On Average Baby PIH and Its Applications

Yuwei Liu* Yijia Chen† Shuangle Li‡ Bingkai Lin§ Xin Zheng¶

October 14, 2024

Abstract

The *Parameterized Inapproximability Hypothesis* (PIH) asserts that no FPT algorithm can decide whether a given 2CSP instance parameterized by the number of variables is satisfiable, or at most a constant fraction of constraints can be satisfied simultaneously. In a recent breakthrough, Guruswami, Lin, Ren, Sun and Wu (STOC 2024) proved the PIH under the Exponential Time Hypothesis (ETH). But it still remains a major open problem whether the PIH can be established assuming only $\mathsf{W}[1] \neq \mathsf{FPT}$. Guruswami, Ren and Sandeep (CCC 2024) showed a weaker version of the PIH called the *Baby PIH* under $\mathsf{W}[1] \neq \mathsf{FPT}$. They also proposed an intermediate hypothesis called the *Average Baby PIH* that might lead to further progress towards proving the PIH.

Given a 2CSP instance where the number of its variables is the parameter, the Average Baby PIH states that no FPT algorithm can decide whether it is satisfiable or any multi-assignment that satisfies all constraints must assign each variable at least r values on average, for any fixed constant $r > 1$. We prove the $\mathsf{W}[1]$-hardness of this problem by giving an FPT-time self-reduction for 2CSP that turns a gap for multi-assignment in one variable into a gap for multi-assignment on average, thus proving the implication from the Baby PIH to the Average Baby PIH. Since the Baby PIH holds under $\mathsf{W}[1] \neq \mathsf{FPT}$, so does the Average Baby PIH.

As an application, we obtain for the first time the $\mathsf{W}[1]$-hardness of constant approximating k-ExactCover. We also present an alternative proof for the constant inapproximability of parameterized Nearest Codeword Problem with a large hardness factor under the Average Baby PIH.

1 Introduction

In classical complexity theory, the PCP theorem [AS98, ALM+98, Din07] serves as an essential tool for proving most of the existing results in the hardness of approximation. A strengthened version of the PCP theorem is the Multi-Assignment PCP theorem [AMS06, Lemma 11]. It states that for any constant $r > 1$ and $0 < \varepsilon < 1$, there exists no polynomial-time algorithm which can decide

*Shanghai Jiao Tong University, Shanghai, China. Email: yuwei.liu@sjtu.edu.cn
†Shanghai Jiao Tong University, Shanghai, China. Email: yijia.chen@cs.sjtu.edu.cn
‡Nanjing University, Nanjing, China. Email: shuangleli@smail.nju.edu.cn
§Nanjing University, Nanjing, China. Email: lin@nju.edu.cn
¶Nanjing University, Nanjing, China. Email: xinzheng@smail.nju.edu.cn

whether a CSP instance is satisfiable, or any multi-assignment (see Definition 11) that assigns each variable with no more than r values can satisfy less than a $(1-\varepsilon)$-fraction of constraints. The Multi-Assignment PCP theorem was used to show the NP-hardness of approximating SetCover [AMS06]. Although the Multi-Assignment PCP theorem is a simple consequence of the PCP theorem by a straightforward probabilistic argument, it would be ideal to have a direct and purely combinatorial proof. Barto and Kozik [BK22] provided such a proof for the case where $\varepsilon = 0$, that is, there is no gap on the fraction of satisfied constraints, but an arbitrary gap on the "size" of multi-assignment. They call their result the *Baby PCP Theorem*.

As an analog of the PCP theorem in parameterized complexity theory, the Parameterized Inapproximability Hypothesis [LRSZ20], PIH for short, is an important conjecture based on which we can prove many FPT inapproximability results, including the inapproximability of k-Clique, k-ExactCover [GRS24], Direct Odd Cycle Transversal [LRSZ20], etc. It states that for some constant $0 < \varepsilon < 1$, no $f(k) \cdot n^{O(1)}$-time algorithm can distinguish a satisfiable 2CSP instance with k variables from one where less than $(1-\varepsilon)$-fraction of constraints can be satisfied simultaneously [LRSZ20]. Unlike the PCP theorem, the PIH is still a major open problem in parameterized complexity. Currently the state-of-the-art result is that the PIH holds under the Exponential Time Hypothesis (ETH) [GLR+24b, GLR+24a], but a proof of the PIH under the minimum assumption $\mathsf{W}[1] \neq \mathsf{FPT}$ remains elusive. Toward this goal, studying some consequences of the PIH and proving them under $\mathsf{W}[1] \neq \mathsf{FPT}$ might be helpful.

