{aligned} \left\langle h_a, \chi_{z,a}^{\otimes k} \right\rangle_{\mu_1^{\otimes k}} \;&=\; \sum_{j=1}^{p} c_j \prod_{t=1}^{k} \left\langle f_{a,j_t}, \chi_{z,a} \right\rangle_{\mu_1} \;=\; \sum_{j=1}^{p} c_j \prod_{t=1}^{k} \left\langle f_{a,j_t}, \mathbb{E}\left[\chi_{z,a}|\mathcal{B}_a\right] \right\rangle_{\mu_1} \\ &=\; \left\langle h_a, \left(\mathbb{E}\left[\chi_{z,a}|\mathcal{B}_a\right]\right)^{\otimes k} \right\rangle_{\mu_1^{\otimes k}} . \end{aligned}$$

Combining the above with Cauchy-Schwarz, we get

$$\begin{aligned} \frac{\beta}{2} \;\leq\; \left| \mathop{\mathbb{E}}_{a\in\mathbb{F}_q^\times} \left[\left\langle h_a, \chi_{z,a}^{\otimes k} \right\rangle_{\mu_1^{\otimes k}} \right] \right| \;&\leq\; \mathop{\mathbb{E}}_{a\in\mathbb{F}_q^\times} \left[\|h_a\|_{\mu_1^{\otimes k}} \cdot \left\| \left(\mathbb{E}\left[\chi_{z,a}|\mathcal{B}_a\right]\right)^{\otimes k} \right\|_{\mu_1^{\otimes k}} \right] \\ &=\; \mathop{\mathbb{E}}_{a\in\mathbb{F}_q^\times} \left[\|h_a\|_{\mu_1^{\otimes k}} \cdot \|\mathbb{E}\left[\chi_{z,a}|\mathcal{B}_a\right]\|_{\mu_1}^k \right] . \end{aligned}$$

Using $\|h_a\|_{\mu_1^{\otimes k}} \leq 2$ then gives $\max_{a\in\mathbb{F}_q^\times} \|\mathbb{E}[\chi_{z,a}|\mathcal{B}_a]\|_{\mu_1}^2 \geq (\beta/4)^{2/k}$. ∎

Contrary to the binary case, having strong bias with a single $a \in \mathbb{F}_q^\times$ is not necessarily enough to have a majority symbol occurring with sufficiently high frequency in each atom of $\mathcal{B}_a$ in general. By assuming that q is prime, this will be possible as we now show by establishing some claims.

A distribution with large bias with respect to a primitive root (which is the case for non-trivial character of $\mathbb{Z}_q$ with q prime) also has a majority occurring with high probability.

Claim 7.9. *Let $(p_a)_{a\in\mathbb{F}_q}$ be a probability distribution on $\mathbb{F}_q$ with q prime. Let ω be a primitive qth root of unity. Then,*

$$\left| \sum_{a\in\mathbb{F}_q} \omega^a p_a \right|^2 \;\leq\; 1 - C_q\left(1 - \sum_{a\in\mathbb{F}_q} p_a^2\right),$$

where $C_q := 1 - \cos(\pi/q)$.

Proof. Let $v_a = (\cos(2\pi a/q), \sin(2\pi a/q))$. Observe that for $a \neq a'$, we have

$$\begin{aligned} |\langle v_a, v_{a'}\rangle| &= \left|\cos(2\pi a/q)\cos(2\pi a'/q) + \sin(2\pi a/q)\cos(2\pi a'/q)\right| \\ &= \left|\cos(2\pi(a-a')/q)\right| \leq \max_{a'' \in \mathbb{F}_q^\times} \left|\cos(2\pi a''/q)\right| \;\leq\; |\cos(\pi/q)|\,. \end{aligned}$$

Computing, we obtain

$$\left|\sum_{a\in\mathbb{F}_q} v_a p_a\right|^2 \;\leq\; \sum_{a,a'\in\mathbb{F}_q} p_a p_{a'} \,|\langle v_a, v_{a'}\rangle| \;\leq\; \left(\sum_{a\in\mathbb{F}_q} p_a\right)^2 - C_q \sum_{a\neq a'} p_a p_{a'} \;\leq\; 1 - C_q + C_q \sum_a p_a^2\,. \qquad \blacksquare$$

The quantity $\sum_{a\in\mathbb{F}_q} p_a^2$ (collision probability) if large, readily implies that the distribution is close to a delta distribution.

Claim 7.10. *Let $(p_a)_{a\in[q]}$ be a probability distribution. If $\sum_{a=1}^q p_a^2 \geq 1-\nu$, then there exists $a \in [q]$ with $p_a \geq 1-\nu$.*

Proof. Let $p^* = \max_{a\in[q]} p_a$. We have $1-\nu \leq \sum_{a=1}^q p_a^2 \leq p^* \sum_{a=1}^q p_a = p^*$. $\blacksquare$

Combining the two preceding observations, we deduce the following.

Corollary 7.11. *Let $(p_a)_{a\in\mathbb{F}_q}$ be a probability distribution on $\mathbb{F}_q$. Let ω be a primitive qth root of the unit. If $\left|\sum_{a\in\mathbb{F}_q} \omega^a p_a\right|^2 \geq 1 - C_q\cdot\nu$ for some $\nu \in (0,1)$, then $p_a \geq 1-\nu$.*

Proof. Combining our assumption with the bound of Claim 7.9 we obtain

$$1 - C_q\cdot\nu \;\leq\; \left|\sum_{a\in\mathbb{F}_q} \omega^a p_a\right|^2 \;\leq\; 1 - C_q(1 - \sum_{a\in\mathbb{F}_q} p_a^2),$$

from which we deduce that $\sum_{a\in\mathbb{F}_q} p_a^2 \geq 1-\nu$. Using Claim 7.10, we conclude the proof. $\blacksquare$

From the large bias assumption in the base space, we deduce closeness in Hamming distance as follows.

