Maximum Span Hypothesis: A Potentially Weaker Assumption than Gap-ETH for Parameterized Complexity

Karthik C. S.*   Subhash Khot†

Abstract

The Gap Exponential Time Hypothesis rules out FPT algorithms providing (nearly) tight inapproximability results for a host of fundamental problems in parameterized complexity. One of the downsides of working under Gap-ETH is that the assumption is not inherently in the parameterized complexity world, and therefore one of the main research directions is to replace Gap-ETH with weaker assumptions.

In this paper, we propose a hypothesis called the Maximum Span Hypothesis (MSH), which roughly asserts that given a collection of n vectors in $\mathbb{F}_2^{\text{poly}(k)\cdot\log n}$ such that there is a k-dimensional subspace containing $2^{\Omega(k)}$ input vectors, the goal of finding poly(k) input vectors which are contained in some k-dimensional subspace is W[1]-hard.

Assuming MSH, we obtain near optimal inapproximability ratio for the k-clique problem and polynomial inapproximability ratio for the 2-CSP problem (on k variables and alphabet size n). Assuming a strengthening of MSH with additional completeness guarantees, we are able to obtain near optimal inapproximability ratio for the k-biclique problem and some constant inapproximability ratio for the Densest k-subgraph problem. Finally, we prove that Gap-ETH implies a mild version of MSH.

We remark that even a weaker version of MSH (for example, that the task of finding k^2 input vectors which are contained in some k-dimensional subspace, even when we are promised that there is a k-dimensional subspace containing k^5 input vectors, is W[1]-hard) implies improved inapproximability results for the above problems.

*Rutgers University
†New York University

1 Introduction

Approximation and Fixed Parameter Tractability are arguably the two most popular ways to cope with NP-hardness of problems. Thus, Hardness of Approximation and Parameterized Complexity complement the study of approximation algorithms and FPT algorithms respectively. These areas have received significant attention by the Theoretical Computer Science community. This paper deals with problems arising at the intersection of hardness of approximation and parameterized complexity.

Given any computational problem φ (typically NP-hard) and a parameter k of the problem, we say that φ is *Fixed Parameter Tractable* (FPT) if there exists an algorithm that solves φ and runs in time $F(k) \cdot n^c$, where n is the input size, F is some computable function, and c is an absolute constant. For example, the Vertex Cover problem parameterized by the size of the vertex cover is in[1] FPT [DF13, CFK+15]. Another popular example of a problem in FPT is detecting if a graph has a simple path of length k (where k is the parameter of the problem) [AYZ95].

On the other hand, there are many important optimization problems such as the k-Clique problem (for a fixed parameter k, given as input a graph, find a clique of size k in the input) which is believed to not be in FPT. This is formalized by the concept of the W-hierarchy [DF13, CFK+15] and the problems which are W[i]-hard, for any $i \in \mathbb{N}$, are believed to be not in FPT, and k-Clique is the canonical W[1]-complete problem. Thus, a natural way to cope with this hardness is to investigate if there is some computable function T for which approximating k-Clique to $T(k)$ factor, is in FPT.

Beginning with the breakthrough work of Lin [Lin21], and subsequently improved in [KK22, CFLL23], we now know that the k-Clique problem does not admit FPT approximation algorithms for any $k^{o(1)}$-factor of approximation, i.e., assuming W[1] $\neq$ FPT, there is no FPT-time algorithm that given an instance of k-Clique can find a clique of size $k^{1-o(1)}$ (for example, of size $k/\text{polylog}(k)$), even when the input has a clique of size k. However, in terms of FPT algorithms, there is no known non-trivial algorithm which clearly beats (i.e., by more than a constant multiplicative factor) the naive solution of simply outputting a single vertex, which would be a k factor approximation.

Is approximating k-Clique problem to any $o(k)$ factor W[1]-hard?

Gap Exponential Time Hypothesis. Over the last decade, the area of Fine-Grained Complexity has provided deep and refined structural insights on the precise runtime of computational problems by shedding light on distinguishing, say, between problems where exhaustive search is essentially the best possible algorithm, and those that have improved algorithms [Wil15, Wil16, Wil18]. Some of the popular assumptions of Fine-Grained Complexity are the Strong Exponential Time Hypothesis (SETH) [IP01, IPZ01], the Exponential Time Hypothesis (ETH) [IP01, IPZ01], and the Gap Exponential Time Hypothesis (Gap-ETH) [Din16, MR16]. Informally speaking, Gap-ETH asserts that 3-SAT on n variables and $O(n)$ clauses cannot be approximated to $(1 - \delta)$ factor in $2^{o(n)}$ time for some tiny constant $\delta > 0$ (see Section 2.1 for a formal definition).

In fact, the answers to many of the open questions that will be raised in this paper are already completely known if one assumes Gap-ETH instead of W[1] $\neq$ FPT [CCK+20] (including the question above about the approximability of the k-Clique problem). Then again, while Gap-ETH may be plausible, it is a much stronger conjecture than W[1] $\neq$ FPT, and in the works that use them, the hypothesis does much of the work in the proof, as Gap-ETH itself already gives the gap in the hardness of approximation; once they have such a gap, it suffices to design gap preserving reductions to prove other inapproximability results. This is analogous to the NP-world, where once one inapproximability result can be shown, many others follow via relatively simple gap-preserving reductions (see, e.g., [PY91]). However, creating a gap in the first place requires the PCP Theorem [FGL+96, AS98, ALM+98, Din07], which involves several new technical ideas such as local checkability and decodability of codes and proof composition. Hence, it is desirable to bypass Gap-ETH and prove inapproximability results under a standard assumption such as W[1] $\neq$ FPT, that doesn't inherently have a gap.

Another important aspect of ruling out FPT approximation algorithms based on Gap-ETH, that is particularly close to the motivation of this paper is that of *gap amplification.* Typically, for many important problems, assuming Gap-ETH, we can not only rule out non-trivial approximation factor FPT algorithms, but go further, and rule out constant factor approximation algorithms running in time much better than exhaustive search. Case in point, for the k-Clique problem on n vertices, by a simple reduction from Gap 3-SAT, we can rule out algorithms running

[1]We also denote by FPT the class of problems that are fixed parameter tractable.

in time $n^{o(k)}$ which can find a clique of size $(1-\delta)\cdot k$, whenever the graph is promised to have a clique of size k (for some small constant $\delta > 0$). Starting from this strong running time lower bound for the k-Clique problem, we can trade-off the required runtime (as long as we continue to rule out FPT algorithms) to amplify the gap. This is precisely how the authors in [CCK+20] rule out $o(k)$ approximation FPT algorithms for the k-Clique problem (by using standard disperser based arguments).

However, such a gap amplification technique is inherently impossible in the parameterized complexity world, as only a fixed polynomial blowup is allowed in the input size to amplify the gap. Therefore, some of the key techniques that lead to the resolution of the inapproximability of important problems based on Gap-ETH inherently seem useless to prove the same inapproximability results under W[1] $\neq$ FPT. This leads us to the main question addressed in this paper:

Is there a gap assumption that is both rooted in the parameterized complexity world and at the same time yields all the inapproximability results that we know to be true based on Gap-ETH*?*

Parameterized Inapproximability Hypothesis. The celebrated PCP theorem for NP [FGL+96, AS98, ALM+98, Din07] lies at the heart of all NP-hardness of approximation results, and thus provides a framework (along with other ingredients such as the parallel repetition theorem [Raz98] and dictatorship tests [BGS98, Hås01]) through which we can obtain essentially all the known inapproximability results for NP-hard problems.

Inspired by this framework, Lokshtanov et al. put forth a conjecture on constraint satisfaction problems in the parameterized setting, called the *Parameterized Inapproximability Hypothesis* (PIH) [LRSZ20] which would then provide the equivalent PCP theorem based framework for proving non-existence of FPT time approximation algorithms for k-Clique, k-Set Cover, and many other important problems [FKLM20]. Informally speaking, PIH asserts that approximating 2-CSPs on k variables and alphabet size n to $(1-\delta)$ factor is W[1]-hard when parameterized by k, for some tiny constant $\delta > 0$ (see Section 2.1 for a formal definition). Very recently, in a remarkable work, it has been shown that ETH implies PIH, i.e., approximating 2-CSPs on k variables and alphabet size n to $(1-\delta)$ factor is M[1]-hard [GLR+24], adding further credibility to the hypothesis.

At first glance, PIH seems to be exactly the gap assumption that we were looking for, but alas, PIH in the pameterized complexity world is not (known to be) as powerful as the PCP theorem in the NP-world. For instance, PIH immediately implies that approximating k-Clique to any constant factor is W[1]-hard, but it is still open to prove super constant factor W[1]-hardness for k-Clique directly based on PIH. In fact, Lin [Lin21], introduced ideas precisely to prove super constant inapproximability of the k-Clique problem while circumventing proving PIH (and these were heavily used in the works thereafter [KK22, CFLL23]).

Another downside of PIH, is that while Gap-ETH implies optimal inapproximability for k-Biclique and near polynomial inapproximability results for Densest k-Subgraph, PIH cannot even prove constant inapproximability for both these fundamental graph problems. To exacerbate this situation, even if one considers a strengthening of PIH, namely that approximating 2-CSPs on k variables and alphabet size n to δ factor is W[1]-hard even when $\delta = k^{1-o(1)}$, then we would obtain (near)-optimal inapproximability results for 2-CSP (by definition) and Clique (via FGLSS reduction [FGL+96]), but to the best of our knowledge, would imply nothing for k-Biclique and Densest k-Subgraph problems.

It is worth remarking here that even in the NP-world, the Biclique and Densest k-Subgraph problems are notoriously difficult problems to prove hardness of approximation results: the question of even ruling out PTAS for these two problems under NP$\neq$P is still wide open! That said, under strong assumptions, hardness of approximation results for the Biclique problem [Fei02, Kho06, BGH+17, Man17b] and for the Densest k-Subgraph problem [Fei02, Kho06, AAM+11, Man17a] are known.

Thus, while PIH is an important hypothesis whose truth would assert inapproximability results for some important problems in parameterized complexity, it alone does not seem sufficient to prove strong inapproximability W[1]-hardness results for many of the fundamental problems studied in the field.

1.1 Maximum Span Hypothesis (MSH) To remedy the situation, we introduce a problem, called the *Maximum Span Problem* (MSP) in parameterized complexity, which is closely related to the well-studied k-Vector Sum problem and the parameterized Minimum Distance of a Code Problem (a.k.a. Even set problem).

Elaborating, we define MSP as a gap problem, where for any two computable functions Y and N, we define an instance of the (Y, N)-MSP through a set of n vectors $X \subseteq \mathbb{F}_2^{O(c\log n)}$ and a parameter k (the constant c in the dimension of the space can depend on k), and the goal is to determine if either there exist some $Y(k)$ vectors in X that can be contained in a k-dimensional space or if every $N(k)+1$ vectors in X contain $k+1$ linearly independent vectors (see Definition 3.1 for a formal definition).

MSP may be seen as the compliment of the gap version of the parameterized Minimum Distance of a Code Problem, (or more appropriately, as the compliment of the parameterized *Linear Dependent Set Problem*; see [BBE+21] for the formal definition), where given as input a set of vectors over $\mathbb{F}_2$ and a parameter k, the goal is to determine if either there exist some k input vectors that are linearly dependent or if every $\gamma \cdot k$ input vectors are linearly independent (for any $\gamma \geqslant 1$).

On the other hand, we can view (Y,N)-MSP as a generalization of the k-Vector Sum problem, where given as input a set of n vectors $X \subseteq \mathbb{F}_2^{O(c \log n)}$ and a parameter k (the constant c in the dimension of the space can depend on k), the goal is to determine if either there exist some k vectors in X that can be contained in a $(k-1)$-dimensional space or if every k vectors in X are linearly independent vectors (see Section 2.1 for a formal definition). Thus if we have for all $i \in \mathbb{N}$, $Y(i) := N(i) - 1 := i$, then for such Y and N, (Y,N)-MSP is equivalent to the k-Vector Sum problem. In fact, since k-Vector Sum problem is W[1]-hard [ALW13], this implies that for aforementioned Y and N, we also have (Y,N)-MSP is W[1]-hard (See Proposition 3.1).

The main conceptual contribution of this paper is identifying that the hardness of (Y,N)-MSP would imply strong inapproximability results for fundamental problems in parameterized complexity. To this effect, we introduce the *Maximum Span Hypothesis* (MSH), which asserts that (Y,N)-MSP remains W[1]-hard even when $Y(k) = 2^{\Omega(k)}$ and $N(k) := \text{poly}(k)$ (see Hypothesis 3.1 for a formal version).

We provide some evidence that MSH is plausibly true by showing that a weaker version is implied from Gap-ETH.

Theorem 1.1. *Assuming* Gap-ETH*, there exist some polynomial functions $Y, N : \mathbb{N} \to \mathbb{N}$ and a constant $\delta > 0$ such that $\forall i \in \mathbb{N}$, $Y(i) \geqslant (1+\delta) \cdot N(i)$, and no algorithm running in $F(k) \cdot poly(n)$ time (for any computable function $F : \mathbb{N} \to \mathbb{N}$) can decide all instances of* (Y,N)-MSP *on n input vectors*[2].

As we will see in the next subsection, MSH is sufficient to prove strong inapproximability results for k-Clique and 2-CSP (in particular, MSH implies that PIH is true), however, in order to prove inapproximability results for k-Biclique and Densest k-Subgraph, we need a slight strengthening of MSH. Thus, we put forth the *Strong Maximum Span Hypothesis* (Strong MSH), which concerns a generalization of MSP to larger fields.

Elaborating, we first define the *Maximum Totally Span Problem* (MTSP), where for any two computable functions Y and N, we define an instance of the (Y,N)-MTSP through a set of n vectors $X \subseteq \mathbb{F}_q^{O(c \log n)}$ (for some prime q), a threshold value ω, and a parameter k (the constant c in the dimension of the space can depend on k), and the goal is to determine if either there exist some $Y(k)$ vectors $\widetilde{X} \subseteq X$ such that $\widetilde{X}$ is a k-dimensional subspace and moreover every subset of $\widetilde{X}$ of size at least $\omega \cdot Y(k)$ spans a k-dimensional space, or if every $N(k)+1$ vectors in X contain $k+1$ linearly independent vectors (see Definition 7.1 for a formal definition). Then, the Strong MSH asserts that (Y,N)-MTSP remains W[1]-hard even when $Y(k) = q^{\Omega(k)}$ and $N(k) := \text{poly}(k)$ (and $\omega = 1/q^{\Omega(1)}$).

1.2 Consequences of MSH In this subsection, we show that assuming MSH and Strong MSH, we obtain strong inapproximability results for four fundamental problems in parameterized complexity: k-Clique, 2-CSP (on k variables), k-Biclique, and Densest k-Subgraph. The inapproximability results for these four problems imply numerous other hardness of approximation results for various other important problems (for example see the survey [FKLM20]).

In a breakthrough work, Lin [Lin21] proved the W[1]-hardness result for the k-Clique problem for any constant gap factor. This has been subsequently improved to obtain W[1]-hardness result for any $k^{o(1)}$ gap [KK22, CFLL23]. On the other hand, a relatively simple proof[3] based on Gap-ETH implies that approximating k-Clique to any $o(k)$ cannot be done in FPT time. Our result based on MSH nearly matches the conclusion from Gap-ETH.

Theorem 1.2. (Inapproximability of k-Clique) *Assuming* MSH*, there is some constant $C \in \mathbb{N}$ such that given as input a graph G and a parameter k, it is* W[1]*-hard (under randomized reductions) to decide between the following two cases:*

[2]We prove the conditional lower bound for a colored version of MSP. Please see Theorem 8.1 for details.

[3]The proof in [CCK+20] is involved but has been simplified in the exposition in [FKLM20].

Completeness: *G contains a clique of size k.*

Soundness: *G does not contain a clique of size $(\log k)^C$.*

We remark here that $k^{1-o(1)}$ approximation factor FPT algorithms have been recently ruled out under ETH [LRSW23b] (building on [LRSW22]). While this result proves a very strong inapproximability result based on a non-gap assumption, it does perform gap amplification by trading runtime, and as described earlier in this section, such techniques are unlikely to translate to prove W[1]-hardness of approximation for these problems.

Moving our discussion to 2-CSP, Gap-ETH is not only known to imply PIH, but also implies that 2-CSPs on k variables (the parameter) and alphabet size n cannot be approximated to $\frac{k}{2^{(\log k)^{0.5+o(1)}}}$ factor [DM18], i.e., assuming Gap-ETH, given a 2-CSP which admits a satisfying assignment, there is no FPT algorithm which can find an assignment that satisfies $\frac{2^{(\log k)^{0.5+o(1)}}}{k}$ fraction of the constraints. Note that satisfying $O(1/k)$ fraction of the constraints is trivial and also note that PIH is still open, i.e., we do not even know how to prove constant W[1]-hardness of approximation for 2-CSPs. While our result based on MSH doesn't match the conclusion from Gap-ETH, it nevertheless provides a strong polynomial factor gap.

THEOREM 1.3. (INAPPROXIMABILITY OF 2-CSP) *Assuming* MSH*, there is some constant $C \in \mathbb{N}$ such that given as input a 2-*CSP *instance Φ on k variables and alphabet set Σ, it is* W[1]*-hard parameterized by k (under randomized reductions) to decide between the following two cases:*

Completeness: *Φ admits a satisfying assignment.*

Soundness: *Every assignment to Φ satisfies at most $\frac{(\log k)^C}{\sqrt{k}}$ fraction of the constraints.*

We now shift our attention to results based on Strong MSH. Even proving the W[1]-hardness of exactly solving the k-Biclique problem was a major open problem and only resolved in the last decade [Lin18]. So it is not surprising that there is no W[1]-hardness of approximation result known for this problem. On the other hand, we can completely rule out $o(k)$ approximation FPT algorithms based on Gap-ETH (again owing mainly to the gap amplification via runtime trade-off technique that we had earlier discussed). Thus, it comes as a mild surprise, that we are able to prove near optimal W[1]-hardness of approximation result for the k-Biclique problem based on Strong MSH.

THEOREM 1.4. (INAPPROXIMABILITY OF k-Biclique) *Assuming* Strong MSH*, there is some constant $C \in \mathbb{N}$ such that given as input a bipartite graph G and a parameter k, it is* W[1]*-hard (under randomized reductions) to decide between the following two cases:*

Completeness: *G contains a balanced biclique of size k.*

Soundness: *G does not contain a balanced biclique of size $(\log k)^C$.*

Our final result concerns the Densest k-Subgraph problem, a notorious problem even in the NP-world where even a constant factor NP-hardness is not yet known. Even under the powerful Gap-ETH, only almost polynomial factor FPT algorithms can be ruled out. Assuming Strong MSH, we are able to prove that Densest k-Subgraph is W[1]-hard to some small constant factor.

THEOREM 1.5. (INAPPROXIMABILITY OF Densest k-Subgraph) *Assuming* Strong MSH*, there is some constant $C > 1$ such that given as input a graph G, density parameter $\rho \in (0, 1]$, and a parameter k, it is* W[1]*-hard (under randomized reductions) to decide between the following two cases:*

Completeness: *G contains a subgraph on k vertices with $\rho \cdot \binom{k}{2}$ edges.*

Soundness: *G does not contain a subgraph on k vertices with $\frac{\rho}{C} \cdot \binom{k}{2}$ edges.*

In Table 1, we have summarized the discussion above.