As a recent progress in this direction, Guruswami, Ren, and Sandeep [GRS24] proved a parameterized version of the Baby PCP Theorem, named as the *Baby PIH*, under $\mathsf{W}[1] \neq \mathsf{FPT}$. The Baby PIH states that for any constant $r > 1$, no FPT algorithm can distinguish a satisfiable 2CSP instance from one with no satisfying multi-assignment assigning each variable less than r values. The same as the relationship between the PCP theorem and the Baby PCP theorem, the Baby PIH is a direct consequence of the PIH. As a next step, they suggested a further conjecture, i.e., the *Average Baby PIH* [GRS24, Conjecture 3], which seems to be sandwiched between the PIH and the Baby PIH. It postulates the $\mathsf{W}[1]$-hardness of a problem named Avg-r-Gap-2CSP (see Definition 12). The problem asks for distinguishing a satisfiable 2CSP instance from one with no satisfying multi-assignment assigning each variable less than r values *on average*. In [GRS24], the authors gave a counterexample that for all $r > 1$ and $0 < \varepsilon < 1$, some special 2CSP instance cannot be satisfied by multi-assignments assigning each variable less than r values, but can be satisfied by a multi-assignment assigning $(1+\varepsilon)|X|$ values to X. Compared to proving the PIH under $\mathsf{W}[1] \neq \mathsf{FPT}$, it is apparently easier to show the Average Baby PIH under $\mathsf{W}[1] \neq \mathsf{FPT}$, and studying the Average Baby PIH might bring us further closer to a final proof for the PIH. Furthermore, it turns out that the Average Baby PIH is already sufficient for proving some non-trivial inapproximability results such as k-ExactCover [GRS24]. Finally, we also give a reduction from Avg-r-Gap-2CSP to the parameterized Nearest Codeword Problem (k-NCP), showing that a strengthened version of Average Baby PIH could imply better lower bounds for constant approximating k-NCP.

1.1 Main Results

In this paper, we first present a proof of the Average Baby PIH under $\mathsf{W}[1] \neq \mathsf{FPT}$. A multi-assignment to variables is a relaxed version of assignments. Instead of assigning one value to a variable, a multi-assignment assigns each variable *a set of* values. A 2CSP instance (X, Σ, Φ) is said to be *satisfied* by a multi-assignment if for every constraint φ, one can pick for each variable of φ a value from the set assigned to this variable to satisfy φ. Informally, we say a multi-assignment $\hat{\sigma} : X \to 2^{\Sigma}$ assigns $\sum_{x \in X} |\hat{\sigma}(x)|$ values (see Definition 11). Our main result is the following $\mathsf{W}[1]$-hardness of the Average Baby PIH.

Theorem 1 (Informal, See Theorem 21)**.** *Assuming* $\mathsf{W}[1] \neq \mathsf{FPT}$*, for any constant* $r > 1$*, given a 2CSP instance* $\Pi = (X, \Sigma, \Phi)$ *parameterized by* $|X|$*, no* FPT*-time algorithm can distinguish between:*

- Π *is satisfiable.*
- *No multi-assignment to* X *assigning less than* $r|X|$ *values satisfies* Π*.*

Applying the reduction in [GRS24], we obtain the $\mathsf{W}[1]$-hardness of constant approximating k-ExactCover problem (Definition 13), improving previous hardness result under stronger assumption such as the Gap-ETH [Man20].

Theorem 2 (Theorem 35 restated)**.** *For any constant* $r > 1$*,* r*-approximating* k*-ExactCover is* $\mathsf{W}[1]$*-hard.*

The $\mathsf{W}[1]$-hardness of approximating k-ExactCover problem is a long-standing open problem in parameterized complexity. Although the $\mathsf{W}[1]$, $\mathsf{W}[2]$, ETH-hardness of approximating k-SetCover problem has been established [CL19, KLM19, Lin19, LRSW23], as a special case, the hardness of approximating k-ExactCover is only known under the PIH [Man20]. Our work gives the first proof of its $\mathsf{W}[1]$-hardness.

Another contribution of this paper is a reduction from Avg-r-Gap-2CSP to constant approximating parameterized Maximum Likelihood Decoding problem (Gap-k-MLD), which is a problem asking for k vectors from k sets spanning to a target vector (see Definition 14). This problem is equivalent to the parameterized Nearest Codeword Problem (Gap-k-NCP, Definition 15).