Lemma 7.12. *If $\|\mathbb{E}[\chi_{z,a}|\mathcal{B}_a]\|_{\mu_1}^2 \geq 1 - (C_q\cdot\nu)^2$ for some $\nu \in (0,1)$, then there exists a $\mathcal{B}_a$-measurable function $\widetilde{z} \in \mathbb{F}_q^n$ such that $\Delta(\widetilde{z}, z) \leq (1+C_q)\cdot\nu$.*

Proof. Let $i_1, \ldots, i_r \in [n]$ be a set of representatives of the factor $\mathcal{B}_a$. Each part $\mathcal{B}_a(i)$ has measure $\mathbb{P}_{\mu_1}[\mathcal{B}_a(i)] = |\mathcal{B}_a(i)|/n$ and the multiset of elements $\{z_j \mid j \in \mathcal{B}_a(i)\}$ gives rise to a probability distribution on $\mathbb{F}_q$. By assumption, we have

$$1 - \left(C_q\cdot\nu\right)^2 \;\leq\; \|\mathbb{E}\left[\chi_{z,a}|\mathcal{B}_a\right]\|_{\mu_1}^2 \;=\; \sum_{i\in\{i_1,\ldots,i_r\}} \left|\mathbb{E}_{j\in\mathcal{B}_a(i)}\omega^{a\cdot z_j}\right|^2 \cdot \mathop{\mathbb{P}}_{\mu_1}[\mathcal{B}_a(i)].$$

Since $\left|\mathbb{E}_{j\in\mathcal{B}_a(i)}\omega^{a\cdot z_j}\right|^2 \leq 1$, the above guarantee implies that with probability at least $1 - C_q\cdot\nu$ over the choice of factor $\mathcal{B}_a(i)$ we have $\left|\mathbb{E}_{j\in\mathcal{B}_a(i)}\omega^{a\cdot z_j}\right|^2 \geq 1 - C_q\cdot\nu$. Using Corollary 7.11, we deduce that the distribution on $\mathbb{F}_q$ induced by each such $\mathcal{B}_a(i)$ has an element which we denote by $\widetilde{z}_i$ occurring with probability at least $1-\nu$. The values $\widetilde{z}_i$ can be seeing as defining a partial function from $[n]$ to $\mathbb{F}_q$. By extending it to a $\mathcal{B}_a$-measurable function $\widetilde{z}$ in an arbitrary way, we conclude the claim. $\blacksquare$

Using the above results, we can now complete the analysis of the algorithm.

Proof of Theorem 7.5. Using Lemma 7.8, there exists $a \in \mathbb{F}_q^\times$ such that $\|\mathbb{E}[\chi_{z,a}|\mathcal{B}_a]\|_{\mu_1}^2 \geq (\beta/4)^{2/k}$. By our choice of β, we have $\beta \geq 4(1-(C_q \cdot \delta_0/2)^2)^{k/2}$ implying that

$$\|\mathbb{E}\left[\chi_{z,a}|\mathcal{B}_a\right]\|_{\mu_1}^2 \;\geq\; 1-(C_q \cdot \delta_0/2)^2\,.$$

By Lemma 7.12, for each codeword $z \in \mathcal{C}_0$ such that $\mathrm{dsum}_W(z) \in \mathcal{L}_\beta$, there exists a $\mathcal{B}_a$-measurable word $\widetilde{z}$ such that $\Delta(z,\widetilde{z}) \leq \delta_0$. When the algorithm invokes the decoder of $\mathcal{C}_0$ on $\widetilde{z}$ (each $\mathcal{B}_a$-measurable function on $\mathbb{F}_q$ is considered), by assumption we are guaranteed to retrieve z. Therefore, we can indeed recover each codeword $\mathrm{dsum}_W(z) \in \mathcal{L}_\beta$.

Running time. Using Theorem 5.12, the regularity decomposition of each h_a can be computed in time $\tilde{O}(C_{\beta,k,\varepsilon_0,q} \cdot |W|)$. Given the functions $f_{a,1},\ldots,f_{a,r}$ forming the decomposition of h_a, the factor $\mathcal{B}_a$ can be computed in time $O(q^{kr} \cdot n)$. Also, the distance $\Delta(\tilde{y}, \mathrm{dsum}_W(z))$ can be computed in time $O(|W|)$. Since the total number of decoding steps is at most $q^{|B_a|}$ and the number of functions in the decomposition of h_a is $O(k^3/\beta^2)$ from Theorem 5.12, we get that the number of decoding steps is $q^{q^{O(k^3/\beta^2)}}$ for the factor $\mathcal{B}_a$. Thus, the total running time is bounded by $\tilde{O}(C_{\beta,k,\varepsilon_0,q} \cdot (|W| + \mathcal{T}_0))$, where $C_{\beta,k,\varepsilon_0,q} = 2^{q^{O(k^3/\beta^2)}}$. ∎

7.3 Instantiating the Decoder with a Base Code

We now combine all the pieces in order to obtain our main result establishing a near-linear time *unique* decoding algorithm for Ta-Shma's codes [TS17] over any constant sized prime fields $\mathbb{F}_q$. It will follow from the new generalized regularity based list decoding algorithm for direct sum codes over prime $\mathbb{F}_q$, Theorem 7.5, applied to the decoding of a slight modification of Ta-Shma's construction from [JQST20] that yields a splittable collection of tuples[28].

Theorem 1.1 (Main I - Near-linear Time Unique Decoding over $\mathbb{F}_q$)**.** *Let q be a prime. For every $\varepsilon > 0$ sufficiently small, there are explicit linear Ta-Shma codes $\mathcal{C}_{N,q,\varepsilon} \subseteq \mathbb{F}_q^N$ for infinitely many values $N \in \mathbb{N}$ with*

(i) distance at least $(1-1/q)(1-\varepsilon)$ (actually ε-balanced),

(ii) rate $\Omega_q(\varepsilon^{2+\alpha})$ where $\alpha = O(1/(\log_2(1/\varepsilon))^{1/6})$, and

(iii) an $r(q/\varepsilon) \cdot \tilde{O}(N)$ time randomized unique decoding algorithm that decodes within radius $((1-1/q)(1-\varepsilon))/2$,

where $r(x) = \exp(\exp(\mathrm{poly}(x)))$.