State-of-the-art Inapproximability Results for Some Parameterized Complexity Problems

Assumption	k-Clique	2-CSP	k-Biclique	Densest k-Subgraph
W[1] $\neq$FPT	$k^{o(1)}$ [KK22, CFLL23]	Exact[4] Folklore	Exact [Lin18]	Exact Folklore
ETH	$\frac{k}{(\log k)^{\Omega(\log\log\log k)}} = k^{1-o(1)}$ [LRSW23b]	$(\log k)^{O(1/\sqrt{\log\log k})}$ [GLR+24][5]	Exact [Lin18]	Exact Folklore
Gap-ETH	$o(k)$ [CCK+20]	$k^{1-o(1)}$ [DM18]	$o(k)$ [CCK+20]	$k^{o(1)}$ [CCK+20]
MSH	$\frac{k}{(\log k)^{\Omega(1)}} = k^{1-o(1)}$ This paper	$k^{0.5+o(1)}$ This paper	–	–
Strong MSH	$\frac{k}{(\log k)^{\Omega(1)}} = k^{1-o(1)}$ This paper	$k^{0.5+o(1)}$ This paper	$k^{1-o(1)}$ This paper	$\Omega(1)$ This paper

Table 1: Summary of previous results ruling out FPT algorithms (under W[1] $\neq$ FPT, ETH, Gap-ETH) and our results (under MSH and Strong MSH). Note that MSH is a stronger assumption than W[1] $\neq$ FPT, a potentially weaker assumption than Gap-ETH, and incomparable to ETH. We have excluded listing some other important problems, such as parameterized set cover from the table; please see Remark 1.1 for an explanation.

Remark 1.1. *In Table 1, we have excluded listing the inapproximability results some important problems such as for parameterized Set Cover [CL19, KLM19, Lin19, KN21, LRSW23a], k-max coverage [KLM19, Man20, KLM24], parameterized Minimum Distance of a Code [BGKM18, BBE+21, BCGR23, Man20, LLL24], parameterized Shortest Vector in a Lattice [BGKM18, BBE+21, BCGR23, Man20, LLL24], and parameterized Set Intersection problem [Lin18, BKN21]. This is because for all the aforementioned problems, the inapproximability results (to rule out* FPT *algorithms) achieved under* W[1] $\neq$ FPT, *match the hardness factors obtained under* Gap-ETH.

1.3 Merits of Maximum Span Hypothesis Here, we briefly discuss and summarize some of the merits of MSH (and Strong MSH).

A hypothesis which is inside FPT world. First, the advantage of MSH (or Strong MSH) is that it is inherently a parameterized complexity hypothesis (unlike Gap-ETH, ETH, or SETH). Thus, it has the additional advantage that MSH is not only a tool but a target; Attempts to prove MSH has a direct impact on our understanding of parameterized complexity. It is also potentially strictly weaker than Gap-ETH– we proved Theorem 1.1 mainly to show some indications that MSH might be true, but we believe it might be possible to show the full implication that Gap-ETH implies MSH (and leave this for future work).

Advantages over PIH. The four theorems in the previous subsection are ample evidence that MSH is much more powerful and has broader applicability than PIH, which itself has been very useful in understanding the

[4] Technically speaking, it is possible to obtain hardness of approximation factor of $1-\frac{1}{F(k)}$ for any increasing computable function F by simple padding construction but we do not write down this detail as the factor that can be achieved is subconstant. Also, it is possible to interpret the results in [Lin21, KK22, LRSW22, LRSW23b, GLR+24] as progress on understanding the W[1]-hardness of approximating 2-CSP, although there is no succinct way to quantify this over the inapproximability factor of the simple padding argument. Recently, in [GRS24], the authors proved a (weak) minimization version of PIH.

[5] In [GLR+24] the authors prove that assuming ETH, 2-CSP cannot be approximated to some constant factor in $n^{\Omega(\sqrt{\log\log k})}$ time. We can then apply parallel repetition theorem [Raz98] $\Omega(\sqrt{\log\log k})$-fold to obtain the stated inapproximability factor against FPT algorithms.

landscape of FPT (in)-approximability. Additionally, the arithmetic structure in MSP makes it more suitable as a starting problem (as compared to generic 2-CSPs), and this brings us to our next point.

An approximate version of k-Vector Sum problem. As briefly touched upon, MSP can be seen as a gap version of k-Vector Sum, in fact one may even think of MSP as an approximate version of the counting version of k-Vector Sum, i.e., in the completeness case, there are many $k+1$ tuples of input vectors that sum to $\vec{0}$ and in the soundness case there are fewer such tuples, of course, we need to enforce that in both cases we are looking for solutions in a subspace.

Another important aspect that we would like to bring to light is that all the recent results on inapproximability of k-Clique, or even PIH, have either started from the hardness of exact k-Vector Sum or have reached it as an intermediate step (this is true explicitly in [Lin21, LRSW22, KK22] and implicitly through Sidon sets in [CFLL23] and Vector-Valued CSPs in [LRSW23b, LLL24]). MSH and its consequences thus highlight that proving hardness of approximation results for k-Vector Sum is potentially a very fertile approach to pursue. These previous papers also help place MSP and MSH in the bigger picture of the problems studied in the area. For example, it is conceivable that the threshold graph technique might prove useful in making progress on MSH (as it proved useful in obtaining the constant inapproximability of the Minimum Distance of a code problem, which we had discussed earlier).

Partial Progress on Proving MSH implies Improved Inapproximability Results. One of the most important remarks that we would like to make is that proving *any* hardness of approximation result for MSP (resp. MTSP) implies hardness of approximation results for k-Clique and 2-CSP (resp. k-Biclique and Densest k-Subgraph). In particular, Theorems 5.1, 6.1, 7.1 and 7.2 are written in a way where they translate the gap in MSP and MTSP to the corresponding target problems. To the best of our knowledge, this is the first concrete approach to tackle parameterized inapproximability of k-Biclique and Densest k-Subgraph (in a holistic way where we make progress on understanding the approximability of a host of problems simultaneously).

1.4 Proof Techniques

In this subsection, we provide an overview of the proof techniques involved in proving the theorems mentioned in Section 1.2.

1.4.1 Inapproximability of k-Clique

The proof of Theorem 1.2 proceeds by reducing an instance of (Y,N)-MSP to an instance of k-Clique problem. Given a set of n vectors $X \subseteq \mathbb{F}_2^{c\log n}$ and a parameter k, we first mentally associate with it two matrices. The first matrix $A_X \in \mathbb{F}_2^{c\log n \times k}$, which we refer to as the "short hand matrix" contains as columns the basis vectors whose span supposedly contains $Y(k)$ vectors in X (if we are in the completeness case). The second matrix $\widetilde{A}_X \in \mathbb{F}_2^{c\log n\times 2^k}$, which we refer to as the "full matrix" contain as columns the span of the columns of A_X, and in particular allegedly contains $Y(k)$ vectors from X as its columns (again if we are in the completeness case). Then, we build a 2-CSP Φ whose assignment together corresponds to the entries of A_X and constraints correspond to local checks via row-column consistency checks, that there is an underlying matrix $\widetilde{A}_X$ that is encoded by the assignment to its variables.

Elaborating, we first encode the vectors in X using some good code, so that every pair of vectors in X have some constant relative Hamming distance (say 0.1). Then, we construct a 2-CSP Φ where for every $i\in Y(k)$, we have a variable v_i in Φ. We identify the alphabet set with the set $\mathbb{F}_2^{\log n}\times\mathbb{F}_2^k\times X$ and thus it contains $O_k(n^2)$ many labels. We will now explain what this alphabet set is capturing.

Imagine that the matrices A_X and $\widetilde{A}_X$ are not filled in. For each variable v_i, uniformly and independently at random sample a set of m rows of A_X and denote this set of rows by S_i (where $m := \frac{\log n}{k}$). As part of the assignment to variable v_i we assign to it the entries in the matrix A_X corresponding only to the rows in S_i. This is a submatrix in $\mathbb{F}_2^{m\times k}$ which we think of as an element in $\mathbb{F}_2^{\log n}$.

In addition, to a variable being assigned to m entire rows of A_X, each variable is also assigned one column in $\widetilde{A}_X$ which corresponds to some vector in X and the column is indexed by its linear combination coefficients, i.e., a vector in $\mathbb{F}_2^k$. Thus, each variable is assigned an element in $\mathbb{F}_2^{\log n}\times\mathbb{F}_2^k\times X$.

Therefore, we can think of the labels to a variable v_i to be of the form $(\vec{\alpha}_1^i,\ldots,\vec{\alpha}_k^i,\vec{\beta}^i,\vec{x}^i)$, where $\vec{\alpha}_j^i\in\mathbb{F}_2^m$ corresponds to the entries of the j^{th} row of the m rows in S_i, $\vec{\beta}^i\in\mathbb{F}_2^k$ corresponds to the index of the column of $\widetilde{A}_X$ and $\vec{x}^i$ corresponds to the column in that location which is also a vector in X.

The constraints of Φ simply check that:

- Every pair of variables are assigned distinct vectors in X and they are given distinct column indices in $\widetilde{A}_X$

where they would fill.

- Every pair of variables consistently fill the rows of A_X that they have in common.

- The row entries of one variable for the full matrix are consistent with the column entries of every other variable for the short form matrix. More formally, for every pair of variables v_i and v_j we have that:

$$\sum_{r\in[k]} \vec{\beta}^i(r)\cdot\vec{\alpha}^j_r = \vec{x}^i\mid_{S_j},$$

i.e., the linear combination of the $\vec{\alpha}^j$s as specified by the coefficient vector $\vec{\beta}^i$ is exactly equal to $\vec{x}^i$ when restricted to the coordinates of S_j.

The $Y(k)$-Clique instance is defined as the so called FGLSS graph of Φ, where for every variable and an assignment to that variable we have a node in the graph. A pair of nodes, say $(v_i,\vec{\alpha}^i_1,\ldots,\vec{\alpha}^i_k,\vec{\beta}^i,\vec{x}^i)$ and $(v_j,\vec{\alpha}^j_1,\ldots,\vec{\alpha}^j_k,\vec{\beta}^j,\vec{x}^j)$ are adjacent in the graph if and only if, the assignments $(\vec{\alpha}^i_1,\ldots,\vec{\alpha}^i_k,\vec{\beta}^i,\vec{x}^i)$ and $(\vec{\alpha}^j_1,\ldots,\vec{\alpha}^j_k,\vec{\beta}^j,\vec{x}^j)$ to v_i and v_j respectively, satisfy the constraint between those two variables.

It is easy to see that if there exist some $Y(k)$ vectors in X that can be contained in a k-dimensional space, then we can fill up the matrices A_X and $\widetilde{A}_X$ using the k-dimensional basis and the $Y(k)$ vectors in X and from that extract an assignment to Φ that satisfies all the constraints. Thus, we would have a clique of size $Y(k)$ in the above graph.

On the other hand, if every $N(k)+1$ vectors in X contain $k+1$ linearly independent vectors then, we will show that there is no $N(k)+1$ sized clique in the graph. The proof approach is by contradiction, suppose there was a clique of size $t>N(k)$ in the graph then they must correspond to the assignment to t distinct variables of Φ. Let us denote them by $v_1,\ldots,v_t$. Look at their corresponding sets of rows in A_X, i.e., $S_1,\ldots,S_t$. Since t is large enough (we will assume that $N(k)=\omega(ck)$), we know that these t columns essentially cover all the rows of A_X. This is referred to as the *Disperser property* (see Section 2.4). Therefore through these t rows, one can extract almost the entire entries of $\widetilde{A}_X$. Let the union of these rows be denoted by the set $\widetilde{S}$ and we can conclude that $|\widetilde{S}|\geqslant 0.95c\log n$.

However, each of these t variables have also committed to a column of $\widetilde{A}_X$. Thus, we have committed to t distinct columns of $\widetilde{A}_X$ (which are vectors in X), and in these t columns are $k+1$ linearly independent vectors but their restriction to the coordinates of $\widetilde{S}$ makes us think of them as contained in a k-dimensional space. This is impossible, because we had encoded the vectors in X using a code of relative distance 0.1, and thus any linear combination of these $k+1$ linearly independent vectors must have 1 entries on at least 10% of their coordinates, but at the same time if they sum to $\vec{0}$ on 95% of their coordinates, we have a contradiction.

1.4.2 Inapproximability of 2-CSP The completeness case analysis of Φ was already done above. So we will focus on the soundness analysis, i.e., if every $N(k)+1$ vectors in X contain $k+1$ linearly independent vectors then, we will show that no assignment to Φ can satisfy $\delta:=\frac{3N(k)}{\sqrt{Y(k)}}$ fraction of clauses. Suppose otherwise, and we have an assignment satisfying δ fraction of the constraints.

We would have ideally liked to prove some kind of direct product theorem here (a standard proof strategy for these kinds of analysis) to say that each constraint reveals m entries of the matrix $\widetilde{A}_X$ (i.e., the rows-column intersection), and if δ fraction of the constraints are revealed then, we have obtained commitment from the assignment to $m\cdot\delta\cdot(Y(k))^2$ entries of the matrix, which is sufficient to recover $N(k)+1$ vectors in X which lie in a k-dimensional space. Unfortunately, we ran into several issues to make these arguments go through, as there is not strong commitment to the same information from sufficiently many variables.

To remedy this, we took a more "local" approach and proved the following technical lemma.

Informal Version of Deceitful Subspaces Lemma (See Lemma 4.3). Let $X\subseteq\mathbb{F}_2^{c\log n}$ be a set of vectors whose pairwise relative distance is at least 0.1. Let $S_1,\ldots,S_{Y(k)}$ be a random collection of m rows of A_X. For every $R\subseteq X$, we say a set of m rows S is bad w.r.t. R if $\mathsf{rank}(R)>k$ but $\mathsf{rank}(R|_S)\leqslant k$. Let $\mathcal{S}_R\subseteq\{S_1,\ldots,S_{Y(k)}\}$ be the collection of set of coordinates each of which is bad w.r.t. R. Then with high probability, we have that for all $R\subseteq X$ of size γ, we have $|\mathcal{S}_R|\leqslant 8\gamma k$.

Given the Deceitful Subspaces Lemma, the soundness analysis proceeds as follows. By an averaging argument, we pick a set $\widetilde{V}$ of $\frac{\delta}{3}$ fraction of the variables all of whom have the guarantee that at least $\frac{2\delta}{3}$ fraction of the constraints incident on them are satisfied. Let $R \subseteq X$ be the vectors assigned as columns to the variables in $\widetilde{V}$.

We now apply the Deceitful Subspaces Lemma on this set R and obtain the collection of bad sets of rows $\mathcal{S}_R$. Each of the set of rows in $\mathcal{S}_R$ itself corresponds to a variable in Φ. Let V_R be the set of variables corresponding to $\mathcal{S}_R$. The crucial observation is that for every variable not in V_R there are at most $N(k)$ variables in $\widetilde{V}$ with whom the variable can share a satisfied constraint (follows from the soundness assumption). On the other hand, by our choice, we had that every variable in $\widetilde{V}$ has (relatively speaking) a lot of constriants incident on it being satisfied. This leads to a contradiction.

1.4.3 Inapproximability of k-Biclique and Densest k-Subgraph The analysis for the hardness of approximation of k-Biclique and Densest k-Subgraph are quite involved and we only provide here the rough construction of the hard instance from (Y, N)-MTSP (with some details skipped).

For both the problems, we build the same bipartite graph on vertex set $R \dot\cup C$ where the size of R is $Y(k) \cdot |\Gamma|$ and the size of C is $Y(k) \cdot n$ (Γ is the collection of all k-dimensional spaces over the vectors in $\mathbb{F}_2^m$). Both R and C are equipartitioned into $Y(k)$ parts, say $R_1, \ldots, R_{Y(k)}$ and $C_1, \ldots, C_{Y(k)}$ respectively (we can think of these set of nodes as assignments to a variable, i.e., we may visualize this as a 2-CSP on $2 \cdot Y(k)$ variables). For all $i \in [Y(k)]$, we think of the vertices in R_i as a copy of Γ and the vertices in C_i as a copy of X. The edges simply check the rows-column consistency as before.

The first distinction here is that for every $i \in [Y(k)]$ we only allow variable v_i (on the R side) to fill the rows S_i of A_X with k linearly independent vectors. The second distinction here is that each variable no longer holds both a set of rows and a column, and thus doing consistency checks is even more cumbersome, but luckily the Deceitful Subspaces Lemma (along with the Disperser property) is powerful enough for our purposes.

1.4.4 Gap-ETH Guarantee: Reduction from 3-SAT to MSP Given a 3-SAT formula φ on n variables, we equipartition it's variable set to k parts. We look at the set of all partial assignments to each of the k parts as vectors in $\mathbb{F}_2^n$ where for every partial assignment to the i^{th} part is viewed as a vector in $\mathbb{F}_2^n$, and the entry for a variable/coordinate not in the i^{th} part is 0 by default. Thus, we have k collections of vectors in $\mathbb{F}_2^n$ where each collection is of size $2^{n/k}$. For some small constant $t \in \mathbb{N}$ (for example, think of t as 10), the idea is to look at all t parts out of the k parts and take the bit-wise XORs of all possible t-tuples of vectors, one from each part. Thus, we now have $\binom{k}{t}$ collections of vectors in $\mathbb{F}_2^n$ where each collection is of size $2^{nt/k}$. Finally for each t-tuple of collections, we remove all the points that do not satisfy all the clauses all of whose variables appear in the t-tuple of parts of variables. The final point-set is our instance of MSP. The intuition is that by construction, in the completeness case, we have initially k vectors corresponding to the k parts of the satisfying assignment, and then they are made to span to obtain $\binom{k}{t}$ input points of MSP instance (and these points would never been thrown away because they satisfy all clauses). The soundness analysis is more intricate but follows from looking at $(t-1)$-tuples to decode an almost satisfying assignment.

1.5 Organization of Paper In Section 2 we introduce the problems, hypotheses, and other tools relevant to this paper. In Section 3, we formally define MSH. In Section 4, we develop some tools that will be helpful for our proofs. In Section 5, we prove the near optimal inapproximability of the k-Clique problem under MSH. In Section 6 we prove the polynomial factor inapproximability for the 2-CSP problem under MSH. In Section 7 we prove the near optimal inapproximability of the k-Biclique problem under Strong MSH and also prove the constant inapproximability of Densest k-Subgraph under the same. In Section 8, we show that Gap-ETH implies a weak version of MSH.

2 Preliminaries

Notations. Let $\mathbb{N}$ be the set of natural numbers and $\mathbb{P}$ be the set of prime numbers. For every non-negative integers n and k, we denote by $\binom{[n]}{k}$, the collection of all subsets of $[n]$ of size exactly k.

2.1 Problems In this subsection, we recall the definitions of the relevant computational problems to this paper.

3-SAT. In the 3-SAT problem, we are given a CNF formula φ over n variables $x_1, \ldots x_n$, such that each clause contains at most 3 literals. Our goal is to decide if there exist an assignment to $x_1, \ldots x_n$ which satisfies φ.

k-Clique problem. In the k-Clique problem we receive a graph $H = (V, E)$ with $|V| = n$, and our goal is to decide if H contains a clique of size k, i.e. there exists $v_1, \ldots v_k \in V$ such that for every $i \neq j \in [k]$, $(v_i, v_j) \in E$.

Densest k-Subgraph problem. In the Densest k-Subgraph problem we receive a graph $H = (V, E)$ with $|V| = n$ and $\alpha := \alpha(k) \in (0, 1]$, and our goal is to decide if H contains a subgraph on k vertices with at least $\alpha \cdot \binom{k}{2}$ edges.

k-Biclique problem. In the k-Biclique problem we receive a bipartite graph $H = (V, E)$ with $|V| = n$, and our goal is to decide if H contains a balanced complete bipartite graph of size $2k$, i.e. there exists a copy of $K_{k,k}$ as an induced subgraph in H.