Theorem 3 (Informal, see Theorem 32)**.** *For any prime* p*, constants* $r > \gamma > 1$*, there is a reduction from* Avg-r-Gap-2CSP *instance* $\Pi = (X, \Sigma, \Phi)$ *to* γ-Gap-k'-MLD$_p$ *instance* $(V, \vec{t})$ *with* $k' = O(|\Phi|)$ *and satisfies:*

- *If* Π *is satisfiable, then there is a solution to* $(V, \vec{t})$ *with* k' *vectors;*
- *If satisfying multi-assignment to* X *requires more than* $r|X|$ *values, then solution to* $(V, \vec{t})$ *need at least* $\gamma k'$ *vectors.*

This immediately gives a close relation between running time lower bounds for γ-Gap-k'-NCP$_p$ and Avg-r-Gap-2CSP (e.g., [GLR+24a]).

Corollary 4. *For any prime* p*, if no* $f(k) \cdot n^{o(k)}$*-time algorithm can decide* Avg-r-Gap-2CSP *with* k **constraints** *for any computable function* f*, then* γ-Gap-k-NCP$_p$ *cannot be solved in time* $g(k) \cdot n^{o(k)}$ *for any constant* $\gamma < r$ *and any computable function* g*.*

Previous proofs for the parameterized inapproximability of k-NCP (resp. k-MLD) with large approximation factor [BGKM18, BBE+21, LLL24] all have some blow-up of the parameters that prevent us from establishing tight time lower bounds for the inapproximability of k-NCP under the ETH. Specifically, [BGKM18] first presents a reduction from Gap-2CSP to γ-Gap-k-NCP, which only works *for some constant* $1 < \gamma < \frac{4}{3}$, and then uses tensor product to amply the gap γ to arbitrary constant, with the price of increasing the parameter k. Hence, It is suggested in [Man20,LLL24] to look for an "one-shot" reduction that gives arbitrary constant gap for γ-Gap-k-NCP. Theorem 3 presents such a one-step reduction from Avg-r-Gap-2CSP to γ-Gap-k-MLD *for any constant* $r > \gamma > 1$. Therefore, to establish tighter lower bound for it, it suffices to create large gap for Avg-r-Gap-2CSP, which might be easier than doing directly for γ-Gap-k-NCP.

1.2 Technical Overview

We briefly introduce our techniques to proving the Average Baby PIH and the reduction to Gap-k-MLD.

1.2.1 Local-to-Global Reduction For 2CSP

We give a local-to-global reduction which proves the Average Baby PIH assuming the Baby PIH. Since the Baby PIH holds under $\mathsf{W[1]} \neq \mathsf{FPT}$ (Theorem 18, see also [GRS24]), this proves our main result.

The Baby PIH states that, for all $r > 1$, it's $\mathsf{W[1]}$-hard to decide whether a 2CSP instance is satisfiable, or it cannot be satisfied by any multi-assignment assigning each variable less than r values. The reduction is said to be "local-to-global" in the sense that, the gap in the Baby PIH is local, since it only concerns the gap on one variable assigned with the maximum number of values. On the other hand, the Average Baby PIH concerns gap on the total number of values assigned to all variables, which is "global".

Our construction is simple. Fix some Reed-Solomon code $\mathcal{C} : \mathbb{F}_p^k \to \mathbb{F}_p^m$ for prime $n^{1/k} \le p < 2n^{1/k}$ and $m = \Theta(k^5)$. Given a 2CSP instance $\Pi_0 = (X_0, \Sigma_0, \Phi_0)$, we construct a new 2CSP instance $\Pi = (A \dot\cup B, \Sigma, \Phi)$ as:

- $A = \{a_1, \cdots, a_{|\Phi_0|}\}$, $B = \{b_1, \cdots, b_m\}$.
- Each a_j takes value from the (encoding of) satisfying assignments of $\varphi_j = (x_{i_1} x_{i_2}, C_j) \in \Phi_0$, i.e., $\{(\mathcal{C}(a), \mathcal{C}(b)) : (a, b) \in C_j\}$. Each b_ℓ takes value from $\mathbb{F}_p^k$.
- For each $a_j \in A$ and $b_\ell \in B$, there is a constraint between them. Let $\varphi_j = (x_{i_1} x_{i_2}, C_j) \in \Phi_0$, then the constraint checks whether a_j's assignment $(\vec{u}, \vec{u}')$ and b_ℓ's assignment $\vec{v}$ satisfy $\vec{u}[\ell] = \vec{v}[i_1]$ and $\vec{u}'[\ell] = \vec{v}[i_2]$.