We now state the properties and guarantees needed in our work of this slightly modified version of Ta-Shma's direct sum construction of near optimal ε-balanced codes. To make the decoding task more transparent, we will additionally require the base code in Ta-Shma's construction to have the following technical property.

[28] Recently, Blanc and Doron [BD22] use a weaker expansion condition. In fact, our framework only relies on the splittable mixing lemma.

Definition 7.13. *We say that a code has symbol multiplicity $m \in \mathbb{N}$ if it can be obtained from another code by repeating each symbol of its codeword m times.*

The parameter trade-offs in Ta-Shma's construction [TS17] over $\mathbb{F}_2$ are (essentially) the same as those over $\mathbb{F}_q$ as analyzed by Jalan and Moshkovitz in [JM21]. In particular, the decay in bias is the same. For this reason, we can almost reuse the following theorem from [TS17] which provides an interface to Ta-Shma's parameters except that we will need to be more careful with the rate of the base code which now depends on the alphabet size q.

Theorem 7.14 (Ta-Shma's Codes (implicit in [TS17] following q-ary analysis of [JM21]))**.** *Let $c > 0$ be any constant. For every $\varepsilon > 0$ sufficiently small, there exists $k = k(\varepsilon)$ satisfying $\Omega(\log(1/\varepsilon)^{1/3}) \leq k \leq O(\log(1/\varepsilon))$, $\varepsilon_0 = \varepsilon_0(\varepsilon) > 0$, and positive integer $m = m(\varepsilon) \leq (1/\varepsilon)^{o(1)}$ such that Ta-Shma's construction yields a collection of τ-splittable tuples $W = W \subseteq [n]^k$ satisfying:*

(i) For every linear ε_0-balanced code $\mathcal{C}_0 \subseteq \mathbb{F}_q^n$ with symbol multiplicity m, the direct sum code $\mathrm{dsum}_W(\mathcal{C}_0)$ is:

(i.1) ε-balanced (parity sampling).

(i.2) if $\mathcal{C}_0$ has rate $\Omega(\varepsilon_0^c/m)$, then $\mathrm{dsum}_W(\mathcal{C}_0)$ has rate $\Omega(\varepsilon^{2+o_c(1)})$ (near optimal rate)

(ii) $\tau \leq \exp(-\Theta(\log(1/\varepsilon)^{1/6}))$ (splittability).

(iii) W is constructible in $\mathrm{poly}(|W|)$ time (explicit construction).

Theorem 7.15. *There exists a constant $\varepsilon_0 > 0$ such that for every $\varepsilon > 0$ sufficiently small and constant size prime field $\mathbb{F}_q$, there is an explicit family of codes over this field such that every member $\mathcal{C}_0 \subseteq \mathbb{F}_q^n$ in the family has $\mathrm{bias}(\mathcal{C}_0) \leq \varepsilon$, rate $\Omega_q(\varepsilon^{O(1)})$ and is unique decodable in time $\widetilde{O}(\exp(\exp(\mathrm{poly}(q/\varepsilon))) \cdot n)$.*

We will prove the (gentle) *list* decoding result of Ta-Shma's codes over prime $\mathbb{F}_q$ of which our main result Theorem 1.1 is a particular case.

Theorem 1.2 (Near-linear time List Decoding over $\mathbb{F}_q$)**.** *Let q be a prime. For every $\varepsilon > 0$ sufficiently small, there are explicit binary linear Ta-Shma codes $\mathcal{C}_{N,q,\varepsilon} \subseteq \mathbb{F}_q^N$ for infinitely many values $N \in \mathbb{N}$ with*

(i) distance at least $(1 - 1/q)(1 - \varepsilon)$ (actually ε-balanced),

(ii) rate $\Omega_q(\varepsilon^{2+\alpha})$ where $\alpha = O(1/(\log_2(1/\varepsilon))^{1/6})$, and

(iii) an $r(q/\varepsilon) \cdot \tilde{O}(N)$ time randomized list decoding algorithm that decodes within radius $1 - 1/q - 2^{-\Theta_q((\log_2(1/\varepsilon))^{1/6})}$ and works with high probability,

where $r(x) = \exp(\exp(\mathrm{poly}(x)))$.

Proof. We start by dealing with a simple technical issue of making the base code in Ta-Shma's construction have the required symbol multiplicity. Let $\mathcal{C}_0' \subseteq \mathbb{F}_q^{n'}$ be an ε_0-balanced

code from Theorem 7.15 which we will use to obtain a base code in Ta-Shma's construction where $\varepsilon_0 > 0$ is a suitable value prescribed by this construction.

Ta-Shma's construction then takes $\mathcal{C}_0' \subseteq \mathbb{F}_q^{n'}$ and forms a new code $\mathcal{C}_0 \subseteq \mathbb{F}_q^n$ by repeating each codeword symbol $m \leq (1/\varepsilon)^{o(1)}$ times. By Claim A.1, $\mathcal{C}_0$ is an ε_0-balanced code that can be unique decoded within the same (fractional) radius of $\mathcal{C}_0'$ in time $\mathcal{T}_0(n) = r \cdot \mathcal{T}_0(n') + \widetilde{O}(r^2 \cdot n')$, where $\mathcal{T}_0(n')$ is the running time of an unique decoder for $\mathcal{C}_0'$. Since by Theorem 7.15 $\mathcal{T}_0(n') = O(\exp(\exp(\mathrm{poly}(q/\varepsilon_0))) \cdot n')$, the decoding time of $\mathcal{C}_0$ can be (crudely) bounded as $\mathcal{T}_0(n) = O(\exp(\exp(\mathrm{poly}(q/\varepsilon))) \cdot n)$.