2-CSP problem. An instance Γ of 2-CSP consists of

- an undirected graph $G = (V, E)$, which is referred to as the *constraint graph*,
- an *alphabet set* Σ,
- for each edge $e = (u, v) \in E$, a *constraint* $C_{uv} \subseteq \Sigma \times \Sigma$.

An *assignment* of Γ is simply a function from V to Σ. An edge $e = (u, v) \in E$ is said to be *satisfied* by an assignment $\psi : V \to \Sigma$ if $(\psi(u), \psi(v)) \in C_{uv}$. A *value* of an assignment ψ, denoted by $\mathsf{val}(\psi)$, is the fraction of edges satisfied by ψ, i.e., $\mathsf{val}(\psi) = \frac{1}{|E|} \cdot \{(u, v) \in E \mid (\psi(u), \psi(v)) \in C_{uv}\}$. The value of the instance Γ, denoted by $\mathsf{val}(\Gamma)$, is the maximum value among all possible assignments, i.e., $\mathsf{val}(\Gamma) = \max_{\psi:V\to\Sigma} \mathsf{val}(\psi)$.

k-Vector Sum Given k sets $U_1, \ldots, U_k$ of vectors in $\mathbb{F}_2^m$, the goal of k-Vector Sum problem is to decide whether there exist $\vec{u}_1 \in U_1, \ldots, \vec{u}_k \in U_k$ such that

$$\sum_{i\in[k]} \vec{u}_i = \vec{0}.$$

It is known that the above problem is W[1]-hard over finite fields [ALW13]. We direct the reader to [Lin21] for a short proof[6].

THEOREM 2.1. *[ALW13, Lin21] The k-Vector Sum over $\mathbb{F}_2$ and $m = \Theta(k^2 \log n)$ is W[1]-hard parameterized by k.*

By a simple gadget reduction, we can show that for some constant c, given a set U of vectors in $\mathbb{F}_2^{ck^2 \log n}$, deciding whether there exist $\vec{u}_1, \ldots, \vec{u}_k \in U$ such that

$$\sum_{i\in[k]} \vec{u}_i = \vec{0},$$

is still W[1]-complete. This is referred to as the *monochromatic* k-Vector Sum problem.

2.2 Hypotheses

In this subsection, we recall the relevant computational hypotheses that will be used in this paper.

HYPOTHESIS 2.1. (W[1] $\neq$ FPT) *For any computable function $F : \mathbb{N} \to \mathbb{N}$, there is no $F(k)poly(n)$-time algorithm which solves the k-Clique problem over n vertices.*

HYPOTHESIS 2.2. (EXPONENTIAL TIME HYPOTHESIS (ETH) [IP01, IPZ01, TOV84]) *There exists $\varepsilon > 0$ such that no algorithm can solve 3-SAT on n variables in time $O(2^{\varepsilon n})$.*

HYPOTHESIS 2.3. (GAP EXPONENTIAL TIME HYPOTHESIS (Gap-ETH) [DIN16, MR16]) *There exist constants $\varepsilon, \delta > 0$ such that any randomized algorithm that, on input a 3-SAT formula φ on n variables and $O(n)$ clauses, can distinguish between $\mathsf{val}(\varphi) = 1$ and $\mathsf{val}(\varphi) < 1 - \delta$, must run in time at least $2^{\varepsilon n}$.*

Moreover, using standard expander replacement arguments (for example, see [PY91, Fei98]), we may assume that Gap-ETH holds even when each variable appears in at most 5 clauses.

[6] [Lin21] proves the hardness result for a version of k-Vector Sum where a target vector is given as input, but that version reduces to the version given in this paper by simply including an extra collection containing only the negative of the target vector.

Hypothesis 2.4. (Parameterized Inapproximability Hypothesis (PIH) [LRSZ20]) *There exists a constant $\varepsilon > 0$ such that any algorithm that, on input a 2-CSP instance Γ on k variables and alphabet size n, can distinguish between $\mathsf{val}(\Gamma) = 1$ and $\mathsf{val}(\Gamma) < 1 - \varepsilon$, cannot run in FPT time when parameterized by k.*

2.3 Error Correcting Codes In this subsection, we recall the definition of error correcting codes and some standard code constructions known in the literature. We define below a notion of distance used in coding theory (called *Hamming* distance) and then define error correcting codes with its various parameters.

Definition 2.1. (Distance) *Let $q \in \mathbb{P}$. Let $\ell \in \mathbb{N}$. The distance[7] between $\vec{x}, \vec{y} \in \mathbb{F}_q^\ell$, denoted by $\Delta(\vec{x}, \vec{y})$, is defined to be:*

$$\Delta(\vec{x}, \vec{y}) = \frac{1}{\ell} \cdot |\{i \in [\ell] \mid \vec{x}(i) \neq \vec{y}(i)\}| .$$

Definition 2.2. (Error Correcting Code) *Let $q \in \mathbb{P}$. For every $\ell \in \mathbb{N}$, a subset $C \subseteq \mathbb{F}_q^\ell$ is said to be an error correcting code with block length ℓ, message length k, and relative distance δ if $|C| = q^k$ and for every $\vec{x}, \vec{y} \in C$, $\Delta(\vec{x}, \vec{y}) \geq \delta$. We refer to C as a $[k, \ell, \delta]$ code and to the elements of a code C as codewords. Sometimes, we think of a code C as some canonical bijective function $E_C : \mathbb{F}_q^k \to C$. Moreover, given an injective function from $\mathbb{F}_q^k$ to $\mathbb{F}_q^\ell$, the code associated with it is simply the image set of the function.*

Definition 2.3. (Linear Codes) *We say that a $[k, \ell, \delta]$ code C is a linear code if for all $\vec{x}, \vec{y} \in C$, we have that $\vec{x} + \vec{y}$ also is in C (where the addition is done coordinate-wise over $\mathbb{F}_q$).*

Fact 2.1. (See Appendix E.1.2.5 in [Gol08]) *For every $q \in \mathbb{P}$ and every constant $\delta < 0.5$, there is an integer ℓ such that for all $k \in \mathbb{N}$, there is a linear code C with message length k, block length $\ell \cdot k$, and relative distance δ, and an encoding algorithm which on input a point in $\mathbb{F}_q^k$, outputs in time polynomial in k the image of the input under E_C.*

2.4 Disperser In this subsection, we recall the definition of a disperser and some acheivable parameters for the same.

Definition 2.4. (Disperser [CW89, Zuc96a, Zuc96b]) *For positive integers $m, k, \ell, r \in \mathbb{N}$ and constant $\varepsilon \in (0, 1)$, an $(m, k, \ell, r, \varepsilon)$-disperser is a collection $\mathcal{S}$ of k subsets $S_1, \ldots, S_k \subseteq [m]$, each of size ℓ, such that the union of any r different subsets from the collection has size at least $(1 - \varepsilon)m$.*

Dispersers could be constructed efficiently by probabilistic methods, as in the following Lemma.

Lemma 2.1. (For example, see [LRSW22]) *For positive integers $m, \ell, r \in \mathbb{N}$ and constant $\varepsilon \in (0, 1)$, let $\ell = \lceil \frac{3m}{\varepsilon r} \rceil$ and let $S_1, \ldots, S_k$ be random ℓ-subsets of $[m]$. If $\ln k \leqslant \frac{m}{r}$ then $\mathcal{S} = \{S_1, \ldots, S_k\}$ is an $(m, k, \ell, r, \varepsilon)$-disperser with probability at least $1 - e^{-m}$.*

3 k-Maximum Span Problem and Maximum Span Hypothesis

In this section, we formally define the k-Maximum Span Problem and Maximum Span Hypothesis.

Let $d \in \mathbb{N}$. For any sets of vectors $S, T \subseteq \mathbb{F}_2^d$ we say that T spans S if and only if $S \subseteq \mathsf{span}(T)$.

Definition 3.1. ((Y, N)-Maximum Span Problem ((Y, N)-MSP)) *Given two computable functions $Y, N : \mathbb{N} \to \mathbb{N}$, an instance of (Y, N)-MSP is specified by the tuple (X, d, k) where $X \subseteq \mathbb{F}_2^d$, k is the parameter, and the goal is to distinguish between the two cases:*

Completeness: *There exists a set $T \subseteq \mathbb{F}_2^d$ of k vectors such that*

$$|X \cap \mathsf{span}(T)| \geqslant Y(k).$$

[7] We use the normalized notion of distance for the sake of exposition. In coding theory literature, our notion of distance is referred to as *relative distance*.

Soundness: *For every set $T \subseteq \mathbb{F}_2^d$ of k vectors we have*

$$|X \cap \mathsf{span}(T)| \leqslant N(k).$$

We remark that while we define MSP over $\mathbb{F}_2$, in Section 7.1, we define the problem over $\mathbb{F}_q$ with a stronger completeness guarantee. Next, we note that the exact version of the problem is W[1]-hard.

PROPOSITION 3.1. *Let $Y, N : \mathbb{N} \to \mathbb{N}$ be two computable functions such that for all $i \in \mathbb{N}$ we have $N(i) := Y(i) - 1$ and $Y(i) \geqslant i$. Then (Y, N)-MSP is W[1]-hard even for instances (X, d, k) where $d = O(k^2 \log |X|)$.*

Proof. Fix $Y, N : \mathbb{N} \to \mathbb{N}$ such that for all $i \in \mathbb{N}$ we have $N(i) := Y(i) - 1$ and $Y(i) \geqslant i$. The proof is by a reduction from the monochromatic k-Vector Sum problem. Given a set U of vectors in $\mathbb{F}_2^{ck^2 \log n}$ of the monochromatic k-Vector Sum problem, we can equivalently think of it as an $(|U|, ck^2 \log |U|, k - 1)$ instance of (Y', N')-MSP where for all $i \in \mathbb{N}$, $Y'(i) := i$ and $N'(i) = i - 1$. We can extend this hardness result for every other $Y : \mathbb{N} \to \mathbb{N}$ by a simple padding argument. □

Finally, we define the Maximum Span Hypothesis which informally asserts that the above W[1]-hardness extends to a gap version.

HYPOTHESIS 3.1. (MAXIMUM SPAN HYPOTHESIS (MSH)) *There exist constants $\delta, \rho > 0$, $\eta, \zeta \in \mathbb{N}$, computable functions $Y, N : \mathbb{N} \to \mathbb{N}$, and a polynomial function $D : \mathbb{N} \to \mathbb{N}$ such that the following holds.*

- *For all $i \in \mathbb{N}$ we have $Y(i) \geqslant \rho \cdot 2^{\delta \cdot i}$ and $N(i) \leqslant \eta \cdot i^{\zeta}$.*
- *(Y, N)-MSP is W[1]-hard even for instances (X, d, k) where $d = D(k) \cdot \log |X|$.*

4 Mise-en-place for Reductions based on MSH

In this section, we construct a few technical tools that will be used throughout the paper. In Section 4.1 we prove that we can assume that the hard instances of (Y, N)-MSP have the additional property that every pair of input vectors are at relative distance 0.1 or more from each other. In Section 4.2, we prove that we can assume that the hard instances of (Y, N)-MSP have the additional property that the input set can be partitioned into $Y(k)$ color classes and that we demand only for solutions in which we pick one vector from each color class. In Section 4.3, we describe a mapping of (Y, N)-MSP to instances of the 2-CSP problem. This will be later recalled in Sections 5 and 6. Finally, in Section 4.4, we prove a lemma about "deceitful" subspaces which will be heavily used throughout the paper.

4.1 Preprocessing Step I: Encoding Using Error Correcting Codes Recall the definitions of Codes from Section 2.3. We have the below transformation for (Y, N)-MSP instances.

LEMMA 4.1. *There exists a polynomial time algorithm and a constant $\ell \in \mathbb{N}$ which takes as input an instance (X_0, d, k) of (Y, N)-MSP and outputs an instance $(X, \ell \cdot d, k)$ of (Y, N)-MSP with the following guarantees:*

Size: $|X_0| = |X|$.

Distance: *For every $\vec{\lambda} \in \mathbb{F}_2^{|X|}$, let $\vec{x}^{\vec{\lambda}} := \sum_{\vec{x} \in X} \vec{\lambda}(\vec{x}) \cdot \vec{x}$. If $\vec{x}^{\vec{\lambda}} \neq \vec{0}$ then $\Delta(\vec{x}^{\vec{\lambda}}, \vec{0}) \geqslant 0.1$.*

Gap Preservation: *For every $t \geqslant k$ the following holds. There are t vectors in X_0 contained in a k-dimensional subspace of $\mathbb{F}_2^d$ if and only if there are t vectors in X contained in a k-dimensional subspace of $\mathbb{F}_2^{\ell \cdot d}$.*

Proof. Let C be a linear error correcting code of block length $\ell \cdot d$, message length d and relative distance 0.1 guaranteed by Fact 2.1. The ℓ in the theorem statement is precisely as specified by Fact 2.1.

Given as input an instance (X_0, d, k) of (Y, N)-MSP, the output instance $(X, \ell \cdot d, k)$ of (Y, N)-MSP is given by

$$X := \{E_C(\vec{x}_0) \mid \vec{x}_0 \in X_0\}.$$

The distance property in the theorem statement immediately follows from the distance property of C.

Suppose there exists a set $T := \{\vec{z}_1, \ldots, \vec{z}_k\} \subseteq \mathbb{F}_2^d$ of k vectors such that there exists $\widetilde{X}_0 \subseteq X_0 \cap \mathsf{span}(T)$, where $|\widetilde{X}_0| = t$. Let $T' := \{E_C(\vec{z}_i) \mid i \in [k]\} \subseteq \mathbb{F}_2^{\ell \cdot d}$ and $\widetilde{X} := \{E_C(\vec{x}_0) \mid \vec{x}_0 \in \widetilde{X}_0\} \subseteq X$. Then we claim that $\widetilde{X} \subseteq \mathsf{span}(T')$. To see this note that for every $\vec{x}_0 \in \widetilde{X}_0$, if $\vec{x}_0 = \sum_{i \in [k]} \lambda_i \cdot \vec{z}_k$ for some scalars $\lambda_1, \ldots, \lambda_k \in \mathbb{F}_2$, then we have the following for $y := E_C(\vec{x}_0) \in \widetilde{X}$:

$$\vec{x} = E_C(\vec{x}_0) = E_C\left(\sum_{i \in [k]} \lambda_i \cdot \vec{z}_k\right) = \sum_{i \in [k]} \lambda_i \cdot E_C(\vec{z}_k),$$

where we used that C is a linear code for the last equality. Also note that $|\widetilde{X}| = t$, because E_C is bijective.

On the other hand, suppose there exists a set $T \subseteq \mathbb{F}_2^{\ell \cdot d}$ of k vectors such that there exists $\widetilde{X} \subseteq X \cap \mathsf{span}(T)$, where $|\widetilde{X}| = t$. Let $A_X \subseteq \mathbb{F}_2^{d \times t}$ be the matrix whose columns are the vectors in $\widetilde{X}$. Let $r := \mathsf{rank}(A_X) \leqslant k$. Let $\vec{x}_1, \ldots, \vec{x}_r$ be the r linearly independent column vectors of A_X. Let $\widetilde{X}_0 := \{E_C^{-1}(\vec{x}) \mid \vec{x} \in \widetilde{X}\} \subseteq X_0$. For every $\vec{x} \in \widetilde{X}$ we have that there are scalars $\lambda_1^{\vec{x}}, \ldots, \lambda_r^{\vec{x}} \in \mathbb{F}_2$, such that:

$$\vec{x} = \sum_{i \in [r]} \lambda_i^{\vec{x}} \cdot \vec{x}_i.$$

Thus, we have that for every $\vec{x}_0 \in \widetilde{X}_0$,

$$\vec{x}_0 = E_C^{-1}(\vec{x}) = E_C^{-1}\left(\sum_{i \in [r]} \lambda_i^{\vec{x}} \cdot \vec{x}_i\right) = \sum_{i \in [r]} \lambda_i^{\vec{x}} \cdot E_C^{-1}(\vec{x}_i).$$

This implies that $\widetilde{X}_0$ is spanned by the vectors $E_C^{-1}(\vec{x}_1), \ldots, E_C^{-1}(\vec{x}_r)$.

Finally note that the procedure runs in polynomial time as E_C requires poly(d) time for constructing each point in X. ∎

4.2 Preprocessing Step II: Product Structure from Color Coding First, we define below a colored version of (Y, N)-MSP.

Definition 4.1. ((Y, N)-Colored Maximum Span Problem ((Y, N)-Colored-MSP)) *Given two computable functions $Y, N : \mathbb{N} \to \mathbb{N}$, an instance of (Y, N)-Colored-MSP is specified by the tuple $(X := X_1 \dot\cup X_2 \dot\cup \cdots \dot\cup X_{Y(k)}, d, k)$ where for all $i \in [Y(k)]$, we have $X_i \subseteq \mathbb{F}_2^d$ (for all $i, i' \in [Y(k)]$, we have $|X_i| = |X_{i'}|$), k is the parameter, and the goal is to distinguish between the two cases:*

Completeness: *There exists a set $T \subseteq \mathbb{F}_2^d$ of k vectors such that for all $i \in [Y(k)]$,*

$$|X_i \cap \mathsf{span}(T)| \geqslant 1.$$

Soundness: *For every set $T \subseteq \mathbb{F}_2^d$ of k vectors we have*

$$\sum_{i \in [Y(k)]} |X_i \cap \mathsf{span}(T)| \leqslant N(k).$$

We obtain the following product structure for the hard instances of (Y, N)-MSP for free as a simple application of the color coding technique [AYZ95] (which can be derandomized using perfect hash families [NSS95]).

Lemma 4.2. *There exists a randomized polynomial time algorithm and a constant $\ell \in \mathbb{N}$ which takes as input an instance (X, d, k) of (Y, N)-MSP and outputs an instance $(X := X_1 \dot\cup X_2 \dot\cup \cdots \dot\cup X_{Y(k)}, d, k)$ of (Y, N)-Colored-MSP with the following guarantees:*

Completeness: *If there exists a set $T \subseteq \mathbb{F}_2^d$ of k vectors such that*

$$|X \cap \mathsf{span}(T)| \geqslant Y(k).$$

Then with probability at least $2^{-\widetilde{\Omega}(Y(k))}$, for all $i \in [Y(k)]$,

$$|X_i \cap \mathsf{span}(T)| \geqslant 1.$$

Soundness: *Suppose for every set $T \subseteq \mathbb{F}_2^d$ of k vectors we have*

$$|X \cap \mathsf{span}(T)| \leqslant N(k).$$

Then for every set $T \subseteq \mathbb{F}_2^d$ of k vectors we have

$$\sum_{i \in [Y(k)]} |X_i \cap \mathsf{span}(T)| \leqslant N(k).$$

Proof. The proof follows by simply taking a uniformly random equipartition of X to $Y(k)$ parts. ∎

4.3 Construction of the Hard 2-CSP Instance Φ In this subsection, we describe the construction of a 2-CSP instance that we will construct for each instance of (Y, N)-MSP.

For some $c_0 \in \mathbb{Z}$, starting from an instance of $(X_0, c_0 \log n, k)$ of (Y, N)-MSP (where $n := |X_0|$), we apply the algorithm in Lemma 4.1 to obtain an instance $(X, c \log n, k)$ of (Y, N)-MSP, where $c := c_0 \cdot \ell$, for some constant ℓ. Moreover, we have that every pair of vectors in X are at relative Hamming distance greater than or equal to 0.1.

From such an instance $(X, c \log n, k)$ of MSP, we construct a 2-CSP Φ on variable set V and alphabet set Σ as follows. For every $i \in [Y(k)]$, we have a variable v_i in V. We identify the alphabet set Σ with the set $\mathbb{F}_2^{\log n} \times \mathbb{F}_2^k \times X$ and thus it contains $O_k(n^2)$ many labels.

Let $m := \frac{\log n}{k}$. Let $S_1, \ldots, S_{Y(k)}$ be random subsets of $[c \log n]$ of size m. Let $\pi_i : S_i \to [m]$ be some canonical 1-to-1 mapping.

We are now ready to define the constraints. Fix $i, j \in [Y(k)]$ where $i \neq j$. We define the constraint $C_{i,j} \subseteq \Sigma \times \Sigma$ between variables v_i and v_j as follows. Given label $(\vec{\alpha}_1^i, \ldots, \vec{\alpha}_k^i, \vec{\beta}^i, \vec{x}^i)$ for v_i and label $(\vec{\alpha}_1^j, \ldots, \vec{\alpha}_k^j, \vec{\beta}^j, \vec{x}^j)$ for v_j, where $\vec{\alpha}_1^i, \ldots, \vec{\alpha}_k^i \in \mathbb{F}_2^m$, $\vec{\alpha}_1^j, \ldots, \vec{\alpha}_k^j \in \mathbb{F}_2^m$, $\vec{\beta}^i, \vec{\beta}^j \in \mathbb{F}_2^k$, and $\vec{x}^i, \vec{x}^j \in X$, we have that the pair of labels is in $C_{i,j}$ if and only if all of the following hold:

Distinctness Check: $\vec{\beta}^i \neq \vec{\beta}^j$ and $\vec{x}^i \neq \vec{x}^j$.

Row-Column Intra-Consistency Check:

$$\sum_{r=1}^{k} \vec{\beta}^i(r) \cdot \vec{\alpha}_r^i = \left(\vec{x}^i\left(\pi_i^{-1}(1)\right), \vec{x}^i\left(\pi_i^{-1}(2)\right), \ldots, \vec{x}^i\left(\pi_i^{-1}(m)\right)\right)$$

and

$$\sum_{r=1}^{k} \vec{\beta}^j(r) \cdot \vec{\alpha}_r^j = \left(\vec{x}^j(\pi_j^{-1}(1)), \vec{x}^j(\pi_j^{-1}(2)), \ldots, \vec{x}^j\left(\pi_j^{-1}(m)\right)\right).$$

Row-Column Inter-Consistency Check:

$$\sum_{r=1}^{k} \vec{\beta}^i(r) \cdot \vec{\alpha}_r^j = \left(\vec{x}^i\left(\pi_j^{-1}(1)\right), \vec{x}^i\left(\pi_j^{-1}(2)\right), \ldots, \vec{x}^i\left(\pi_j^{-1}(m)\right)\right)$$

and

$$\sum_{r=1}^{k} \vec{\beta}^j(r) \cdot \vec{\alpha}_r^i = \left(\vec{x}^j(\pi_i^{-1}(1)), \vec{x}^j(\pi_i^{-1}(2)), \ldots, \vec{x}^j\left(\pi_i^{-1}(m)\right)\right).$$

Column-Column Consistency Check: $\forall s \in S_i \cap S_j$ and $\forall r \in [k]$ we have,

$$\vec{\alpha}_r^i(\pi_i(s)) = \vec{\alpha}_r^j(\pi_j(s)).$$

4.4 Techincal Tool: Deceitful Subspaces Lemma In this subsection, we prove a lemma about the probability that a random axis-parallel projection of a set of linearly independent vectors are not independent under the projection.

LEMMA 4.3. (DECEITFUL SUBSPACES LEMMA) *Let $q \in \mathbb{P}$, $k \in \mathbb{N}$, $c := c(k) \in \mathbb{N}$, $n \in \mathbb{N}$, $Y : \mathbb{N} \to \mathbb{N}$, and $m := \frac{\log n}{k \log q}$. Let $\kappa, \gamma \in \mathbb{N}$ be such that $\kappa \geqslant 8\gamma k \log q$. Let $X := \{\vec{x}^1, \ldots, \vec{x}^n\} \subseteq C \subseteq \mathbb{F}_q^{c \log n}$ be a set of vectors where C is a $[t, c\log n, 0.1]$ linear code, for any $t \geqslant \log_q n$. Let $S_1, \ldots, S_{Y(k)}$ be a random collection of m-sized subsets of $[c \log n]$, where $Y(k) > \kappa$. For every $R := \{\vec{x}^{i_1}, \ldots, \vec{x}^{i_\gamma}\} \subseteq X$, we say $S \in \binom{[c \log n]}{m}$ is bad w.r.t. R if there exists $\vec{\lambda} \in \mathbb{F}_q^\gamma$ such that the following holds:*

$$\sum_{r \in [\gamma]} \vec{\lambda}(r) \cdot \vec{x}^{i_r} \neq \vec{0} \text{ but } \sum_{r \in [\gamma]} \vec{\lambda}(r) \cdot \left(\vec{x}^{i_r}\big|_S\right) = \vec{0}.$$

Let $\mathcal{S}_R \subseteq \{S_1, \ldots, S_{Y(k)}\}$ be the collection of set of coordinates each of which is bad w.r.t. R. Then with probability at least $1 - \frac{2^{Y(k)+\kappa^2 \log q}}{n^{\gamma/7}}$, we have that for all $R \subseteq X$ of size γ, we have $|\mathcal{S}_R| \leqslant \kappa$.

Proof. Pick $S_1, \ldots, S_{Y(k)}$ to be a random collection of m-sized subsets of $[c \log n]$, where $Y(k) > \kappa$. Fix $R := \{\vec{x}^{i_1}, \ldots, \vec{x}^{i_\gamma}\} \subseteq X$ such that $|R| = \gamma$. Note that for every $\vec{\lambda} \in \mathbb{F}_q^n$, we have from the linearity of C:

$$\sum_{i=1}^{n} \vec{\lambda}(i) \cdot \vec{x}^i \neq \vec{0} \implies \Delta\left(\sum_{i=1}^{n} \vec{\lambda}(i) \cdot \vec{x}^i, \vec{0}\right) \geqslant 0.1.$$

Now fix any $\vec{\lambda} \in \mathbb{F}_q^\gamma \setminus \{\vec{0}\}$. Let $\vec{z} := \sum_{h \in [\gamma]} \vec{\lambda}(h) \cdot \vec{x}^{i_h}$. If $\vec{z} \neq \vec{0}$ then we have $\Delta\left(\vec{z}, \vec{0}\right) \geqslant 0.1$. Let $Z \subseteq [c \log n]$ be defined as follows: $\forall t \in [c \log n]$, we have $t \in Z$ if and only if $\vec{z}(t) = 0$. Note that $|Z| \leqslant 0.9c \log n$. The probability that for a random subset S of $[c \log n]$ of size m, we have $S \subseteq Z$ is at most:

$$\frac{\binom{0.9c \log n}{m}}{\binom{c \log n}{m}} = \frac{\prod_{a \in [m]} (0.9c \log n - a + 1)}{\prod_{a \in [m]} (c \log n - a + 1)} = \prod_{a \in [m]} \left(\frac{0.9c \log n - a + 1}{c \log n - a + 1}\right) \leqslant 0.9^m.$$

Thus, the probability that a random set $S \in \binom{[c \log n]}{m}$ is bad w.r.t. R is at most:

$$(q^\gamma - 1) \cdot 0.9^m.$$

The probability of having κ or more subsets in $S_1, \ldots, S_{Y(k)}$ which are all bad w.r.t. R is:

$$\sum_{\ell=\kappa}^{Y(k)} \binom{Y(k)}{\ell} \cdot \left((q^\gamma - 1) \cdot 0.9^m\right)^\ell \leqslant 2^{Y(k)} \cdot \left(q^\gamma \cdot 0.9^m\right)^\kappa.$$

The probability of having some $R \subseteq X$ of size γ which has more than κ subsets in $S_1, \ldots, S_{Y(k)}$ which are all bad w.r.t. R is at most:

$$\begin{aligned} n^\gamma \cdot 2^{Y(k)} \cdot \left(q^\gamma \cdot 0.9^m\right)^\kappa &\leqslant 2^{Y(k)+\gamma \log n + \gamma\kappa \log q} \cdot 0.9^{m\kappa} \\ &\leqslant 2^{Y(k)+\gamma \log n + \gamma\kappa \log q - \frac{m\kappa}{7}} \\ &\leqslant 2^{Y(k)+\gamma \log n + \kappa^2 \log q - \frac{8}{7}\gamma \log n} \\ &= \frac{2^{Y(k)+\kappa^2 \log q}}{n^{\gamma/7}}. \end{aligned}$$

□

5 Near Optimal Inapproximability of k-Clique Problem

In this subsection, we prove the near optimal inapproximability of the k-Clique problem.

THEOREM 5.1. *There is a randomized algorithm running in* FPT *time which takes as input an instance* $(X, d := c\log|X|, k)$ *of* (Y,N)-MSP *and outputs an instance* G *(of order* $O_k(|X|^2)$*) of* $Y(k)$-Clique *problem such that the following holds:*

Completeness: *If there exists a set* $T \subseteq \mathbb{F}_2^d$ *of* k *vectors such that*

$$|X \cap \mathsf{span}(T)| \geqslant Y(k),$$

then, with probability 1, G *has a clique of size* $Y(k)$*.*

Soundness: *If for every set* $T \subseteq \mathbb{F}_2^d$ *of* k *vectors we have*

$$|X \cap \mathsf{span}(T)| \leqslant N(k),$$

then, if $N(k) \geqslant 60ck$*, with probability* $1 - \frac{1}{|X|^{\Omega(1)}}$*, we have that* G *does not have a clique of size* $N(k)$*.*

Assuming the above theorem, we have the proof of Theorem 1.2.

Proof. [Proof of Theorem 1.2]

As promised by MSH, let there be computable functions $Y, N : \mathbb{N} \to \mathbb{N}$, polynomial function $D : \mathbb{N} \to \mathbb{N}$, and constants $\delta, \rho > 0$ and $\eta, \zeta \in \mathbb{N}$ such that for all $i \in \mathbb{N}$ we have $Y(i) \geqslant \rho \cdot 2^{\delta \cdot i}$, $N(i) \leqslant \eta \cdot i^{\zeta}$, and (Y,N)-MSP being W[1]-hard even for instances (X,d,k) where $d = D(k)\cdot\log|X|$. We may assume that $N(k) \geqslant 60D(k)\cdot k$, otherwise we revise the constants η and ζ accordingly.

We then apply the FPT reduction to instances of $Y(k)$-Clique problem as given in Theorem 5.1 and obtain that it is W[1]-hard to distinguish instances of $Y(k)$-Clique problem where there is a $\rho \cdot 2^{\delta k}$ sized clique versus instances of $Y(k)$-Clique problem where there is no $1 + \eta \cdot k^{\zeta}$ sized clique. Thus, if the parameter is set to $k' := Y(k) \geqslant \rho \cdot 2^{\delta k}$, then distinguishing the case where there is a k' sized clique versus the case there is no clique of size $1 + \eta \cdot k^{\zeta} \leqslant 1 + \frac{\eta}{\delta^{\zeta}}(\log k'/\rho)^{\zeta} = (\log k')^{O(1)}$, is W[1]-hard. ∎

The rest of this section is dedicated to proving Theorem 5.1.

Construction of k-Clique instance. Let $V := \{v_1, \ldots, v_{Y(k)}\}$ be the variable set of the 2-CSP Φ from Section 4.3. We build the so-called FGLSS graph[8] G of Φ [FGL+96] whose vertex set is $V \times \Sigma$ and we think of the vertex set as $|V|$ copies of the set Σ where the i^{th} copy of Σ corresponding to the labels of variable v_i is denoted by U_i.

Each U_i is an independent set. Given $(\vec{\alpha}_1^i, \ldots, \vec{\alpha}_k^i, \vec{\beta}^i, \vec{x}^i) \in U_i$ and $(\vec{\alpha}_1^j, \ldots, \vec{\alpha}_k^j, \vec{\beta}^j, \vec{x}^j) \in U_j$ where $\vec{\alpha}_1^i, \ldots, \vec{\alpha}_k^i \in \mathbb{F}_2^m$, $\vec{\alpha}_1^j, \ldots, \vec{\alpha}_k^j \in \mathbb{F}_2^m$, $\vec{\beta}^i, \vec{\beta}^j \in \mathbb{F}_2^k$, and $\vec{x}^i, \vec{x}^j \in X$, we have an edge between them if and only if the pair of labels belongs to $C_{i,j}$.

5.1 Completeness

Suppose there exists a set $T := \{\vec{z}_1, \ldots, \vec{z}_k\} \subseteq \mathbb{F}_2^{c\log n}$ of k vectors such that there exists $\widetilde{X} \subseteq X \cap \mathsf{span}(T)$, such that $|\widetilde{X}| = Y(k)$. Then we define 2^k vectors in $\mathbb{F}_2^{c\log n}$ as follows:

$$\forall \vec{\beta} \in \mathbb{F}_2^k,\ \vec{\psi}_{\vec{\beta}} := \sum_{\ell \in [k]} \vec{\beta}(\ell) \cdot \vec{z}_\ell.$$

Let $\pi : \widetilde{X} \to \mathbb{F}_2^k$ be the mapping defined as follows: $\forall \vec{x} \in \widetilde{X}$, we set $\pi(\vec{x}) := \vec{\beta}$ where we have $\vec{\psi}_{\vec{\beta}} = \vec{x}$. Moreover, let $\widetilde{X} := \{\vec{x}^1, \ldots, \vec{x}^{Y(k)}\}$.

For every $i \in [Y(k)]$, we pick the vertex $u_i^* := (\vec{\alpha}_1^i, \ldots, \vec{\alpha}_k^i, \vec{\beta}^i, \vec{x}^i) \in U_i$ defined as follows:

$$\forall r \in [k],\ \vec{\alpha}_r^i := \vec{z}_r|_{S_i} \text{ and } \vec{\beta}^i := \pi(\vec{x}^i).$$

To finish the analysis in the completeness case, we need to show that $\{u_1^*, \ldots, u_{Y(k)}^*\}$ is a clique in G. Fix some arbitrary $i, j \in [Y(k)]$, such that $i \neq j$. We will show that $\{u_i^*, u_j^*\} \in E(G)$.

[8]It is more accurate to think of it as an extended label version of Φ.

Passing Distinctness Check: By the distinctness of the elements in $\widetilde{X}$, we have that $\vec{x}^i \neq \vec{x}^j$. Moreover, note that $\vec{\beta}^i = \pi(\vec{x}^i)$ and $\vec{\beta}^j = \pi(\vec{x}^j)$. Thus, we have that $\vec{\psi}_{\vec{\beta}^i} = \vec{x}^i$ and $\vec{\psi}_{\vec{\beta}^j} = \vec{x}^j$. Since $\vec{x}^i \neq \vec{x}^j$ we have that $\vec{\psi}_{\vec{\beta}^i} \neq \vec{\psi}_{\vec{\beta}^j}$. This implies that $\vec{\beta}^i \neq \vec{\beta}^j$. Thus, $\{u_i^*, u_j^*\}$ passes the distinctness check.

Passing Row-Column Intra-Consistency Check: Fix some $t \in [m]$. We want to show that:

$$\sum_{r=1}^{k} \vec{\beta}^i(r) \cdot \vec{\alpha}_r^i(t) = \vec{x}^i\left(\pi_i^{-1}(t)\right) \text{ and } \sum_{r=1}^{k} \vec{\beta}^j(r) \cdot \vec{\alpha}_r^j(t) = \vec{x}^j(\pi_j^{-1}(t)).$$

We focus first on showing the former of the above two as the latter follows in a similar way.

$$\sum_{r=1}^{k} \vec{\beta}^i(r) \cdot \vec{\alpha}_r^i(t) = \sum_{r=1}^{k} \vec{\beta}^i(r) \cdot \vec{z}_r(\pi_i^{-1}(t)) = \sum_{r=1}^{k} \pi(\vec{x}^i)_r \cdot \vec{z}_r(\pi_i^{-1}(t)) = \vec{x}^i(\pi_i^{-1}(t))$$

Since the choice of t was arbitrary, it holds for all $t \in [m]$. Thus, $\{u_i^*, u_j^*\}$ passes the row-column intra-consistency check.

Passing Row-Column Inter-Consistency Check: Fix some $t \in [m]$. We want to show that:

$$\sum_{r=1}^{k} \vec{\beta}^i(r) \cdot \vec{\alpha}_r^j(t) = \vec{x}^i\left(\pi_j^{-1}(t)\right) \text{ and } \sum_{r=1}^{k} \vec{\beta}^j(r) \cdot \vec{\alpha}_r^i(t) = \vec{x}^j(\pi_i^{-1}(t)).$$

We focus first on showing the former of the above two as the latter follows in a similar way.

$$\sum_{r=1}^{k} \vec{\beta}^i(r) \cdot \vec{\alpha}_r^j(t) = \sum_{r=1}^{k} \vec{\beta}^i(r) \cdot \vec{z}_r(\pi_j^{-1}(t)) = \sum_{r=1}^{k} \pi(\vec{x}^i)_r \cdot \vec{z}_r(\pi_j^{-1}(t)) = \vec{x}^i(\pi_j^{-1}(t)).$$

Since the choice of t was arbitrary, it holds for all $t \in [m]$. Thus, $\{u_i^*, u_j^*\}$ passes the row-column inter-consistency check.

Passing Column-Column Consistency Check: Fix some $s \in S_i \cap S_j$ and fix some $r \in [k]$.

$$\vec{\alpha}_r^i(\pi_i(s)) = \vec{z}_r|_s = \vec{\alpha}_r^j(\pi_j(s)).$$

Since the choice of s and r were arbitrary, it holds for all $s \in S_i \cap S_j$ and for all $r \in [k]$. Thus, $\{u_i^*, u_j^*\}$ passes the column-column consistency check.

Since all the checks have passed we have that $\{u_i^*, u_j^*\} \in E(G)$. Since the choice of i and j was arbitrary, we have that $\{u_1^*, \ldots, u_{Y(k)}^*\}$ is a clique in G.

5.2 Soundness Let t be any integer greater than $N(k) \geqslant 60ck$. To do the soundness analysis, suppose $\{u_{a_1}, \ldots, u_{a_t}\} \in U_{a_1} \times \cdots \times U_{a_t}$ is a clique in G. We will relabel the V_is and assume that $\{u_1, \ldots, u_t\} \in U_1 \times \cdots \times U_t$ is a clique in G, where for all $i \in [t]$, we have $u_i := (\vec{\alpha}_1^i, \ldots, \vec{\alpha}_k^i, \vec{\beta}^i, \vec{x}^i) \in U_i$.

From the distinctness and membership checks, we know there exists a set $X' = \{\vec{x}^1, \ldots, \vec{x}^t\}$ of t vectors in X and t many distinct coefficient vectors $B := \{\vec{\beta}^1, \ldots, \vec{\beta}^t\}$.