Let $W = W$ be a collection of tuples from Ta-Shma's construction Theorem 7.14 so that $\mathcal{C} = \mathrm{dsum}_W(\mathcal{C}_0)$ is ε-balanced, $\tau \leq \exp(-\Theta_q(\log(1/\varepsilon)^{1/6}))$ and $k = \Omega_q(\log(1/\varepsilon)^{1/3})$. We will invoke our list decoding algorithm Theorem 7.5 whose list decoding radius $1 - 1/q - \beta$ has to satisfy

$$\beta \geq \max\left\{ \left(2^{20} \cdot \tau \cdot k^3\right)^{1/2}, 4(1 - (C_q \cdot \varepsilon_0/2)^2)^{k/2} \right\}.$$

Using our values of τ and k together with the fact that $(C_q \cdot \varepsilon_0)^2 > 0$ is bounded away from 0 by a constant amount (depending on q) gives

$$\beta \geq \max\left\{ \exp(-\Theta((\log(1/\varepsilon))^{1/6})), \exp(-\Theta_q((\log(1/\varepsilon))^{1/3})) \right\}.$$

Hence, we can take $\beta = \exp(-\Theta_q(\log(1/\varepsilon)^{1/6}))$. Now, we compute the list decoding running time proving a (crude) upper bound on its dependence on ε and q. By Theorem 7.5, the list decoding time

$$\widetilde{O}(C_{\beta,k,\varepsilon_0,q} \cdot (|W| + \mathcal{T}_0(n))),$$

where $C_{k,\beta,\varepsilon_0,q} = 2^{q^{O(k^3/\beta^2)}}$. For our choices of parameters, this decoding time can be (crudely) bounded by $\widetilde{O}(\exp(\exp(\mathrm{poly}(q/\varepsilon))) \cdot N)$. ∎

Choosing the Base Code

We will now describe the family of base codes used in the amplification. To obtain a code $\mathcal{C}_0$ in this family, we will take a code $\mathcal{C}_0'$ from another expander based family (with constant rate, distance and near-linear time decoding) amplify its bias using expander walks to obtain $\mathcal{C}_0$ which will be near-linear time decodable with our decoder Theorem 7.5. In [JST21], they found a suitable "off-the-shelf" family of base codes [GI05].

More precisely, the bias amplified family of base codes will be the following.

Theorem 7.15. *There exists a constant $\varepsilon_0 > 0$ such that for every $\varepsilon > 0$ sufficiently small and constant size prime field $\mathbb{F}_q$, there is an explicit family of codes over this field such that every member $\mathcal{C}_0 \subseteq \mathbb{F}_q^n$ in the family has* $\mathrm{bias}(\mathcal{C}_0) \leq \varepsilon$, *rate $\Omega_q(\varepsilon^{O(1)})$ and is unique decodable in time $\widetilde{O}(\exp(\exp(\mathrm{poly}(q/\varepsilon))) \cdot n)$.*

For the base code $\mathcal{C}_0'$ over $\mathbb{F}_q$, we will use Zemor's Tanner code construction [Zem01] whose rate, distance and decoding analysis is independent of the field size except for the cost of field operations in decoding (see Rao's notes [Rao19]). The local codes in the Tanner construction are of constant size and can be found by brute-force search and the family of expander graphs can be taken to be bipartite Ramanujan graphs from [LPS88]. In summary, we have the following corollary from their results.

Theorem 7.16 (Corollary of [Zem01] and [LPS88]). *There are universal constants $r_0 > 0$ and $\delta_0 > 0$ such that for any finite field $\mathbb{F}_q$, there is an explicit family of codes over this field and each member $\mathcal{C}'_0 \subseteq \mathbb{F}_q^n$ in the family has rate r_0 and can be uniquely decoded from ε_0 fraction of errors in $\widetilde{O}(\mathrm{poly}(q) \cdot n)$ time.*

Proof. **Starting Code:** Let $\mathbb{F}_q$ be a field with q prime. Use Theorem 7.16 to obtain a good family of codes. Each member in $\mathcal{C}'_0 \subseteq \mathbb{F}_q^{n'}$ has constant relative distance at least $2\varepsilon_0$. We will slightly modify this code to ensure that its bias is a constant bounded away from 1. By zeroing out the last $\varepsilon_0/2$ symbols of each codeword in $\mathcal{C}'_0$, we obtain a new linear code $\mathcal{C}''_0$ such that each non-zero codeword has at least $3\varepsilon_0/2$ non-zero symbols and ε_0 zero symbols. Since q is prime, using Claim 7.9, this implies that $\mathrm{bias}(\mathcal{C}''_0) \leq 1 - \eta$, where $\eta = \eta(q, \varepsilon_0) > 0$. Note that the code $\mathcal{C}''_0$ is still unique decodable from $\varepsilon_0/2$ fraction of errors.

Direct-sum Amplification: We will use the simpler expander walk construction of Rozenman and Wigderson (as analyzed by Ta-Shma [TS17] and we use the q-ary version from [JM21]) to amplify this bias to ε using $k = O_\theta(\log_2(1/\varepsilon))$. Let W be the collection of walks and $\mathcal{C}_0 = \mathrm{dsum}_W(\mathcal{C}''_0)$. With this choice of k, we have that dsum_W is $(1-\theta, \varepsilon)$-bias reducer. By choosing an expander with constant but sufficiently small spectral expansion, we obtain W having arbitrarily small splittbility parameter τ as show in [AJQ+20]. By assuming that ε is sufficiently small we can make k arbitrarily large. This implies that we can choose the decoding parameter β, in Theorem 7.5, a fixed constant as small as we want. Hence, we can list decode from radius $(1 - 1/q)(1 - \beta)$.

Rate: Since we can take the expander in this amplification to be of constant degree d_0, we obtain a rate of $r_0 \cdot d_0^{-k} = \Omega(\varepsilon^{O_q(1)})$.