Let $A_{X'} \subseteq \mathbb{F}_2^{c \log n \times t}$ be the matrix whose columns are the vectors in X'. Let $\gamma := \mathsf{rank}(A_{X'}) \geqslant k+1$ (from the soundness assumption). Let $\vec{x}^{i_1}, \ldots, \vec{x}^{i_\gamma}$ be the γ linearly independent column vectors of $A_{X'}$. Thus, for every $\vec{\lambda} \in \mathbb{F}_2^\gamma \setminus \vec{0}$, we have:

$$\sum_{w \in [\gamma]} \vec{\lambda}(w) \cdot \vec{x}^{i_w} \neq \vec{0}.$$

Since $\vec{x}^{i_1}, \dots, \vec{x}^{i_\gamma}$ are codewords of a linear code with relative distance greater than 0.1, we have that for every $\vec{\lambda} \in \mathbb{F}_2^\gamma \setminus \vec{0}$, we have:

$$\Delta\left(\sum_{w\in[\gamma]} \vec{\lambda}(w) \cdot \vec{x}^{i_w}, \vec{0}\right) \geqslant 0.1.$$

This further implies that for any set $T \subseteq [c\log n]$ of coordinates, if $|T| > 0.9c\log n$ then,

$$\left(\sum_{w\in[\gamma]} \vec{\lambda}(w) \cdot \vec{x}^{i_w}\right)\Big|_T \neq \vec{0}. \tag{5.1}$$

We will use the above result later in this proof. On a different note, since for every non-decreasing positive computable function $\Lambda : \mathbb{N} \to \mathbb{N}$, we have that for every large enough n, that $\ln Y(k) \leqslant \frac{\log n}{\Lambda(k)}$, from Lemma 2.1, we have with high probability (i.e., with probability $1 - \frac{1}{n^{\Omega(c)}}$) that:

$$S_1, \dots, S_{Y(k)} \text{ is a } (c\log n, Y(k), m, \lceil 60ck\rceil, 0.05)\text{-disperser.}$$

Thus, we have that:

$$\left|\widetilde{S}\right| \geqslant 0.95c\log n \text{ where } \widetilde{S} := \bigcup_{i\in[t]} S_i.$$

Consider the map $\rho : \widetilde{S} \to [t]$, where $\rho(s) = i$ where i is the smallest integer such that $s \in S_i$.

From the column-column consistency check, we have that for all $i \in [t]$, for all $s \in S_i$, and for all $r \in [k]$, we have $\vec{\alpha}^i_r(\pi_i(s)) = \vec{\alpha}^{\rho(s)}_r(\pi_{\rho(s)}(s))$.

We now construct k points, namely $\vec{z}_1, \dots, \vec{z}_k$ in $\mathbb{F}_2^{c\log n}$ as follows.

$$\forall r \in [k],\ \forall s \in [c\log n],\ \vec{z}_r(s) := \begin{cases} \vec{\alpha}^{\rho(s)}_r(\pi_{\rho(s)}(s)) & \text{if } s \in \widetilde{S} \\ 0 & \text{otherwise} \end{cases},$$

From the row-column inter-consistency checks and the row-column intra-consistency checks, we further have that for all $i \in [t]$:

$$\left(\sum_{r\in[k]} \vec{\beta}^i(r) \cdot \vec{z}_r\right)\Big|_{\widetilde{S}} = \vec{x}^i|_{\widetilde{S}}.$$

Now, we note that $\vec{\beta}^{i_1}, \dots, \vec{\beta}^{i_\gamma}$ are γ vectors in $\mathbb{F}_2^k$, and since $\gamma > k$, we have that these γ vectors are linearly dependent. Thus, there exists some $\vec{\lambda}^* \in \mathbb{F}_2^\gamma \setminus \{\vec{0}\}$ such that:

$$\sum_{w\in[\gamma]} \vec{\lambda}^*(w) \cdot \vec{\beta}^{i_w} = \vec{0}. \tag{5.2}$$

Now consider the vectors $\vec{x}^{i_1}, \dots, \vec{x}^{i_\gamma}$ restricted to the coordinates of $\widetilde{S}$. We have:

$$\begin{aligned}\left(\sum_{w\in[\gamma]} \vec{\lambda}^*(w) \cdot \vec{x}^{i_w}\right)\Big|_{\widetilde{S}} &= \sum_{w\in[\gamma]} \vec{\lambda}^*(w) \cdot \left(\vec{x}^{i_w}|_{\widetilde{S}}\right) \\ &= \sum_{w\in[\gamma]} \vec{\lambda}^*(w) \cdot \left(\sum_{r\in[k]} \vec{\beta}^{i_w}(r) \cdot \vec{z}_r\right)\Big|_{\widetilde{S}} \\ &= \sum_{r\in[k]} \left(\vec{z}_r\Big|_{\widetilde{S}} \cdot \sum_{w\in[\gamma]} \left(\vec{\lambda}^*(w) \cdot \vec{\beta}^{i_w}(r)\right)\right) = \vec{0},\end{aligned}$$

where the last equality follows from (5.2).

However, $\left(\sum_{w\in[\gamma]} \vec{\lambda}^*(w) \cdot \vec{x}^{i_w}\right)\Big|_{\widetilde{S}} = \vec{0}$ contradicts (5.1). Thus, there is no clique of size greater than $N(k)$ in G.

6 Inapproximability of Parameterized 2-CSP Problem

In this subsection, we prove the strong inapproximability of 2-CSP problem.

THEOREM 6.1. *There is a randomized algorithm running in FPT time which takes as input an instance $(X, c\log|X|, k)$ of (Y,N)-MSP and outputs an instance Φ of 2-CSP problem on $Y(k)$ variables and $O_k(|X|^2)$ sized alphabet set such that the following holds:*

Completeness: *If there exists a set $T \subseteq \mathbb{F}_2^d$ of k vectors such that*

$$|X \cap \mathsf{span}(T)| \geqslant Y(k),$$

then, with probability 1, Φ has a satisfying assignment.

Soundness: *If for every set $T \subseteq \mathbb{F}_2^d$ of k vectors we have*

$$|X \cap \mathsf{span}(T)| \leqslant N(k),$$

then, if $Y(k) > (8k + 3\cdot N(k))^2$ and $N(k) > 8k$, with probability $1 - \frac{1}{|X|^{\Omega(1)}}$, we have that every assignment to Φ satisfies at most $\frac{3\cdot N(k)}{\sqrt{Y(k)}}$ fraction of the constraints.

Assuming the above theorem, we have the proof of Theorem 1.3.

Proof. [Proof of Theorem 1.3]

As promised by MSH, let there be computable functions $Y, N : \mathbb{N} \to \mathbb{N}$, polynomial function $D : \mathbb{N} \to \mathbb{N}$, and constants $\delta, \rho > 0$ and $\eta, \zeta \in \mathbb{N}$ such that for all $i \in \mathbb{N}$ we have $Y(i) \geqslant \rho \cdot 2^{\delta \cdot i}$, $N(i) \leqslant \eta \cdot i^{\zeta}$, and (Y,N)-MSP being W[1]-hard even for instances (X, d, k) where $d = D(k)\cdot \log|X|$. We may assume that $Y(k) > (8k + 3\cdot N(k))^2$ and $N(k) > 8k$ (which is true for every large enough k).

We then apply the FPT reduction to instances of 2-CSP problem as given in Theorem 6.1 and obtain that it is W[1]-hard to distinguish instances of 2-CSP problem on $Y(k)$ variables where there is a satisfying assignment versus instances of 2-CSP problem where any assignment satisfies at most $3/k$ fraction of the constraints. Thus, if the parameter is set to $k' := Y(k) \geqslant \rho \cdot 2^{\delta k}$, then distinguishing the case where there is a satisfying assignment versus the case where any assignment satisfies at most $\frac{3\cdot N(k)}{\sqrt{Y(k)}} \leqslant \frac{3\cdot\eta\cdot(\log\frac{k'}{\rho})^{\zeta}}{\delta^{\zeta}\sqrt{k'}} = \frac{(\log k')^{O(1)}}{\sqrt{k'}}$ fraction of the constraints, is W[1]-hard. ∎

The rest of this section is dedicated to proving Theorem 6.1 where the reduction from (Y,N)-MSP to 2-CSP is the one described in Section 4.3.

6.1 Completeness Analysis

Suppose there exists a set $T := \{\vec{z}_1, \ldots, \vec{z}_k\} \subseteq \mathbb{F}_2^{c\log n}$ of k vectors such that there exists $\widetilde{X} \subseteq X \cap \mathsf{span}(T)$, such that $|\widetilde{X}| = Y(k)$. Then we define 2^k vectors in $\mathbb{F}_2^{c\log n}$ as follows:

$$\forall \vec{\beta} \in \mathbb{F}_2^k,\ \vec{\psi}_{\vec{\beta}} := \sum_{\ell\in[k]} \vec{\beta}(\ell)\cdot \vec{z}_\ell.$$

Let $\pi : \widetilde{X} \to \mathbb{F}_2^k$ be the mapping defined as follows: $\forall \vec{x} \in \widetilde{X}$, we set $\pi(\vec{x}) := \vec{\beta}$ where we have $\vec{\psi}_{\vec{\beta}} = \vec{x}$. Moreover, let $\widetilde{X} := \{\vec{x}^1, \ldots, \vec{x}^{Y(k)}\}$.

For every $i \in [Y(k)]$, we pick the label $(\vec{\alpha}^i_1, \ldots, \vec{\alpha}^i_k, \vec{\beta}^i, \vec{x}^i)$ for the variable v_i defined as follows:

$$\forall r \in [k],\ \vec{\alpha}^i_r := \vec{z}_r|_{S_i} \text{ and } \vec{\beta}^i := \pi(\vec{x}^i).$$

To finish the analysis in the completeness case, we need to show that every constraint is satisfied. To do so, fix some arbitrary $i, j \in [Y(k)]$, such that $i \neq j$. We can show as in Section 5.1 that the assigned pair of labels for variables v_i and v_j satisfy $C_{i,j}$. Details are omitted.

6.2 Soundness Analysis Let $\gamma := \sqrt{Y(k)}$. Let $Y(i) > (8 \cdot i + 3 \cdot N(i))^2$ and $N(i) > 8i$ for all $i \in \mathbb{N}$. Let $\delta := \frac{3 \cdot N(k)}{\sqrt{Y(k)}}$. To do the soundness analysis, suppose there exists an assignment $\sigma : V \to \Sigma$ that satisfies at least δ fraction of the constraints. Let $\sigma(v_i) := (\vec{\alpha}_1^i, \ldots, \vec{\alpha}_k^i, \vec{\beta}^i, \vec{x}^i)$ for all $i \in [Y(k)]$.

Let $V^* \subseteq V$ contain all $v \in V^*$, such that at least $\frac{2\delta}{3}$ fraction of constraints incident of v are satisfied under σ. By an averaging argument, we have that $|V^*| \geqslant \frac{\delta}{3} \cdot Y(k) = N(k) \cdot \sqrt{Y(k)}$. Pick some arbitrary $\widetilde{V} \subseteq V^*$ of size γ over $\mathbb{F}_2$. We apply the Deceitful Subspaces Lemma (Lemma 4.3) with $\kappa = 8\gamma k$ and $q = 2$ for the following $R \subseteq X$:

$$R := \{\vec{x}^i \mid v_i \in \widetilde{V}\}.$$

For every variable $v_i \in V$, let $Z_i \subseteq V$ be the set of variables each of whose constraint with v_i is satisfied under σ. Fix some $v_i \in \widetilde{V}$. From the definition of V^*, and since $\widetilde{V} \subseteq V^*$, we have that $|Z_i| \geqslant Y(k) \cdot 2\delta/3$. Let $\mathcal{S}_R \subseteq \binom{[c \log n]}{m}$ be the collection of sets of coordinates each of which is bad w.r.t. R. From the Deceitful Subspaces Lemma, with probability at least $1 - \frac{2^{(64k^2+1) \cdot Y(k)}}{n^{\frac{\sqrt{Y(k)}}{7}}}$, we have that $|\mathcal{S}_R| \leqslant \kappa$ as $|R| \leqslant |\widetilde{V}| = \gamma$. Note that we don't have to be worried if $|R|$ is smaller than γ, as we would then apply the Deceitful Subspaces Lemma with the same value of κ but a smaller value of γ. Alternatively, we could apply the transformation in Lemma 4.2 before constructing Φ in Section 4.3 to avoid duplicates in R.

We remove $v_j \in Z_i$ if $S_j \in \mathcal{S}_R$. Let the resulting subset of Z_i be denoted $\widetilde{Z}_i$.

CLAIM 6.1. *For every $v_j \in V$, if $S_j \notin \mathcal{S}_R$ then $|\widetilde{Z}_j \cap \widetilde{V}| \leqslant N(k)$.*

We now upper and lower bound the sum of number of satisfied constraints over each variable in $\widetilde{V}$. For the upper bound we have:

$$\begin{aligned}
\sum_{v_i \in \widetilde{V}} |Z_i| &\leqslant |\widetilde{V}| \cdot |Z_i \setminus \widetilde{Z}_i| + \sum_{v_i \in \widetilde{V}} |\widetilde{Z}_i| \\
&\leqslant \gamma\kappa + \sum_{v_i \in \widetilde{V}} |\widetilde{Z}_i| \\
&= 8k \cdot Y(k) + \sum_{\substack{v_j \in V \\ S_j \notin \mathcal{S}_R}} |Z_j \cap \widetilde{V}| \\
&\leqslant 8k \cdot Y(k) + Y(k) \cdot N(k)
\end{aligned}$$

Next, we have the lower bound:

$$\sum_{v_i \in \widetilde{V}} |Z_i| \geqslant \frac{2\delta \cdot Y(k)}{3} \cdot |\widetilde{V}| = 2 \cdot Y(k) \cdot N(k)$$

This leads to a contradiction as $N(k) > 8k$.

Proof. [Proof of Claim 6.1] Suppose the contrary, and assume there is some $v_j \in V$ such that $S_j \notin \mathcal{S}_R$ and $|\widetilde{Z}_j \cap \widetilde{V}| > N(k)$. Then for every $v_i \in \widetilde{Z}_j \cap \widetilde{V}$ we have that:

$$\sum_{r=1}^{k} \vec{\beta}^i(r) \cdot \vec{\alpha}_r^j = \left(\vec{x}^i(\pi_j^{-1}(1)), \vec{x}^i(\pi_j^{-1}(2)), \ldots, \vec{x}^i\left(\pi_j^{-1}(m)\right)\right).$$

Let $\widetilde{Z}_j \cap \widetilde{V} = \{v_{i_1}, \ldots, v_{i_w}\}$ where $w > N(k)$. Let A be the matrix whose columns are the vectors $\vec{x}^{i_1}, \ldots, \vec{x}^{i_w}$ (all of which are in X). From the soundness case assumption, we know that $\mathsf{rank}(A) \geqslant k+1$. Thus, there are $k+1$ linearly independent columns in A, say, $\vec{x}^{i'_1}, \ldots, \vec{x}^{i'_{k+1}}$. However, we know that $\vec{\beta}^{i'_1}, \ldots, \vec{\beta}^{i'_{k+1}}$ are linearly dependent (as these are $k+1$ vectors in $\mathbb{F}_2^k$). Thus, there exists $\vec{\lambda} \in \mathbb{F}_2^{k+1}$ such that:

$$\sum_{h=1}^{k+1} \vec{\lambda}(h) \cdot \vec{\beta}^{i'_h} = \vec{0}.$$

$$\forall r \in [k],\ \sum_{t=h}^{k+1} \vec{\lambda}(h) \cdot \vec{\beta}^{i'_h}(r) = 0.$$

For all $t \in S_j$ we have:

$$\sum_{h=1}^{k+1} \vec{\lambda}(h) \cdot \vec{x}^{i'_h}(t) = \sum_{h=1}^{k+1} \left(\vec{\lambda}(h) \cdot \left(\sum_{r=1}^{k} \vec{\beta}^{i'_h}(r) \cdot \vec{\alpha}_r^j \right) \right)$$
$$= \sum_{r=1}^{k} \left(\left(\sum_{h=1}^{k+1} \vec{\lambda}(h) \cdot \vec{\beta}^{i'_h}(r) \right) \cdot \vec{\alpha}_r^j \right) = 0.$$

Since the above is true for all $t \in [m]$, we have that:

$$\sum_{h=1}^{k+1} \vec{\lambda}(h) \cdot \left(\vec{x}^{i'_h} \right) |_{S_j} = \vec{0}.$$

However, this contradicts $S_j \notin \mathcal{S}_R$ as $\vec{x}^{i'_1}, \ldots, \vec{x}^{i'_{k+1}}$ are linearly independent. □

7 Inapproximability of k-Biclique and Densest k-Subgraph

In this subsection, we prove the strong inapproximability of k-Biclique problem and the Densest k-Subgraph problem. However, to do so we need to introduce a strengthening of MSH which we do below.

7.1 Strong Maximum Span Hypothesis In this subsection, we formally define (the colored version of) the k-Maximum Totally Span Problem and the Strong Maximum Span Hypothesis.

DEFINITION 7.1. ((Y, N)-MAXIMUM TOTALLY SPAN PROBLEM ((Y, N)-MTSP)) *Given two computable functions $Y, N : \mathbb{N} \to \mathbb{N}$, an instance of (Y, N)-MTSP is specified by the tuple $(q, X := X_1 \dot\cup X_2 \dot\cup \cdots \dot\cup X_{Y(k)}, d, k, \omega)$ where $q \in \mathbb{P}$, for all $i \in [Y(k)]$, we have $X_i \subseteq \mathbb{F}_q^d$ (for all $i, i' \in [Y(k)]$, we have $|X_i| = |X_{i'}|$), $\omega := \omega(k) \in (0, 1]$ is a constant, k is the parameter, and the goal is to distinguish between the two cases:*

Completeness: *There exists a subset $\widetilde{X} \subseteq X$ such that for all $i \in [Y(k)]$ we have $|\widetilde{X} \cap X_i| = 1$ and for every subset $\widetilde{X}_0 \subseteq \widetilde{X}$, if $|\widetilde{X}_0| \geqslant \omega \cdot Y(k)$ then the rank of the space spanned by $\widetilde{X}_0$ is exactly k.*

Soundness: *For every set $T \subseteq \mathbb{F}_q^d$ of k vectors we have*

$$\sum_{i \in [Y(k)]} |X_i \cap \mathsf{span}(T)| \leqslant N(k).$$

REMARK 7.1. *Every instance (X, d, k) of (Y, N)-MSP is also an instance $(2, X, d, k, 1)$ of (Y, N)-MTSP (after applying Lemma 4.2) with the additional soundness guarantee that in the completeness case, the $Y(k)$ input points span a k-dimensional space (instead of the weaker property in (Y, N)-MSP where they are only contained in a k-dimensional space).*

HYPOTHESIS 7.1. (STRONG MAXIMUM SPAN HYPOTHESIS (Strong MSH)) *There exist computable functions $Y, N : \mathbb{N} \to \mathbb{N}$, polynomial function $D : \mathbb{N} \to \mathbb{N}$, and constants $\delta, \rho > 0$, $\eta, \zeta \in \mathbb{N}$, and $\xi \geqslant 16$, such that for all $i \in \mathbb{N}$ we have $Y(i) \geqslant \rho \cdot q^{\delta \cdot i}$ and $N(i) \leqslant \eta \cdot i^{\zeta}$. Then (Y, N)-MTSP is W[1]-hard even for instances (q, X, d, k, ω) where $d = D(k) \cdot \log |X|$ and $\omega = \frac{1}{\xi k \log q}$.*

7.2 Construction of Hard Graph Instances In this subsection, we describe how for each instance of (Y, N)-MTSP we construct a bipartite graph that will serve as both the hard instance of the k-Biclique problem and the Densest k-Subgraph problem.

For some constant $c_0 \in \mathbb{Z}$, starting from an instance of $(q, X_0, c_0 \log n, k, \omega)$ of (Y, N)-MTSP (where $n := |X_0|$), we apply the algorithm[9] in Lemma 4.1 to obtain an instance $(q, X, c \log n, k, \omega)$ of (Y, N)-MTSP, where $c := c_0 \cdot \ell$, for some constant ℓ. Moreover, we have that every pair of vectors in X are at relative Hamming distance greater than or equal to 0.1. Recall that $X := X_1 \dot\cup X_2 \dot\cup \cdots \dot\cup X_{Y(k)}$, where $|X_i| = n/Y(k)$.