Running Time: We now compute the running time. By Theorem 7.5, the list decoding time is

$$\tilde{O}(C_{\beta,k,\varepsilon_0/2,q} \cdot (|W| + \mathcal{T}_0(n))),$$

where $C_{k,\beta,\varepsilon_0,q} = 2^{q^{O(k^3/\beta^2)}}$. For our choices of parameters, this decoding time can be (crudely) bounded by $\tilde{O}(\exp(\exp(\mathrm{poly}(q/\varepsilon))) \cdot n)$.

To construct a code of rate $\Omega_q(\varepsilon^{O(1)})$ (without the dependence of q in the exponent), we can apply the above construction twice as follows. First, to construct a code of constant bias ε' independent of q and rate $\Omega_q(1)$. From this code, we apply the construction again to obtain a final code with bias ε but this time using $k = O(\log(1/\varepsilon))$ independent of q. This final code has rate $\Omega_q(\varepsilon^{O(1)})$ as desired. ∎

8 Tuple versus Set Constraints

We now show how to obtain a splittable collection of tuples from the hyperedges of a spectral high-dimensional expanders (HDXs) [DK17]. This shows that splittable collections of sets can be seen as a particular case of splittable collections of tuples. This allows us to handle CSPs supported on the edges of high-dimensional expanders as CSPs supported on tuples.

Lemma 8.1. *Let $X = X(\leq d)$ be a τ-splittable HDX with uniform measure on each $X(i)$ for $i \in [d]$.Then, each $\vec{X}(i) := \{(i_1, \ldots, i_j) \mid \{i_1, \ldots, i_j\} \in X(i)\}$ is τ-splittable.*

Proof. We show that if the swap walk $S_{k,k}$ is τ-splittable, then we can find an ordered collection of tuples that is also splittable. We define the corresponding swap walk $\vec{S}_{k,k}$ from $\vec{X}(k)$ to $\vec{X}(k)$ as the normalized (to have largest eigenvalue 1) version of the following operator

$$\left(\vec{S}_{k,k}\right)_{(i_1,\ldots,i_k),(i_{k+1},\ldots,i_{2k})} \propto \mathbf{1}\left[(i_1,\ldots,i_{2k}) \in \vec{X}(2k)\right] .$$

We claim $\vec{S}_{k,k}$ is also τ-splittable. Note that we can (simultaneously) reorder the rows and columns of $\vec{S}_{k,k}$ such that all orientations of a set appear contiguously using a permutation matrix Π in which case we have the following block form

$$\Pi \vec{S}_{k,k} \Pi^t = S_{k,k} \otimes J_{k!\times k!}/k! ,$$

implying that the second largest singular value of $\vec{S}_{k,k}$ is the same as the one of $S_{k,k}$. ∎

Remark 8.2. *Note that the higher-order treshold rank (cf., [AJT19]) is preserved in this translation from sets to tuples.*

Acknowledgement

We thank Vedat Alev and Shravas Rao for stimulating discussions during the initial phase of this project. We thank Shashank Srivastava and Madhur Tulsiani for stimulating discussions leading to [JST21].

References

[AJQ+20] Vedat Levi Alev, Fernando Granha Jeronimo, Dylan Quintana, Shashank Srivastava, and Madhur Tulsiani. List decoding of direct sum codes. In *Proceedings of the 31st ACM-SIAM Symposium on Discrete Algorithms*, pages 1412–1425. SIAM, 2020. 2, 7, 17, 49

[AJT19] Vedat Levi Alev, Fernando Granha Jeronimo, and Madhur Tulsiani. Approximating constraint satisfaction problems on high-dimensional expanders. In *Proceedings of the 60th IEEE Symposium on Foundations of Computer Science*, pages 180–201, 2019. 2, 3, 4, 9, 50

[AK07] Sanjeev Arora and Satyen Kale. A combinatorial, primal-dual approach to semidefinite programs. In *Proceedings of the 39th ACM Symposium on Theory of Computing*, STOC '07, pages 227–236, 2007. 7, 33

[AN04] Noga Alon and Assaf Naor. Approximating the cut-norm via grothendieck's inequality. In *Proceedings of the 36th ACM Symposium on Theory of Computing*, pages 72–80, 2004. 7, 33, 34, 35, 38

[BD22] Guy Blanc and Dean Doron. New near-linear time decodable codes closer to the GV bound. Technical Report TR22-027, Electronic Colloquium on Computational Complexity, 2022. 2, 4, 46

[Bil95] Patrick Billingsley. *Probability and Measure*. J. Wiley and Sons, 1995. 18

[BKS16] Afonso S. Bandeira, Christopher Kennedy, and Amit Singer. Approximating the little grothendieck problem over the orthogonal and unitary groups. *Math. Program.*, 160(1-2), 2016. 40

[BRS11] Boaz Barak, Prasad Raghavendra, and David Steurer. Rounding semidefinite programming hierarchies via global correlation. In *Proceedings of the 52nd IEEE Symposium on Foundations of Computer Science*, pages 472–481, 2011. 3

[Bub15] Sébastien Bubeck. Convex optimization: Algorithms and complexity. *Found. Trends Mach. Learn.*, 8(3-4):231–357, November 2015. 21

[CMR13] Sixia Chen, Cristopher Moore, and Alexander Russell. Small-bias sets for nonabelian groups - derandomizations of the Alon–Roichman theorem. In *APPROX-RANDOM*, volume 8096 of *Lecture Notes in Computer Science*, pages 436–451, 2013. 22

[DD19] Yotam Dikstein and Irit Dinur. Agreement testing theorems on layered set systems. In *Proceedings of the 60th IEEE Symposium on Foundations of Computer Science*, 2019. 4

[DHK+19] Irit Dinur, Prahladh Harsha, Tali Kaufman, Inbal Livni Navon, and Amnon Ta-Shma. List decoding with double samplers. In *Proceedings of the 30th ACM-SIAM Symposium on Discrete Algorithms*, pages 2134–2153, 2019. 2