Let $m := \frac{\log n}{k \log q}$. Let $\Gamma \subseteq \mathbb{F}_q^{\frac{\log n}{\log q}}$ be a subset that is constructed as follows. We identify $\mathbb{F}_q^{\frac{\log n}{\log q}}$ with the set of all $\mathbb{F}_q^{m \times k}$ matrices. We go over each $\mathbb{F}_q^{m \times k}$ matrix A and include it in Γ if both of the following conditions hold:

- A is full rank (i.e., $\mathsf{rank}(A) = k$).
- There is no matrix A' already in Γ such that the span of the columns of A and A' are the same, i.e., $\mathsf{span}(A) = \mathsf{span}(A')$.

We can check the above two conditions for each matrix A in $\widetilde{O}(q^k n + k \cdot m^2) = \widetilde{O}_k(n)$ time. Thus we can construct the set Γ in $\widetilde{O}_k(n^2)$ time.

From such an instance $(q, X, c \log n, k, \omega)$ of (Y, N)-MTSP, we construct at most $\binom{q^k - 1}{Y(k)}$ many bipartite graphs. In particular, for every collection of vectors $B = \{\vec{\beta}^1, \ldots, \vec{\beta}^{Y(k)}\} \subseteq \mathbb{F}_q^k \setminus \{\vec{0}\}$ of size $Y(k)$, if B is full rank, we construct bipartite graph G_B on partite sets R and C (i.e., $V(G_B) := R \dot\cup C$ and $E(G_B) \subseteq R \times C$). The size of R is $Y(k) \cdot n$ and the size of C is $Y(k) \cdot |\Gamma|$. Both R and C are equipartitioned into $Y(k)$ parts, say $R_1, \ldots, R_{Y(k)}$ and $C_1, \ldots, C_{Y(k)}$ respectively. For all $i \in [Y(k)]$, we think of the vertices in R_i as a copy of X_i and the vertices in C_i as a copy of Γ (we view each matrix in Γ through it's k linearly independent column vectors).

Let $S_1, \ldots, S_{Y(k)}$ be random subsets of $[c \log n]$ of size m. Let $\pi_i : S_i \to [m]$ be some canonical 1-to-1 mapping.

We are now ready to define the edges. Fix $i, j \in [Y(k)]$, $\vec{\alpha}_1^j, \ldots, \vec{\alpha}_k^j \in \mathbb{F}_q^m$ and $\vec{x}^i \in X_i$. We define an edge between $\vec{x}^i$ in R_i and $(\vec{\alpha}_1^j, \ldots, \vec{\alpha}_k^j)$ in C_j if and only if the following holds:

$$\sum_{r=1}^{k} \vec{\beta}^i(r) \cdot \vec{\alpha}_r^j = \left(\vec{x}^i(\pi_j^{-1}(1)), \vec{x}^i(\pi_j^{-1}(2)), \ldots, \vec{x}^i\left(\pi_j^{-1}(m)\right)\right).$$

7.3 W[1]-Hardness of Approximating k-Biclique and Densest-k-Subgraph

In this subsection, we first prove the strong inapproximability of the k-Biclique problem.

THEOREM 7.1. *There is a randomized algorithm running in* FPT *time which takes as input an instance $(q, X, c \log |X|, k, \omega)$ of (Y, N)-MTSP and outputs at most $\binom{q^k - 1}{Y(k)}$ many instances $\{G_B\}_{\substack{B \subseteq \mathbb{F}_q^k \setminus \{\vec{0}\} \\ |B| = Y(k) - 1 \\ \mathsf{rank}(B) = k}}$ (of order $O_k(|X|)$) of Y^*-Biclique problem, where $Y^* = \frac{Y(k)}{16k \log q}$, such that the following holds:*

Completeness: *If there exists a subset $\widetilde{X} \subseteq X$ such that for all $i \in [Y(k)]$ we have $|\widetilde{X} \cap X_i| = 1$ and for every subset $\widetilde{X}_0 \subseteq \widetilde{X}$, if $|\widetilde{X}_0| \geqslant \omega \cdot Y(k)$ then the rank of the space spanned by $\widetilde{X}_0$ is exactly k, then, if $Y(k) \geqslant 16k^2 \log q$, there is some $B \subseteq \mathbb{F}_q^k \setminus \{\vec{0}\}$ of size $Y(k)$ such that with probability $1 - \frac{1}{|X|^{\Omega(1)}}$, G_B has a biclique K_{Y^*, Y^*} as a subgraph.*

Soundness: *If for every set $T \subseteq \mathbb{F}_q^d$ of k vectors we have*

$$\sum_{i \in [Y(k)]} |X_i \cap \mathsf{span}(T)| \leqslant N(k).$$

then, if $N(k) \geqslant 60ck$, we have that for all $B \subseteq \mathbb{F}_q^k \setminus \{\vec{0}\}$ of size $Y(k)$, with probability $1 - \frac{1}{|X|^{\Omega(1)}}$, G_B does not have a biclique of size $N(k)$.

Assuming the above theorem, we have the proof of Theorem 1.4.

---

[9]Lemma 4.1 was only proved for vectors over $\mathbb{F}_2$ but it can be easily seen that it extends to all prime order finite fields as Fact 2.1 holds for all prime order fields.

Proof. [Proof of Theorem 1.4] As promised by Strong MSH, let there be computable functions $Y, N : \mathbb{N} \to \mathbb{N}$, polynomial function $D : \mathbb{N} \to \mathbb{N}$, and constants $\delta, \rho > 0$, $\eta, \zeta \in \mathbb{N}$, and $\xi \geqslant 16$, such that for all $i \in \mathbb{N}$ we have $Y(i) \geqslant \rho \cdot q^{\delta \cdot i}$, $N(i) \leqslant \eta \cdot i^{\zeta}$, and (Y, N)-MTSP being W[1]-hard even for instances (q, X, d, k, ω) where $d = D(k) \cdot \log |X|$ and $\omega = \frac{1}{\xi k \log q}$. We may assume that $N(k) \geqslant 60D(k) \cdot k$, otherwise we revise the constants η and ζ accordingly.

We then apply the FPT reduction to instances of $Y(k)$-Biclique problem as given in Theorem 7.1 and obtain that it is W[1]-hard to distinguish instances of $Y(k)$-Biclique problem where there is a $\frac{\rho \cdot q^{\delta k}}{16k \log q}$ sized biclique versus instances of $Y(k)$-Biclique problem where there is no $1 + \eta \cdot k^{\zeta}$ sized biclique. Thus, if the parameter is set to $k' := \frac{\rho \cdot q^{\delta k}}{16k \log q}$, then distinguishing the case where there is a k' sized biclique versus the case there is no biclique of size $1 + \eta \cdot k^{\zeta} = (\log k')^{O(1)}$, is W[1]-hard. ◻

Next, we first prove the near optimal inapproximability of Densest k-Subgraph problem.

THEOREM 7.2. *There is a randomized algorithm running in* FPT *time which takes as input an instance* $(q, X, c \log |X|, k, \omega)$ *of* (Y, N)-MTSP *and outputs at most* $\binom{q^k - 1}{Y(k)}$ *many instances* $\{G_B\}_{\substack{B \subseteq \mathbb{F}_q^k \setminus \{\vec{0}\} \\ |B| = Y(k) - 1 \\ \mathsf{rank}(B) = k}}$ *(of order* $O_k(|X|)$*) of* Densest $Y(k)$-Subgraph *problem such that the following holds:*

Completeness: *If there exists a subset* $\widetilde{X} \subseteq X$ *such that for all* $i \in [Y(k)]$ *we have* $|\widetilde{X} \cap X_i| = 1$ *and for every subset* $\widetilde{X}_0 \subseteq \widetilde{X}$*, if* $|\widetilde{X}_0| \geqslant \omega \cdot Y(k)$ *then the rank of the space spanned by* $\widetilde{X}_0$ *is exactly* k*, then, if* $Y(k) \geqslant 16k^2 \log q$*, there is some* $B \subseteq \mathbb{F}_q^k \setminus \{\vec{0}\}$ *of size* $Y(k)$ *such that with probability* $1 - \frac{1}{|X|^{\Omega(1)}}$*,* G_B *has a biclique* K_{Y^*, Y^*} *as a subgraph, where* $Y^* = \frac{Y(k)}{16k \log q}$.

Soundness: *If for every set* $T \subseteq \mathbb{F}_q^d$ *of* k *vectors we have*

$$\sum_{i \in [Y(k)]} |X_i \cap \mathsf{span}(T)| \leqslant N(k).$$

then, if $Y(k) > (10k + 3 \cdot N(k))^2$ *and* $N(k) \geqslant 60ck$*, we have that for all* $B \subseteq \mathbb{F}_q^k \setminus \{\vec{0}\}$ *of size* $Y(k)$*, with probability* $1 - \frac{1}{|X|^{\Omega(1)}}$*,* G_B *does not have a subgraph on* $2Y(k)$ *vertices with more than* $(2Y(k))^{2 - \frac{1}{k}}$ *edges.*

Assuming the above theorem, we have the proof of Theorem 1.5.

Proof. [Proof of Theorem 1.5] As promised by Strong MSH, let there be computable functions $Y, N : \mathbb{N} \to \mathbb{N}$, polynomial function $D : \mathbb{N} \to \mathbb{N}$, and constants $\delta, \rho > 0$, $\eta, \zeta \in \mathbb{N}$, and $\xi \geqslant 16$, such that for all $i \in \mathbb{N}$ we have $Y(i) \geqslant \rho \cdot q^{\delta \cdot i}$, $N(i) \leqslant \eta \cdot i^{\zeta}$, and (Y, N)-MTSP being W[1]-hard even for instances (q, X, d, k, ω) where $d = D(k) \cdot \log |X|$ and $\omega = \frac{1}{\xi k \log q}$. We may assume that $N(k) \geqslant 60D(k) \cdot k$, otherwise we revise the constants η and ζ accordingly. We may further assume that $Y(k) > (10k + 3 \cdot N(k))^2$ (which is true for every large enough k).

We then apply the FPT reduction to instances of Densest $Y(k)$-Subgraph problem as given in Theorem 7.2 and obtain that it is W[1]-hard to distinguish instances of Densest $Y(k)$-Subgraph problem where there is a $\frac{\rho \cdot q^{\delta k}}{16k \log q}$ sized biclique versus instances of Densest $Y(k)$-Subgraph problem where there is no subgraph on $2Y(k)$ vertices with more than $\left(\frac{\rho \cdot q^{\delta k}}{16k \log q}\right)^{2 - \frac{1}{k}}$ edges. Thus, if the parameter is set to $k' := \frac{\rho \cdot q^{\delta k}}{16k \log q}$, then distinguishing the case where there is a k' sized biclique versus the case there is no subgraph on $2k'$ vertices with at most $(k')^{2 - \frac{1}{k}} = (k')^{2 - \frac{1}{\mu \cdot \log k'}} = \frac{(k')^2}{2^{1/\mu}}$ edges (for some $\mu > 0$), is W[1]-hard. ◻

The rest of this section is dedicated to proving Theorems 7.1 and 7.2.

7.4 Completeness Analysis of k-Biclique and Densest-k-Subgraph Problems Let $Y(i) \geqslant 16i^2 \log q$, for all $i \in \mathbb{N}$. Suppose there exists a set $T := \{\vec{z}_1, \ldots, \vec{z}_k\} \subseteq \mathbb{F}_q^{c \log n}$ of k linearly independent vectors such that there exists $\widetilde{X} \subseteq X \cap \mathsf{span}(T)$, such that $|\widetilde{X}| = Y(k)$. We can even assume that $T \subseteq \widetilde{X}$ by a change of basis (as the rank of the space spanned by vectors in $\widetilde{X}$ is equal to k). Then we define q^k vectors in $\mathbb{F}_q^{c \log n}$ as follows:

$$\forall \vec{\beta} \in \mathbb{F}_q^k,\ \vec{\psi}_{\vec{\beta}} := \sum_{\ell \in [k]} \vec{\beta}(\ell) \cdot \vec{z}_\ell.$$

Let $\pi : \widetilde{X} \to \mathbb{F}_q^k$ be the mapping defined as follows: $\forall \vec{x} \in \widetilde{X}$, we set $\pi(\vec{x}) := \vec{\beta}$ where we have $\vec{\psi}_{\vec{\beta}} = \vec{x}$. Moreover, let $\widetilde{X} := \{\vec{x}^1, \dots, \vec{x}^{Y(k)}\}$. Fix $\widetilde{X}'$ to be some subset of $\widetilde{X}$ of size $Y^* := \frac{Y(k)}{16k \log q}$ containing T.

We apply the Deceitful Subspaces Lemma (Lemma 4.3) to $\widetilde{X}'$ of size $\gamma := Y^*$ (i.e., $|\widetilde{X}'| = \gamma$) over $\mathbb{F}_q$ with $\kappa := Y(k)/2$. Then, we know that $\mathcal{S}_{\widetilde{X}'} \subseteq \{S_1, \dots, S_{Y(k)}\}$ in the lemma statement is at most of size $Y(k)/2$ with probability at least $1 - \frac{1}{n^{\Omega_k(1)}}$. Let $\mathcal{T}_{\widetilde{X}'}$ be an arbitrary subset of $\{S_1, \dots, S_{Y(k)}\} \setminus \mathcal{S}_{\widetilde{X}'}$ of size $|\widetilde{X}'| = Y^*$.

We know that for every $S \in \mathcal{T}_{\widetilde{X}'}$, we have that for all $\vec{\lambda} \in \mathbb{F}_q^\gamma$ the following holds:

$$\sum_{r \in [\gamma]} \vec{\lambda}(r) \cdot \vec{x}^{i_r} \neq \vec{0} \implies \sum_{r \in [\gamma]} \vec{\lambda}(r) \cdot \left(\vec{x}^{i_r}\big|_S\right) \neq \vec{0}. \tag{7.3}$$

For every $S_j \in \mathcal{T}_{\widetilde{X}'}$, we pick the vertex $c_j^* \in C_j$ (the choice of the graph will be soon specified) defined as follows. First we define k vectors in $\mathbb{F}_q^{c \log n}$ restricted to the coordinates in S_j.

$$\forall r \in [k],\ \vec{a}_r^j := \vec{z}_r|_{S_j}$$

We know that $(\vec{a}_1^j, \dots, \vec{a}_k^j)$ is a k-tuple of vectors that are linearly independent because $T \subseteq \widetilde{X}'$ is a collection of k linearly independent vectors and then applying (7.3), since we have $S_j \in \mathcal{T}_{\widetilde{X}'}$.

Thus there exists some unique $(\vec{\alpha}_1^j, \dots, \vec{\alpha}_k^j)$ in Γ such that $\mathsf{span}(\{\vec{\alpha}_1^j, \dots, \vec{\alpha}_k^j\}) = \mathsf{span}(\{\vec{a}_1^j, \dots, \vec{a}_k^j\})$. We define $c_j^* := (\vec{\alpha}_1^j, \dots, \vec{\alpha}_k^j)$.

Let A_T be the full rank $\mathbb{F}_q^{k \times k}$ matrix such that:

$$A_T \cdot \begin{bmatrix} \vec{\alpha}_1^j \\ \vdots \\ \vec{\alpha}_k^j \end{bmatrix} = \begin{bmatrix} \vec{a}_1^j \\ \vdots \\ \vec{a}_k^j \end{bmatrix}.$$

For every $\vec{x}^i \in \widetilde{X}'$ let:

$$\vec{\beta}^i := \pi(\vec{x}^i) \cdot A_T$$

Let $B := \{\vec{\beta}^1, \dots, \vec{\beta}^{Y(k)}\}$. We are now interested in the graph G_B. We pick the vertex $r_i^* := \vec{x}^i$ in R_i of G_B.

To finish the analysis in the completeness case, we need to show that for all $\vec{x}^i \in \widetilde{X}'$ and $S_j \in \mathcal{T}_{\widetilde{X}'}$, we have an edge between r_i^* and c_j^* in G_B. Fix some arbitrary $\vec{x}^i \in \widetilde{X}'$ and $S_j \in \mathcal{T}_{\widetilde{X}'}$ and some $t \in [m]$. We want to show that:

$$\sum_{r=1}^{k} \vec{\beta}^i(r) \cdot \vec{\alpha}_r^j(t) = \vec{x}^i(\pi_j^{-1}(t)).$$

This follows from the below computation.

$$\sum_{r=1}^{k} \vec{\beta}^i(r) \cdot \vec{\alpha}_r^j(t) = \sum_{r=1}^{k} \vec{\beta}^i(r) \cdot A_T^{-1} \cdot \vec{z}_r(\pi_j^{-1}(t)) = \sum_{r=1}^{k} \pi(\vec{x}^i)_r \cdot A_T \cdot A_T^{-1} \cdot \vec{z}_r(\pi_j^{-1}(t)) = \vec{x}^i(\pi_j^{-1}(t)).$$

Since the choice of t was arbitrary, it holds for all $t \in [m]$. Thus, G_B has the complete bipartite graph K_{Y^*,Y^*} as an induced subgraph which also means that there are $2 \cdot Y^*$ vertices in G_B with $(Y^*)^2$ edges contained inside those vertices.

7.5 Soundness Analysis of k-Biclique Problem Fix some $B := \{\vec{\beta}^1, \ldots, \vec{\beta}^{Y(k)}\} \subseteq \mathbb{F}_q^k$. Let t be any integer greater than $N(k) \geqslant 60ck$. To do the soundness analysis, suppose there is a $K_{t,t}$ as an induced subgraph in G_B. Without loss of generality, let the vertices in this subgraph be $\{r_1, \ldots, r_t\} \subseteq R$ and $\{c_1, \ldots, c_t\} \subseteq C$, where for all $i \in [t]$, we have $r_i := \vec{x}^i$ and $c_i := (\vec{\alpha}_1^i, \ldots, \vec{\alpha}_k^i)$.

Suppose there existed some $j, j' \in [t]$ and $j^* \in [Y(k)]$ such that $c_j \in C_{j^*}$ and $c_{j'} \in C_{j^*}$, then we have that for all $i \in [t]$ we have:

$$\sum_{r=1}^{k} \vec{\beta}^i(r) \cdot \vec{\alpha}_r^j = \left(\vec{x}^i(\pi_{j^*}^{-1}(1)), \vec{x}^i(\pi_{j^*}^{-1}(2)), \ldots, \vec{x}^i\left(\pi_{j^*}^{-1}(m)\right)\right) = \sum_{r=1}^{k} \vec{\beta}^i(r) \cdot \vec{\alpha}_r^{j'}$$

This leads to a contradiction, as $(\vec{\alpha}_1^j, \ldots, \vec{\alpha}_k^j) \neq (\vec{\alpha}_1^{j'}, \ldots, \vec{\alpha}_k^{j'})$, and at the same time both these sets of k vectors span a k-dimensional space (as they are members of Γ), and thus cannot be mapped to a null space from all $\vec{\beta}^i$s (as that would contradict that $\mathsf{rank}(B) = k$).