[DK17] Irit Dinur and Tali Kaufman. High dimensional expanders imply agreement expanders. In *Proceedings of the 58th IEEE Symposium on Foundations of Computer Science*, pages 974–985, 2017. 49

[Elk01] Noam D. Elkies. Excellent codes from modular curves. In *Proceedings of the 33rd ACM Symposium on Theory of Computing*, 2001. 1

[FK96] A. Frieze and R. Kannan. The regularity lemma and approximation schemes for dense problems. In *Proceedings of the 37th IEEE Symposium on Foundations of Computer Science*, 1996. 2

[FK98] Uriel Feige and Joe Kilian. Zero knowledge and the chromatic number. *Journal of Computer and System Sciences*, 57(2):187–199, 1998. 3, 5, 9

[For66] David Forney. *Concatenated Codes*. PhD thesis, MIT, 1966. 1

[GI04] Venkatesan Guruswami and Piotr Indyk. Efficiently decodable codes meeting Gilbert-Varshamov bound for low rates. In *Proceedings of the 15th ACM-SIAM Symposium on Discrete Algorithms*, SODA '04, pages 756–757, 2004. 4

[GI05] V. Guruswami and P. Indyk. Linear-time encodable/decodable codes with near-optimal rate. *IEEE Transactions on Information Theory*, 51(10):3393–3400, 2005. 48

[Gil52] E.N. Gilbert. A comparison of signalling alphabets. *Bell System Technical Journal*, 31:504–522, 1952. 1

[GKO+17] Sivakanth Gopi, Swastik Kopparty, Rafael Oliveira, Noga Ron-Zewi, and Shubhangi Saraf. Locally testable and locally correctable codes approaching the Gilbert-Varshamov bound. In *Proceedings of the 28th ACM-SIAM Symposium on Discrete Algorithms*, SODA '17, pages 2073–2091, 2017. 4

[GRS19] Venkatesan Guruswami, Atri Rudra, and Madhu Sudan. Essential coding theory. Available at https://cse.buffalo.edu/faculty/atri/courses/coding-theory/book/index.html, 2019. 1

[GRZ22] Zeyu Guo and Noga Ron-Zewi. Efficient list-decoding with constant alphabet and list sizes. *IEEE Transactions on Information Theory*, 68(3):1663–1682, 2022. 4

[GS11] Venkatesan Guruswami and Ali Kemal Sinop. Lasserre hierarchy, higher eigenvalues, and approximation schemes for graph partitioning and quadratic integer programming with psd objectives. In *FOCS*, pages 482–491, 2011. 3

[GS12] Venkatesan Guruswami and Ali Kemal Sinop. Faster SDP hierarchy solvers for local rounding algorithms. In *Proceedings of the 53rd IEEE Symposium on Foundations of Computer Science*, pages 197–206. IEEE, 2012. 3

[GX13] Venkatesan Guruswami and Chaoping Xing. List decoding reed-solomon, algebraic-geometric, and gabidulin subcodes up to the singleton bound. In *Proceedings of the 45th ACM Symposium on Theory of Computing*, 2013. 4

[HLW06] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc.*, 43(04):439–562, August 2006. 22

[HRW17] B. Hemenway, N. Ron-Zewi, and M. Wootters. Local list recovery of high-rate tensor codes applications. In *Proceedings of the 58th IEEE Symposium on Foundations of Computer Science*, pages 204–215, Oct 2017. 4

[JM21] Akhil Jalan and Dana Moshkovitz. Near-optimal Cayley expanders for Abelian groups, 2021. arXiv:2105.01149. 1, 7, 47, 49

[JQST20] Fernando Granha Jeronimo, Dylan Quintana, Shashank Srivastava, and Madhur Tulsiani. Unique decoding of explicit ε-balanced codes near the Gilbert–Varshamov bound. In *Proceedings of the 61st IEEE Symposium on Foundations of Computer Science*, 2020. 1, 2, 4, 7, 17, 46

[JST21] Fernando Granha Jeronimo, Shashank Srivastava, and Madhur Tulsiani. Near-linear time decoding of Ta-Shma's codes via splittable regularity. 2021. 1, 2, 3, 4, 5, 7, 8, 9, 12, 16, 19, 20, 22, 24, 25, 26, 27, 28, 29, 32, 33, 38, 41, 48, 50, 54, 55

[Kal07] Satyen Kale. *Efficient Algorithms Using The Multiplicative Weights Update Method*. PhD thesis, Princeton, 2007. 33

[KRRZ+21] Swastik Kopparty, Nicolas Resch, Noga Ron-Zewi, Shubhangi Saraf, and Shashwat Silas. On list recovery of high-rate tensor codes. *IEEE Transactions on Information Theory*, 67(1):296–316, 2021. doi:10.1109/TIT.2020.3023962. 4

[LP20] Yin Tat Lee and Swati Padmanabhan. An $\widetilde{O}(m/\varepsilon^{3.5})$-cost algorithm for semidefinite programs with diagonal constraints. In *COLT 2020*, volume 125, pages 3069–3119, 2020. 33

[LPS88] Alexander Lubotzky, R. Phillips, and Peter Sarnak. Ramanujan graphs. *Combinatorica*, 8:261–277, 1988. 48, 49

[MRRW77] R. McEliece, E. Rodemich, H. Rumsey, and L. Welch. New upper bounds on the rate of a code via the Delsarte-MacWilliams inequalities. *IEEE Transactions on Information Theory*, 23(2):157–166, 1977. 1

[NRV13] Assaf Naor, Oded Regev, and Thomas Vidick. Efficient rounding for the noncommutative Grothendieck inequality. In *Proceedings of the 45th ACM Symposium on Theory of Computing*, page 71–80, 2013. 39, 40

[NW19] Anand Kumar Narayanan and Matthew Weidner. Subquadratic time encodable codes beating the gilbert–varshamov bound. *IEEE Transactions on Information Theory*, 65(10), 2019. 1

[OGT15] Shayan Oveis Gharan and Luca Trevisan. A new regularity lemma and faster approximation algorithms for low threshold rank graphs. *Theory of Computing*, 11(9):241–256, 2015. URL: http://www.theoryofcomputing.org/articles/v011a009, doi:10.4086/toc.2015.v011a009. 3, 9

[Rao19] Anup Rao. Expander codes: Tanner codes. Lecture notes, October 2019. URL: https://homes.cs.washington.edu/~anuprao/pubs/codingtheory/lecture6.pdf. 48

[Rud91] W. Rudin. *Functional Analysis*. International series in pure and applied mathematics. McGraw-Hill, 1991. 20

[Sag13] B.E. Sagan. *The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions*. Graduate Texts in Mathematics. Springer New York, 2013. 15, 54

[SS96] L. L. Scott and J. P. Serre. *Linear Representations of Finite Groups*. Graduate Texts in Mathematics. Springer New York, 1996. 15, 54

[Sti08] Henning Stichtenoth. *Algebraic Function Fields and Codes*. Springer Publishing Company, Incorporated, 2nd edition, 2008. 1

[SZY07] Anthony Man-Cho So, Jiawei Zhang, and Yinyu Ye. On approximating complex quadratic optimization problems via semidefinite programming relaxations. *Math. Program.*, 110(1):93–110, jun 2007. 7, 34

[Tho83] C. Thommesen. The existence of binary linear concatenated codes with Reed-Solomon outer codes which asymptotically meet the Gilbert-Varshamov bound. *IEEE Transactions on Information Theory*, 29(6):850–853, November 1983. 4

[TS17] Amnon Ta-Shma. Explicit, almost optimal, epsilon-balanced codes. In *Proceedings of the 49th ACM Symposium on Theory of Computing*, STOC 2017, pages 238–251, New York, NY, USA, 2017. ACM. 1, 2, 4, 5, 7, 17, 41, 46, 47, 49

[TTV09] L. Trevisan, M. Tulsiani, and S. Vadhan. Boosting, regularity and efficiently simulating every high-entropy distribution. In *Proceedings of the 24th IEEE Conference on Computational Complexity*, 2009. 5

[TVN07] Michael Tsfasman, Serge Vladut, and Dmitry Nogin. *Algebraic Geometric Codes: Basic Notions*. American Mathematical Society, 2007. 1

[TVZ82] Michael A. Tsfasman, S. G. Vlădut, and Thomas Zink. Modular curves, shimura curves, and goppa codes, better than varshamov-gilbert bound. *Mathematische Nachrichten*, 109:21–28, 1982. 1

[Var57] R.R. Varshamov. Estimate of the number of signals in error correcting codes. *Doklady Akademii Nauk SSSR*, 117:739–741, 1957. 1

[Zem01] G. Zemor. On expander codes. *IEEE Transactions on Information Theory*, 47(2):835–837, 2001. 48, 49

A Deferred Proofs

For convenience, we include the proof of the following simple fact (assuming some basic facts from representation theory [SS96, Sag13]).

Fact 3.5 ([SS96])**.** *Let* $\mathfrak{g} \in \mathfrak{G}$*. Then*

$$\mathop{\mathbb{E}}_{\rho \sim \mathrm{Irrep}(G)} \left[\frac{\mathrm{Tr}(\rho(\mathfrak{g}))}{\dim(\rho)} \right] = \mathbf{1}_{[\mathfrak{g}=1]} ,$$

where 1 *is the identity element in* $\mathfrak{G}$*.*

Proof. Let $\rho_{\mathrm{reg}} \colon \mathfrak{G} \to \mathbb{C}^{\mathfrak{G} \times \mathfrak{G}}$ be matrix representation of the regular action of $\mathfrak{G}$ on $\mathbb{C}[\mathfrak{G}]$. If $\mathfrak{g} = 1$, then $\mathrm{Tr}(\rho_{\mathrm{reg}}(\mathfrak{g})) = |\mathfrak{G}|$, otherwise, the action of $\mathfrak{g}$ has no fixed points so $\mathrm{Tr}(\rho_{\mathrm{reg}}(\mathfrak{g})) = 0$. Equivalently, we have $\mathrm{Tr}(\rho_{\mathrm{reg}}(\mathfrak{g}))/\,|\mathfrak{G}| = \mathbf{1}_{[\mathfrak{g}=1]}$. It is well-known that ρ_{reg} is unitarily equivalent to

$$\bigoplus_{\rho \in \mathrm{Irrep}(\mathfrak{G})} \dim(\rho) \cdot \rho \,.$$

This implies that $\mathbb{E}_{\rho \sim \mathrm{Irrep}(\mathfrak{G})}[\mathrm{Tr}(\rho(\mathfrak{g}))/\dim(\rho)] = \mathrm{Tr}(\rho_{\mathrm{reg}}(\mathfrak{g}))/\,|\mathfrak{G}| = \mathbf{1}_{[\mathfrak{g}=1]}$. ∎

To make this presentation self-contained, we recall some omitted proofs from [JST21] restated in the splittable mixing lemma section Section 5.1.

A.1 Splittable Mixing Lemmas

We can iterate Lemma 5.6 to obtain the following.

Lemma 5.8 (Splittable Mixing Lemma Iterated [JST21])**.** *Suppose* $W \subseteq [n]^k$ *is a* τ*-splittable collection of tuples. For every* $f = f_1 \otimes \cdots \otimes f_k \in \mathcal{F}_{k-1}$*, we have*

$$\left| \mathop{\mathbb{E}}_{\nu_0} f \;-\; \mathop{\mathbb{E}}_{\nu_{k-1}} f \right| \;\leq\; (k-1) \cdot \tau.$$