On a different note, since for every non-decreasing positive computable function $\Lambda : \mathbb{N} \to \mathbb{N}$, we have that for every large enough n, that $\ln Y(k) \leqslant \frac{\log n}{\Lambda(k)}$, from Lemma 2.1, we have with high probability (i.e., with probability $1 - \frac{1}{n^{\Omega(c)}}$) that:

$$S_1, \ldots, S_{Y(k)} \text{ is a } (c \log n, Y(k), m, \lceil 60ck \rceil, 0.05)\text{-disperser.}$$

Thus, we have that:

$$\left|\widetilde{S}\right| \geqslant 0.95c \log n \text{ where } \widetilde{S} := \bigcup_{i \in [t]} S_i. \tag{7.4}$$

We now remark that if suppose there existed some $i, i' \in [t]$ and $i^* \in [Y(k)]$ such that $\vec{x}^i \in X_{i^*}$ and $\vec{x}^{i'} \in X_{i^*}$, then we have that for all $j \in [t]$ we have:

$$\left(\vec{x}^i(\pi_j^{-1}(1)), \vec{x}^i(\pi_j^{-1}(2)), \ldots, \vec{x}^i\left(\pi_j^{-1}(m)\right)\right) = \sum_{r=1}^{k} \vec{\beta}^{i^*}(r) \cdot \vec{\alpha}_r^j = \left(\vec{x}^{i'}(\pi_j^{-1}(1)), \vec{x}^{i'}(\pi_j^{-1}(2)), \ldots, \vec{x}^{i'}\left(\pi_j^{-1}(m)\right)\right).$$

Thus, we have that $(\vec{x}^i)\,|_{\widetilde{S}} = (\vec{x}^{i'})\,|_{\widetilde{S}}$ which is not possible as $|\widetilde{S}| \geqslant 0.95c \log n$ and $\Delta(\vec{x}^i, \vec{x}^{i'}) \geqslant 0.1$.

Thus, from now on, without loss of generality, we may even assume that $\vec{x}^i \in X_i$ and $c_i \in C_i$.

Let $A_{X'} \subseteq \mathbb{F}_q^{c \log n \times t}$ be the matrix whose columns are the vectors in $X' := \{\vec{x}^1, \ldots, \vec{x}^t\}$. Let $\gamma := \mathsf{rank}(A_{X'}) \geqslant k+1$ (from the soundness assumption). Let $\vec{x}^{i_1}, \ldots, \vec{x}^{i_\gamma}$ be the γ linearly independent column vectors of $A_{X'}$. Thus, for every $\vec{\lambda} \in \mathbb{F}_q^\gamma \setminus \vec{0}$, we have:

$$\sum_{w \in [\gamma]} \vec{\lambda}(w) \cdot \vec{x}^{i_w} \neq \vec{0}.$$

Since $\vec{x}^{i_1}, \ldots, \vec{x}^{i_\gamma}$ are codewords of a linear code with relative distance greater than 0.1, we have that for every $\vec{\lambda} \in \mathbb{F}_q^w \setminus \vec{0}$, we have:

$$\Delta\left(\sum_{w \in [\gamma]} \vec{\lambda}(w) \cdot \vec{x}^{i_w}, \vec{0}\right) \geqslant 0.1.$$

This further implies that for any set $T \subseteq [c \log n]$ of coordinates, if $|T| > 0.9c \log n$ then,

$$\left(\sum_{w \in [\gamma]} \vec{\lambda}(w) \cdot \vec{x}^{i_w}\right)\Big|_T \neq \vec{0}. \tag{7.5}$$

Consider the map $\rho : \widetilde{S} \to [t]$, where $\rho(s) = i$ where i is the smallest integer such that $s \in S_i$. We now construct k points, namely $\vec{z}_1, \ldots, \vec{z}_k$ in $\mathbb{F}_q^{c \log n}$ as follows.

$$\forall r \in [k],\ \forall s \in [c \log n],\ \vec{z}_r(s) := \begin{cases} \vec{\alpha}_r^{\rho(s)}(\pi_{\rho(s)}(s)) & \text{if } s \in \widetilde{S} \\ 0 & \text{otherwise} \end{cases},$$

From the row-column inter-consistency checks and the row-column intra-consistency checks, we further have that for all $i \in [t]$:

$$\left(\sum_{r\in[k]} \vec{\beta}^i(r)\cdot \vec{z}_r\right)\Big|_{\widetilde{S}} = \vec{x}^i|_{\widetilde{S}}.$$

Now, we note that $\vec{\beta}^{i_1},\dots,\vec{\beta}^{i_\gamma}$ are γ vectors in $\mathbb{F}_q^k$, and since $\gamma > k$, we have that these γ vectors are linearly dependent. Thus, there exists some $\vec{\lambda}^* \in \mathbb{F}_q^\gamma \setminus \{\vec{0}\}$ such that:

$$\sum_{w\in[\gamma]} \vec{\lambda}^*(w)\cdot \vec{\beta}^{i_w} = \vec{0}. \tag{7.6}$$

Now consider the vectors $\vec{x}^{i_1},\dots,\vec{x}^{i_\gamma}$ restricted to the coordinates of $\widetilde{S}$. We have:

$$\begin{aligned}\left(\sum_{w\in[\gamma]} \vec{\lambda}^*(w)\cdot \vec{x}^{i_w}\right)\Big|_{\widetilde{S}} &= \sum_{w\in[\gamma]} \vec{\lambda}^*(w)\cdot \left(\vec{x}^{i_w}|_{\widetilde{S}}\right)\\ &= \sum_{w\in[\gamma]} \vec{\lambda}^*(w)\cdot \left(\sum_{r\in[k]} \vec{\beta}^{i_w}(r)\cdot \vec{z}_r\right)\Big|_{\widetilde{S}}\\ &= \sum_{r\in[k]}\left(\vec{z}_r\Big|_{\widetilde{S}}\cdot \sum_{w\in[\gamma]}\left(\vec{\lambda}^*(w)\cdot \vec{\beta}^{i_w}(r)\right)\right) = \vec{0},\end{aligned}$$

where the last equality follows from (7.6).

However, $\left(\sum_{w\in[\gamma]} \vec{\lambda}^*(w)\cdot \vec{x}^{i_w}\right)\Big|_{\widetilde{S}} = \vec{0}$ contradicts (7.5). Thus, there is no $K_{t,t}$ in G_B if $t > N(k)$.

7.6 Soundness Analysis of Densest k-Subgraph Problem Fix some $B := \{\vec{\beta}^1,\dots,\vec{\beta}^{Y(k)}\} \subseteq \mathbb{F}_q^k$. To do the soundness analysis, consider an induced subgraph H in G_B, let the vertices of H be $\widetilde{R} := \{r_1,\dots,r_{t_R}\} \subseteq R$ and $\widetilde{C} := \{c_1,\dots,c_{t_C}\} \subseteq C$, where for all $i \in [t]$, we have $r_i := \vec{x}^i$ and $c_i := (\vec{\alpha}^i_1,\dots,\vec{\alpha}^i_k)$, and $t_R + t_C = 2\cdot Y(k)$. We can assume that both $\min\{t_R, t_C\} > N(k)$, as otherwise the number of edges in H is less than $2\cdot Y(k)\cdot N(k)$.

For any $i^*, j^* \in [Y(k)]$, we first observe that there cannot be $K_{60ck,k}$ in H between the vertices in $(R_{i^*}\cap \widetilde{R})$ and $(C_{j^*}\cap \widetilde{C}))$. This is because of (7.4) and that $\mathsf{rank}(B) = k$ (much like the arguments we made in the previous subsection).

We now recall here a classic result connecting the sparsity of a graph with the presence of biclique.

Theorem 7.3. (Kovári-Sós-Turán [KTST54]) *For every positive integer N and $s,t \leqslant N$, every $K_{s,t}$-free graph on N vertices has at most $O(N^{2-\frac{1}{\min(s,t)}})$ edges.*

Thus, we have

$$|E(H)\cap((R_{i^*}\cap\widetilde{R})\times(C_{j^*}\cap\widetilde{C}))| \leqslant (|R_{i^*}\cap\widetilde{R}| + |C_{j^*}\cap\widetilde{C}|)^{2-\frac{1}{k}} \tag{7.7}$$

On the other hand, if we pick one representative for each R_{i^*} from $\widetilde{R}$ (if $R_{i^*}\cap\widetilde{R}\neq\emptyset$) and one representative for each C_{j^*} from $\widetilde{C}$ (if $C_{j^*}\cap\widetilde{C}\neq\emptyset$), then we obtain the scenario of the soundness analysis in Section 6.2 where the number of edges in this subgraph of H is at most $3\cdot N(k)\cdot(Y(k))^{1.5}$. These edges are negligible in number when compared to (7.7). Thus an upper bound on the total number of edges is given by (7.7) which $(Y(k))^{2-\frac{1}{k}}$.

8 Gap-ETH Based Validation for MSH

In this subsection, we prove that assuming $\mathsf{Gap\text{-}ETH}$, we can obtain a weak version of MSH.

Theorem 8.1. *There is some polynomial $Y : \mathbb{N} \to \mathbb{N}$, constants $C_1, C_2 > 0$, and a randomized algorithm which takes as input an instance of 3-SAT φ on n variables and an integer k and outputs an instance $(X := X_1\dot\cup\cdots\dot\cup X_{Y(k)}, n+k, k)$ of MSP such that the following holds:*

Runtime: *The algorithm runs in time* $2^{C_1 n/k}$.

Completeness: *If there exists a satisfying assignment to* φ *then there is a set* $T \subseteq \mathbb{F}_2^{n+k}$ *of* k *vectors such that for all* $i \in [Y(k)]$,

$$|X_i \cap \mathsf{span}(T)| \geqslant 1.$$

Soundness: *If every assignment to* φ *violates* δ *fraction of the clauses then with high probability, for every set* $T \subseteq \mathbb{F}_2^{n+k}$ *of* k *vectors we have*

$$|\{i \in [Y(k)] : |X_i \cap \mathsf{span}(T)| \geqslant 1\}| \leqslant (1 - C_2 \cdot \delta) \cdot Y(k).$$

Assuming the above theorem, we have the proof of Theorem 1.1.

Proof. [Proof of Theorem 1.1] From Gap-ETH, there exist constants $\varepsilon, \delta > 0$ such that any randomized algorithm that, on input a 3-SAT formula φ on n variables and $O(n)$ clauses, can distinguish between $\mathsf{val}(\varphi) = 1$ and $\mathsf{val}(\varphi) < 1 - \delta$, must run in time at least $2^{\varepsilon n}$. Let $N : \mathbb{N} \to \mathbb{N}$ be defined as $N(i) = (1 - C_2 \cdot \delta) \cdot Y(i)$ for all $i \in \mathbb{N}$.

Suppose there was an algorithm to solve (the colored variant of) (Y, N)-MSP in $F(k) \cdot |X|^\alpha$ time (for some computable function F and constant $\alpha > 1$) then we use the algorithm in Theorem 8.1 for a setting of $k \gg C_1\alpha/\varepsilon$.

Thus, we can convert φ to an instance of (Y, N)-MSP in less than $2^{\varepsilon n}$ time and then solve it in time $F(k) \cdot |X|^\alpha = F(k) \cdot 2^{C_1 \alpha n/k} < 2^{\varepsilon n}$. ◻

The rest of this section is dedicated to proving Theorems 8.1.

8.1 Construction of MSP Instances Let φ be a 3-SAT formula on n variables and $m := cn$ clauses (without loss of generality, we assume $c \geqslant 1$). Moreover, we have that each variable in φ appears in at most 5 clauses. Let ε and δ be the constants from Gap-ETH. We denote the variable set of φ by $V := \{v_1, \ldots, v_n\}$ and the clause set by $\mathcal{C}$.

Let $\mathsf{var} : \mathcal{C} \to \mathcal{P}([n])$ be the function that maps every clause to the set of indices of the variables that it contains. More formally, if variables $v_{i_1}, v_{i_2}, v_{i_3}$ appear in a clause $C \in \mathcal{C}$ then $\mathsf{var}(C) = \{i_1, i_2, i_3\}$. Moreover, for every $C \in \mathcal{C}$ let $\mathsf{sgn}_C : \mathsf{var}(C) \to \mathbb{F}_2^3$, be the function specifying the sign of the variable appearing in the clause C.

For every large enough $k \in \mathbb{N}$, we have the following construction of an instance (X, n, k) of MSP.

Let t be an even integer we will set later. Equipartition the variable set V into k parts, say, $V_1, \ldots, V_k$ uniformly at random. For every $\vec{\beta} \in \mathbb{F}_2^k$, such that $\Delta(\vec{\beta}, \vec{0}) = t$ (and then we say $\vec{\beta}$ is *t-balanced*), we associate a collection of points in $\mathbb{F}_2^{n+k}$, denoted $X_{\vec{\beta}}$ as follows.

Next, for every t-balanced $\vec{\beta} \in \mathbb{F}_2^k$, we have $\mathcal{C}_{\vec{\beta}} \subseteq \mathcal{C}$ where a clause $C \in \mathcal{C}$ is included in $\mathcal{C}_{\vec{\beta}}$ if and only if for every $i \in \mathsf{var}(C)$ we have some $j \in [k]$ such that $v_i \in V_j$ and $\vec{\beta}(j) = 1$.

We have that $(\vec{x}, \vec{\beta}) \in \mathbb{F}_2^n \times \mathbb{F}_2^k$ is in $X_{\vec{\beta}}$ if and only if the following conditions hold:

- For all $C \in \mathcal{C}_{\vec{\beta}}$ we have that there exists $i \in \mathsf{var}(C)$ such that $\mathsf{sgn}_C(i) = \vec{x}(i)$.
- For all $i \in [n]$ if $v_i \in V_j$ and $\vec{\beta}(j) = 0$ then $\vec{x}(i) = 0$.

Note that $|X_{\vec{\beta}}| \leqslant 2^{tn/k}$. We define $X \subseteq \mathbb{F}_2^{n+k}$ as the union of $X_{\vec{\beta}}$ for all t-balanced $\vec{\beta} \in \mathbb{F}_2^k$.

The polynomial function Y, is simply defined as $Y(i) := \binom{i}{t}$ for all $i \in \mathbb{N}$.

8.2 Completeness Analysis Let $\sigma : [n] \to \{0, 1\}$ be a satisfying assignment to φ. Fix a t-balanced vector $\vec{\beta} \in \mathbb{F}_2^k$. We then construct a vector $\vec{x}^{\vec{\beta}} \in \mathbb{F}_2^n$ in the following way:

$$\forall i \in [n],\ \vec{x}^{\vec{\beta}}(i) := \begin{cases} \sigma(i) & \text{if for some } j \in [k],\ v_i \in V_j \text{ and } \vec{\beta}(j) = 1 \\ 0 & \text{otherwise} \end{cases}.$$

We first claim that $(\vec{x}^{\vec{\beta}}, \vec{\beta}) \in X_{\vec{\beta}}$. This is because:

- For all $C \in \mathcal{C}_{\vec{\beta}}$ we have that there exists $i \in \mathsf{var}(C)$ such that $\mathsf{sgn}_C(i) = \sigma(i)$ because σ is a satisfying assignment. For such an i, note that v_i is some V_j such that $\vec{\beta}(j) = 1$ because of the definition of $\mathcal{C}_{\vec{\beta}}$. Thus, we have $\vec{x}^{\vec{\beta}}(i) = \sigma(i) = \mathsf{sgn}_C(i)$.
- For all $i \in [n]$ if $v_i \in V_j$ and $\vec{\beta}(j) = 0$ then $\vec{x}^{\vec{\beta}}(i) = 0$.

Let $S := \{(\vec{x}^{\vec{\beta}}, \vec{\beta}) \mid \vec{\beta} \in \mathbb{F}_2^k \text{ is } t\text{-balanced}\}$. We claim that S is contained in a k-dimensional subspace. To do so, we provide the basis vectors $\vec{z}_1, \dots, \vec{z}_k$ that span S.

For every $j \in [k]$, we define $(\vec{z}_j, \vec{e}_j) \in \mathbb{F}_2^n \times \mathbb{F}_2^k$ as follows:

$$\forall i \in [n],\ \vec{z}_j(i) := \begin{cases} \sigma(i) & \text{if } v_i \in V_j \\ 0 & \text{otherwise} \end{cases},$$

and $\vec{e}_j$ is the standard basis vector with 1 on the j^{th} coordinate and 0 everywhere else.

Fix some $(\vec{x}^{\vec{\beta}}, \vec{\beta}) \in S$. We note below that it can be spanned by $\vec{z}_1, \dots, \vec{z}_k$, or more precisely,

$$\vec{x}^{\vec{\beta}} = \sum_{j \in [k]} \vec{\beta}(j) \cdot \vec{z}_j, \text{ and } \vec{\beta} = \sum_{j \in [k]} \vec{\beta}(j) \cdot \vec{e}_j.$$

Thus, $|\mathsf{span}(\{\vec{z}_1, \dots, \vec{z}_k\}) \cap S| = Y(k)$.

8.3 Soundness Analysis Let $\delta' = \frac{\delta}{35}$ where δ is the $\mathsf{Gap\text{-}ETH}$ constant.

Suppose there exists some set $T \subseteq \mathbb{F}_2^{n+k}$ of size k such that:

$$|\{i \in [Y(k)] : |X_i \cap \mathsf{span}(T)| \geqslant 1\}| \geqslant (1 - \delta') \cdot \binom{k}{t} := \kappa.$$

Let $B \subseteq \mathbb{F}_2^k$ be such that for every $\vec{\beta} \in B$ we have there is some $(\vec{x}, \vec{\beta}) \in X_{\vec{\beta}} \cap \mathsf{span}(T)$. Moreover, let us denote for every $\vec{\beta} \in B$, a vector in $X_{\vec{\beta}} \cap \mathsf{span}(T)$ by $\left(\vec{x}^{\vec{\beta}}, \vec{\beta}\right)$ (if there is more than one choice, then pick one arbitrarily). Without loss of generality, let us suppose that $|B| = \kappa$.

Let $\widetilde{X} = \left\{\left(\vec{x}^{\vec{\beta}}, \vec{\beta}\right) : \vec{\beta} \in B\right\}$ and we may assume that $\widetilde{X}$ has dimension exactly equal to k, i.e., $\mathsf{dim}(\mathsf{span}(\widetilde{X})) = k$.

For every $(t-1)$-balanced $\vec{\beta}_0 \in \mathbb{F}_2^k$, let $\mathsf{hit}(\vec{\beta}_0)$ be defined as follows:

$$\mathsf{hit}(\vec{\beta}_0) := \left\{\vec{\beta} \in \mathbb{F}_2^k \mid \vec{\beta} \text{ is } t\text{-balanced},\ \Delta(\vec{\beta}, \vec{\beta}_0) = 1, \text{ and } \left(\vec{x}^{\vec{\beta}}, \vec{\beta}\right) \in \widetilde{X}\right\}.$$

Since we know that,

$$\sum_{\substack{\vec{\beta}_0 \in \mathbb{F}_2^k \\ \Delta(\vec{\beta}_0, \vec{0}) = t-1}} \left|\mathsf{hit}(\vec{\beta}_0)\right| = |\widetilde{X}| \cdot t = t \cdot (1 - \delta') \cdot \binom{k}{t},$$

there exists a $(t-1)$-balanced $\vec{\beta}_0^* \in \mathbb{F}_2^k$ such that

$$\left|\mathsf{hit}(\vec{\beta}_0^*)\right| \geqslant \frac{t \cdot (1 - \delta') \cdot \binom{k}{t}}{\binom{k}{t-1}} = (1 - \delta') \cdot (k - t + 1).$$

Let $B^* := \mathsf{hit}(\vec{\beta}_0^*)$. We now define a subset $S^* \subseteq [k]$ of size $|B^*|$ where $j \in [k]$ is in S^* if and only if there is some $\vec{\beta} \in B^*$ such that $\vec{\beta} - \vec{\beta}_0^* = \vec{e}_j$. Let $|S^*| := \gamma$.

We now construct an assignment $\sigma^* : V \to \{0, 1\}$ as follows:

$$\forall i \in [n],\ \sigma^*(v_i) := \begin{cases} \vec{x}^{\vec{\beta}_0^* + \vec{e}_j}(i) & \text{if there exists some } j \in S^* \text{ such that } v_i \in V_j \\ 0 & \text{otherwise} \end{cases}.$$

The rest of the soundness analysis is simply lower bounding the number of clauses in $\mathcal{C}$ that are satisfied by σ^*.