Proof. Let $1 \in \mathcal{F}_{k-1}$ be the constant 1 function. Note that for any $t \in \{0,\dots,k-1\}$ the restriction of any $f' \in \mathcal{F}_{k-1}$ to the support of ν_t which we denote by $f'|_t$ belongs to $\mathcal{F}_t$. It is immediate that $\langle f,1\rangle_{\nu_t} = \langle f|_t,1\rangle_{\nu_t}$. Computing we obtain

$$\begin{aligned}\left|\mathop{\mathbb{E}}_{\nu_0} f \;-\; \mathop{\mathbb{E}}_{\nu_{k-1}} f\right| \;&=\; \left|\langle f,1\rangle_{\nu_0} \;-\; \langle f,1\rangle_{\nu_{k-1}}\right| \;\le\; \sum_{i=0}^{k-2}\left|\langle f,1\rangle_{\nu_i} \;-\; \langle f,1\rangle_{\nu_{i+1}}\right| \\ &=\; \sum_{i=0}^{k-2}\left|\langle f|_t,1|_t\rangle_{\nu_i} \;-\; \langle f|_{t+1},1|_{t+1}\rangle_{\nu_{i+1}}\right| \\ &\le\; \sum_{i=0}^{k-2}\tau, \qquad \text{(By Lemma 5.6)}\end{aligned}$$

finishing the proof. ∎

In Section 5.3, we used two corollaries of the splittable mixing lemma which we prove now.

Claim 5.9 ([JST21]). *Let $W \subseteq [n]^k$ be a τ-splittable collection of tuples. Let $t \in \{0,\dots,k-2\}$ and $h_{t+1} \in \mathcal{H}(R_0,R_1,\mathcal{F}_{t+1})$. For every $f \in \mathcal{F}_{t+1}$, we have*

$$\left|\langle h_{t+1},f\rangle_{\nu_{t+1}} \;-\; \langle h_{t+1},f\rangle_{\nu_t}\right| \;\le\; \tau\cdot R_1.$$

Proof. Since $h_{t+1} \in \mathcal{H}(R_0,R_1,\mathcal{F}_{t+1})$, we can write $h_{t+1} = \sum_\ell c_\ell \cdot f_\ell$, where $f_\ell \in \mathcal{F}_{t+1}$ and $\sum_\ell |c_\ell| \le R_1$. By the splittable mixing lemma, *cf.*, Lemma 5.6, we have

$$\left|\langle h_{t+1},f\rangle_{\nu_{t+1}} \;-\; \langle h_{t+1},f\rangle_{\nu_t}\right| \;\le\; \sum_\ell |c_\ell|\cdot\left|\langle f_\ell,f\rangle_{\nu_{t+1}} \;-\; \langle f_\ell,f\rangle_{\nu_t}\right| \;\le\; \tau\cdot R_1. \qquad ∎$$

Claim 5.10 ([JST21]). *Let $W \subseteq [n]^k$ be a τ-splittable collection of tuples. Let $t \in \{0,\dots,k-2\}$ and $h_{t+1} \in \mathcal{H}(R_0,R_1,\mathcal{F}_{t+1})$. Then*

$$\left|\|h_{t+1}\|^2_{\nu_{t+1}} \;-\; \|h_{t+1}\|^2_{\nu_t}\right| \;\le\; \tau\cdot R_1^2.$$

Proof. Since $h_{t+1} \in \mathcal{H}(R_0,R_1,\mathcal{F}_{t+1})$, we can write $h_{t+1} = \sum_\ell c_\ell \cdot f_\ell$, where $f_\ell \in \mathcal{F}_{t+1}$ and $\sum_\ell |c_\ell| \le R_1$. By the splittable mixing lemma, *cf.*, Lemma 5.6, we have

$$\left|\langle h_{t+1},h_{t+1}\rangle_{\nu_{t+1}} \;-\; \langle h_{t+1},h_{t+1}\rangle_{\nu_t}\right| \;\le\; \sum_{\ell,\ell'} |c_\ell|\cdot|c_{\ell'}|\cdot\left|\langle f_\ell,f_{\ell'}\rangle_{\nu_{t+1}} \;-\; \langle f_\ell,f_{\ell'}\rangle_{\nu_t}\right| \;\le\; \tau\cdot R_1^2. \qquad ∎$$

A.2 Decoding

To handle the technical requirement of a base code in Ta-Shma's construction having a symbol multiplicity property (*cf.*, Definition 7.13), we use the following observation.

Claim A.1 ([JST21]). *Let $\mathcal{C}_0 \subseteq \mathbb{F}_q^n$ be an ε_0-balanced linear code of dimension D_0. Suppose that $\mathcal{C}_0$ is uniquely decodable within (fractional) radius $\delta_0 \in (0,1]$ in time $\mathcal{T}_0(n)$. Let $m \in \mathbb{N}$ and $\mathcal{C} \subseteq \mathbb{F}_q^{m\cdot n}$ be the code formed by replicating m times each codeword from $\mathcal{C}_0$, i.e.,*

$$\mathcal{C} := \{z_1\cdots z_m \in \mathbb{F}_q^{m\cdot n} \mid z_1 = \cdots = z_m \in \mathcal{C}_0\}.$$

Then, $\mathcal{C}$ is an ε_0-balanced linear code of dimension D_0 that can be uniquely decoded within (fractional) radius δ_0 in time $m\cdot\mathcal{T}_0(n) + \widetilde{O}(m^2\cdot n)$.

Proof. The only non-immediate property is the unique decoding guarantees of $\mathcal{C}$. Given $\tilde{y} \in \mathbb{F}_q^{m \cdot n}$ within δ_0 (relative) distance of $\mathcal{C}$. Let β_i be the fraction of errors in the ith $\mathbb{F}_q^n$ component $\tilde{y}$. By assumption $\mathbb{E}_{i \in [m]} \beta_i \leq \delta_0$, so there is at least one of such component that can be correctly uniquely decoded. We issue unique decoding calls for $\mathcal{C}_o$ on each component $i \in [m]$. For each successful decoding say $z \in \mathcal{C}_0$, we let $y = z \dots z \in \mathbb{F}_q^{m \cdot n}$ and check whether $\Delta(\tilde{y}, y) \leq \delta_0$ returning y if this succeeds. Finally, observe that this procedure indeed takes at most the claimed running time. ∎