We note that $A := \{(\vec{x}^{\vec{\beta}_0^*+\vec{e}_j}, \vec{\beta}_0^* + \vec{e}_j) \mid j \in S^*\}$ is a set of linearly independent vectors, simply because, $\{\vec{e}_j \mid j \in S^*\}$ is a set of linearly independent vectors. Since $\widetilde{X}$ spans a k-dimensional space, we can pick a set $\widetilde{A}$ of $k-\gamma$ many vectors in $\widetilde{X}$ such that $A \cup \widetilde{A}$ spans $\widetilde{X}$. Let $\widehat{S}^* \subseteq [k]$ be defined as follows: For every $j \in [k]$, we have that $j \in \widehat{S}^*$ if and only if there is some $(\vec{x}^{\vec{\beta}'}, \vec{\beta}') \in \widetilde{A}$ (for some t-balanced vector $\vec{\beta}'$) such that $\vec{\beta}'(j) = 1$. It is clear that $|\widehat{S}^*| \leqslant t \cdot (k-\gamma) \leqslant t \cdot (t-1+\delta'(k-t+1))$. Also since B is k-dimensional, we have that $S^* \cup \widehat{S}^* = [k]$.

We say that a clause $C \in \mathcal{C}$ is hit by A if for all $i \in \mathsf{var}(C)$ we have that if $v_i \in V_j$ (for some $j \in [k]$) then $j \notin \widehat{S}^*$. Also, we say that a clause $C \in \mathcal{C}$ is hit by $\widetilde{X}$ if there is some $(\vec{x}^{\vec{\beta}}, \vec{\beta}) \in \widetilde{X}$ such that for all $i \in \mathsf{var}(C)$ we have that if $v_i \in V_j$ (for some $j \in [k]$) then $\vec{\beta}(j) = 1$.

Claim 8.1. *Suppose a clause $C \in \mathcal{C}$ is hit by both $\widetilde{X}$ and A, then σ^* satisfies C.*

Proof. Let $(\vec{x}^{\vec{\beta}}, \vec{\beta}) \in \widetilde{X}$ such that for all $i \in \mathsf{var}(C)$ we have that if $v_i \in V_j$ (for some $j \in [k]$) then $\vec{\beta}(j) = 1$ and $j \notin \widehat{S}^*$. In other words, we have that if $\vec{\beta}(j) = 1$ then $j \in S^*$. Let $\vec{\lambda} \in \mathbb{F}_2^\gamma \setminus \vec{0}$ where $\vec{\lambda}(j) = 1$ if and only if $\beta(j) = 1$ (where for ease of presentation we assumed S^* to be the first γ positive integers). Notice that:

$$\vec{\beta} = \sum_{w \in [\gamma]} \vec{\lambda}(w) \cdot (\vec{\beta}_0^* + \vec{e}_w),$$

where used the fact that t is even. Thus, we have:

$$\vec{x}^{\vec{\beta}} = \sum_{w \in [\gamma]} \vec{\lambda}(w) \cdot \vec{x}^{\vec{\beta}_0^*+\vec{e}_w}.$$

Since for all $C \in \mathcal{C}_{\vec{\beta}}$ we have that there exists $i^* \in \mathsf{var}(C)$ such that $\mathsf{sgn}_C(i^*) = \vec{x}^{\vec{\beta}}(i^*)$. Next, note that for all $j \in S^* \setminus \widehat{S}^*$, we have that for all $v_i \in V_j$, we have $\vec{x}^{\vec{\beta}_0^*+\vec{e}_j}(i) = \vec{x}^{\vec{\beta}}(i)$. Let $v_{i^*} \in V_{j^*}$. Thus, we have $\mathsf{sgn}_C(i^*) = \vec{x}^{\vec{\beta}}(i^*) = \vec{x}^{\vec{\beta}_0^*+\vec{e}_{j^*}}(i^*) = \sigma^*(v_{i^*})$. ◻

Given Claim 8.1, all we need to now do is bound how many clauses are hit by both $\widetilde{X}$ and A. First the number of clauses not hit by A is at most:

$$5 \cdot \sum_{j \in \widehat{S}^*} |V_j| \leqslant \frac{5n}{k} \cdot t \cdot (t-1+\delta'(k-t+1)) \leqslant 5n\delta' + \frac{5nt^2}{k},$$

where we used that each variable appears in at most 5 clauses.

Next, we estimate the number of clauses not hit by $\widetilde{X}$. We say that a clause $C \in \mathcal{C}$ is hit by a subset R of $[k]$ if for all $i \in \mathsf{var}(C)$ we have that if $v_i \in V_j$ (for some $j \in [k]$) then $j \in R$. With high probability, for every fixing of a non-empty set R of size exactly 3 over the universe $[k]$, the number of clauses hit by R is at most $\frac{12m}{\binom{k}{3}}$.

Consider the bipartite graph H on partite vertex sets $\binom{[k]}{3}$ and $\binom{[k]}{t}$, where we put an edge between $L \in \binom{[k]}{3}$ and $\vec{\beta} \in \binom{[k]}{t}$ (we abuse notation and identify the sets in $\binom{[k]}{t}$ through its characteristic vectors, and thus $\binom{[k]}{t}$ is identified as the set of all t-balanced vectors in $\mathbb{F}_2^k$), if and only if for all $j \in L$ we have $\vec{\beta}(j) = 0$. We would like to estimate the largest subset $\mathcal{R}$ of $\binom{[k]}{3}$ such that there is no edge between $\mathcal{R}$ and B in H. In fact, we simply estimate the largest subset $\mathcal{R}$ of $\binom{[k]}{3}$ such that the size of it's total neighborhood in $\binom{[k]}{t}$ is at most $\delta' \cdot \binom{k}{t}$. By simply counting edges incident on $\mathcal{R}$, we have:

$$|\mathcal{R}| \cdot \binom{k-3}{t-3} \leqslant \delta' \cdot \binom{k}{t} \cdot \binom{t}{3}.$$

Thus we have:

$$|\mathcal{R}| \leqslant \frac{\delta' \cdot \binom{k}{t} \cdot \binom{t}{3}}{\binom{k-3}{t-3}} = \delta' \cdot \frac{k(k-1)(k-2)}{t(t-1)(t-2)} \cdot \binom{t}{3} = \delta' \cdot \binom{k}{3}$$

Putting the above bounds together, we have that the total number of clauses hit by all $R \in \mathcal{R}$ is at most $12\delta' \cdot m$.

Therefore, from Claim 8.1, the total number of clauses not satisfied by σ^* is at most:

$$\delta' \cdot (12m + 5n) + \frac{5nt^2}{k} \leqslant \frac{\delta \cdot m}{2}.$$

Clearly, setting t to be some large even integer (like 100) suffices for the above soundness analysis to go through.

Acknowledgements Karthik C. S. was supported by the National Science Foundation under Grant CCF-2313372, a grant from the Simons Foundation, Grant Number 825876, Awardee Thu D. Nguyen, and was also partially funded by the Ministry of Education and Science of Bulgaria's support for INSAIT, Sofia University "St. Kliment Ohridski" as part of the Bulgarian National Roadmap for Research Infrastructure. Subhash Khot was supported by the NSF Award CCF-1422159, 2130816, the Simons Collaboration on Algorithms and Geometry, and the Simons Investigator Award.

References

[AAM+11] Noga Alon, Sanjeev Arora, Rajsekar Manokaran, Dana Moshkovitz, and Omri Weinstein. Inapproximability of densest κ-subgraph from average case hardness. *Unpublished manuscript*, 1:6, 2011.

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

[ALW13] Amir Abboud, Kevin Lewi, and Ryan Williams. On the parameterized complexity of k-sum. *CoRR*, abs/1311.3054, 2013.

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

[AYZ95] Noga Alon, Raphael Yuster, and Uri Zwick. Color-coding. *J. ACM*, 42(4):844–856, 1995.

[BBE+21] Arnab Bhattacharyya, Édouard Bonnet, László Egri, Suprovat Ghoshal, Karthik C. S., Bingkai Lin, Pasin Manurangsi, and Dániel Marx. Parameterized intractability of even set and shortest vector problem. *J. ACM*, 68(3):16:1–16:40, 2021.

[BCGR23] Huck Bennett, Mahdi Cheraghchi, Venkatesan Guruswami, and João Ribeiro. Parameterized inapproximability of the minimum distance problem over all fields and the shortest vector problem in all lp norms. In Barna Saha and Rocco A. Servedio, editors, *Proceedings of the 55th Annual ACM Symposium on Theory of Computing, STOC 2023, Orlando, FL, USA, June 20-23, 2023*, pages 553–566. ACM, 2023.

[BGH+17] Amey Bhangale, Rajiv Gandhi, Mohammad T Hajiaghayi, Rohit Khandekar, and Guy Kortsarz. Bi-covering: covering edges with two small subsets of vertices. *SIAM Journal on Discrete Mathematics*, 31(4):2626–2646, 2017.

[BGKM18] Arnab Bhattacharyya, Suprovat Ghoshal, Karthik C. S., and Pasin Manurangsi. Parameterized intractability of even set and shortest vector problem from gap-eth. In *45th International Colloquium on Automata, Languages, and Programming, ICALP 2018, July 9-13, 2018, Prague, Czech Republic*, pages 17:1–17:15, 2018.

[BGS98] Mihir Bellare, Oded Goldreich, and Madhu Sudan. Free bits, pcps, and nonapproximability—towards tight results. *SIAM Journal on Computing*, 27(3):804–915, 1998.

[BKN21] Boris Bukh, Karthik C. S., and Bhargav Narayanan. Applications of random algebraic constructions to hardness of approximation. In *62nd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2021, Denver, CO, USA, February 7-10, 2022*, pages 237–244. IEEE, 2021.

[CCK+20] Parinya Chalermsook, Marek Cygan, Guy Kortsarz, Bundit Laekhanukit, Pasin Manurangsi, Danupon Nanongkai, and Luca Trevisan. From gap-exponential time hypothesis to fixed parameter tractable inapproximability: Clique, dominating set, and more. *SIAM J. Comput.*, 49(4):772–810, 2020.

[CFK+15] Marek Cygan, Fedor V. Fomin, Lukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michal Pilipczuk, and Saket Saurabh. *Parameterized Algorithms*. Springer, 2015.

[CFLL23] Yijia Chen, Yi Feng, Bundit Laekhanukit, and Yanlin Liu. Simple combinatorial construction of the $k^{o(1)}$-lower bound for approximating the parameterized k-clique. *CoRR*, abs/2304.07516, 2023.

[CL19] Yijia Chen and Bingkai Lin. The constant inapproximability of the parameterized dominating set problem. *SIAM J. Comput.*, 48(2):513–533, 2019.

[CW89] Aviad Cohen and Avi Wigderson. Dispersers, deterministic amplification, and weak random sources. In *30th Annual Symposium on Foundations of Computer Science*, pages 14–19. IEEE Computer Society, 1989.

[DF13] Rodney G. Downey and Michael R. Fellows. *Fundamentals of Parameterized Complexity.* Texts in Computer Science. Springer, 2013.

[Din07] Irit Dinur. The PCP theorem by gap amplification. *J. ACM*, 54(3):12, 2007.

[Din16] Irit Dinur. Mildly exponential reduction from gap 3sat to polynomial-gap label-cover. *Electronic Colloquium on Computational Complexity (ECCC)*, 23:128, 2016.

[DM18] Irit Dinur and Pasin Manurangsi. Eth-hardness of approximating 2-csps and directed steiner network. In Anna R. Karlin, editor, *9th Innovations in Theoretical Computer Science Conference, ITCS 2018, January 11-14, 2018, Cambridge, MA, USA*, volume 94 of *LIPIcs*, pages 36:1–36:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018.

[Fei98] Uriel Feige. A threshold of ln n for approximating set cover. *J. ACM*, 45(4):634–652, 1998.

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

[FGL+96] Uriel Feige, Shafi Goldwasser, László Lovász, Shmuel Safra, and Mario Szegedy. Interactive proofs and the hardness of approximating cliques. *J. ACM*, 43(2):268–292, 1996.

[FKLM20] Andreas Emil Feldmann, Karthik C. S., Euiwoong Lee, and Pasin Manurangsi. A survey on approximation in parameterized complexity: Hardness and algorithms. *Algorithms*, 13(6):146, 2020.

[GLR+24] Venkatesan Guruswami, Bingkai Lin, Xuandi Ren, Yican Sun, and Kewen Wu. Parameterized inapproximability hypothesis under ETH. In *STOC*, 2024.

[Gol08] Oded Goldreich. *Computational Complexity: A Conceptual Perspective.* Cambridge University Press, New York, NY, USA, 1 edition, 2008.

[GRS24] Venkatesan Guruswami, Xuandi Ren, and Sai Sandeep. Baby PIH: parameterized inapproximability of min CSP. In Rahul Santhanam, editor, *39th Computational Complexity Conference, CCC 2024, July 22-25, 2024, Ann Arbor, MI, USA*, volume 300 of *LIPIcs*, pages 27:1–27:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2024.

[Hås01] Johan Håstad. Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001.

[IP01] Russell Impagliazzo and Ramamohan Paturi. On the complexity of k-SAT. *J. Comput. Syst. Sci.*, 62(2):367–375, 2001.

[IPZ01] Russell Impagliazzo, Ramamohan Paturi, and Francis Zane. Which problems have strongly exponential complexity? *J. Comput. Syst. Sci.*, 63(4):512–530, 2001.

[Kho06] Subhash Khot. Ruling out ptas for graph min-bisection, dense k-subgraph, and bipartite clique. *SIAM Journal on Computing*, 36(4):1025–1071, 2006.

[KK22] Karthik C. S. and Subhash Khot. Almost polynomial factor inapproximability for parameterized k-clique. In Shachar Lovett, editor, *37th Computational Complexity Conference, CCC 2022, July 20-23, 2022, Philadelphia, PA, USA*, volume 234 of *LIPIcs*, pages 6:1–6:21. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.

[KLM19] Karthik C. S., Bundit Laekhanukit, and Pasin Manurangsi. On the parameterized complexity of approximating dominating set. *J. ACM*, 66(5):33:1–33:38, 2019.

[KLM24] Karthik C. S., Euiwoong Lee, and Pasin Manurangsi. On equivalence of parameterized inapproximability of k-median, k-max-coverage, and 2-csp. In *19th International Symposium on Parameterized and Exact Computation, IPEC 2024*, 2024.

[KN21] Karthik C. S. and Inbal Livni Navon. On hardness of approximation of parameterized set cover and label cover: Threshold graphs from error correcting codes. In Hung Viet Le and Valerie King, editors, *4th Symposium on Simplicity in Algorithms, SOSA 2021, Virtual Conference, January 11-12, 2021*, pages 210–223. SIAM, 2021.

[KTST54] P Kővári, Vera T Sós, and Pál Turán. On a problem of zarankiewicz. In *Colloquium Mathematicum*, volume 3, pages 50–57. Polska Akademia Nauk, 1954.

[Lin18] Bingkai Lin. The parameterized complexity of the k-biclique problem. *Journal of the ACM (JACM)*, 65(5):1–23, 2018.

[Lin19] Bingkai Lin. A simple gap-producing reduction for the parameterized set cover problem. In *46th International Colloquium on Automata, Languages, and Programming, ICALP 2019, July 9-12, 2019, Patras, Greece*, pages 81:1–81:15, 2019.

[Lin21] Bingkai Lin. Constant approximating k-clique is w [1]-hard. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1749–1756, 2021.

[LLL24] Shuangle Li, Bingkai Lin, and Yuwei Liu. Improved lower bounds for approximating parameterized nearest codeword and related problems under ETH. *arXiv preprint arXiv:2402.09825*, 2024.

[LRSW22] Bingkai Lin, Xuandi Ren, Yican Sun, and Xiuhan Wang. On lower bounds of approximating parameterized k-clique. In *49th International Colloquium on Automata, Languages, and Programming (ICALP 2022)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2022.

[LRSW23a] Bingkai Lin, Xuandi Ren, Yican Sun, and Xiuhan Wang. Constant approximating parameterized k-setcover is w[2]-hard. In Nikhil Bansal and Viswanath Nagarajan, editors, *Proceedings of the 2023 ACM-SIAM Symposium on Discrete Algorithms, SODA 2023, Florence, Italy, January 22-25, 2023*, pages 3305–3316. SIAM, 2023.

[LRSW23b] Bingkai Lin, Xuandi Ren, Yican Sun, and Xiuhan Wang. Improved hardness of approximating k-clique under eth. *arXiv preprint arXiv:2304.02943*, 2023. To appear in FOCS 2023.

[LRSZ20] Daniel Lokshtanov, MS Ramanujan, Saket Saurab, and Meirav Zehavi. Parameterized complexity and approximability of directed odd cycle transversal. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2181–2200. SIAM, 2020.

[Man17a] Pasin Manurangsi. Almost-polynomial ratio eth-hardness of approximating densest k-subgraph. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 954–961, 2017.

[Man17b] Pasin Manurangsi. Inapproximability of maximum edge biclique, maximum balanced biclique and minimum k-cut from the small set expansion hypothesis. In *44th International Colloquium on Automata, Languages, and Programming (ICALP 2017)*. Schloss-Dagstuhl-Leibniz Zentrum für Informatik, 2017.

[Man20] Pasin Manurangsi. Tight running time lower bounds for strong inapproximability of maximum k-coverage, unique set cover and related problems (via t-wise agreement testing theorem). In Shuchi Chawla, editor, *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms, SODA 2020, Salt Lake City, UT, USA, January 5-8, 2020*, pages 62–81. SIAM, 2020.

[MR16] Pasin Manurangsi and Prasad Raghavendra. A birthday repetition theorem and complexity of approximating dense csps. *CoRR*, abs/1607.02986, 2016.

[NSS95] Moni Naor, Leonard J Schulman, and Aravind Srinivasan. Splitters and near-optimal derandomization. In *Proceedings of IEEE 36th Annual Foundations of Computer Science*, pages 182–191. IEEE, 1995.

[PY91] Christos H. Papadimitriou and Mihalis Yannakakis. Optimization, approximation, and complexity classes. *J. Comput. Syst. Sci.*, 43(3):425–440, 1991.

[Raz98] Ran Raz. A parallel repetition theorem. *SIAM J. Comput.*, 27(3):763–803, 1998.

[Tov84] Craig A. Tovey. A simplified NP-complete satisfiability problem. *Discrete Applied Mathematics*, 8(1):85–89, 1984.

[Wil15] Virginia Vassilevska Williams. Hardness of easy problems: Basing hardness on popular conjectures such as the strong exponential time hypothesis (invited talk). In *10th International Symposium on Parameterized and Exact Computation, IPEC 2015, September 16-18, 2015, Patras, Greece*, pages 17–29, 2015.

[Wil16] Virginia Vassilevska Williams. Fine-grained algorithms and complexity (invited talk). In *33rd Symposium on Theoretical Aspects of Computer Science, STACS 2016, February 17-20, 2016, Orléans, France*, pages 3:1–3:1, 2016.

[Wil18] Virginia Vassilevska Williams. On some fine-grained questions in algorithms and complexity. In *Proc. Int. Cong. of Math.*, volume 3, pages 3431–3472, 2018.

[Zuc96a] David Zuckerman. On unapproximable versions of np-complete problems. *SIAM J. Comput.*, 25(6):1293–1304, 1996.

[Zuc96b] David Zuckerman. Simulating BPP using a general weak random source. *Algorithmica*, 16(4/5):367–391, 1996.