Generally speaking, the assignment to A should form a satisfying multi-assignment to the original instance Π_0, and the assignment to $b_\ell \in B$ tries to guess the ℓ-th entry of the encoding of every x_i. It is easy to see that if the input instance Π_0 is satisfiable, then so does the new 2CSP instance Π.

Now, suppose Π_0 has no satisfying multi-assignment assigning at most r values to each variable (r-multi-assignment), we need to argue that Π has no multi-assignment that assign $r(1-\varepsilon)(|A| + |B|)/2$ values in total (i.e., $r(1-\varepsilon)/2$-average-multi-assignment). For that sake, we use the collision number of code $\mathcal{C}$. Recall that a set $S \subseteq \mathbb{F}_p^m$ collides on an index $\ell \in [m]$ if there exist distinct

$x, y \in S$ such that $x[\ell] = y[\ell]$. Intuitively speaking, if a code $\mathcal{C} \subseteq \mathbb{F}_p^m$ has large relative distance and a set $S \subseteq \mathcal{C}$ collides on at least $(1-\varepsilon)m$ indices for some $\varepsilon \in (0,1)$, then $|S|$ must be larger than $\mathrm{Col}_\varepsilon(\mathcal{C})$, the collision number of $\mathcal{C}$. We can pick a code with $\mathrm{Col}_\varepsilon(\mathcal{C}) \geq \Omega(r|A|)$. Fix a multi-assignment $\hat{\sigma}$ to Π. Let $S_{\hat{\sigma}} \subseteq \mathbb{F}_p^m$ be the set of codewords that are used in the assignment to a_j's by $\hat{\sigma}$. Next, we argue that if Π_0 has no satisfying r-multi-assignment, then for any index $\ell \in [m]$, either b_ℓ takes $r/2$ different guesses under $\hat{\sigma}$, or $S_{\hat{\sigma}}$ has a collision on ℓ. Thus, either we have $r(1-\varepsilon)m/2$ values assigned to b_ℓ's, or $\mathrm{Col}_\varepsilon(\mathcal{C})$ values assigned to a_j's. By duplicating the variables in A and B, we show that Π has no $r(1-\varepsilon)/2$-average-multi-assignment. The details are referred to Section 3.

1.2.2 New Reduction from 2CSP To k-MLD

We present a reduction from Avg-r-Gap-2CSP with rectangular relations to γ-Gap-k-MLD for any $\gamma < r$. Given an instance $\Pi = (X, \Sigma, \Phi)$, the reduction produces for each variable $x_i \in X$ a vector set U_i, and for each constraint $\varphi_j = (x_{i_1}x_{i_2}, C_j) \in \Phi$ a vector set W_j. Using some appropriate coloring technique, we can force any solution to contain vectors from each U_i and W_j.

Each vector $\vec{w} \in W_j$ contains the one-hot encoding of a satisfying assignment (a, b) to φ_j. The one-hot encoding of (a, b) is placed in the j-th part of $\vec{w}$. Each vector $\vec{u} \in U_i$ contains the (minus) one-hot encoding of a value $a \in \Sigma_{x_i}$ assigned to x_i. The one-hot encoding of a is placed in all j-th part of $\vec{u}$ that constraint φ_j depends on variable x_i. Finally, the vector set is $V = U \mathbin{\dot\cup} W$, and the target vector $\vec{t}$ has all zeros in those parts corresponding to constraints, and all ones in the coloring part. See Figure 1 in Section 4 for a pictorial illustration.

Now consider the solution to $(V, \vec{t})$. It will be clear that if Π is satisfiable, then we can pick from each U_i and each W_j a vector according to a satisfying assignment and let them sum to $\vec{t}$. For the soundness, we exploit the rectangularity of constraints in Π. Assume that Π is not r-average-list satisfiable. Let $S \subseteq V$ (along with coefficients) be a solution to $(V, \vec{t})$, we notice that $S \cap U$ naturally induces a multi-assignment to X, then we argue that this multi-assignment satisfies all constraints. For each $j \in [|\Phi|]$, consider the set $S \cap W_j$. Since $\varphi_j = (x_{i_1}x_{i_2}, C_j)$ is a rectangular relation (see Definition 19), let Q denote the underlying set of φ_j and π, ρ be the mapping from $\Sigma_{x_{i_1}}$ and $\Sigma_{x_{i_2}}$ to Q, respectively. Then, $\Sigma_{x_{i_1}}$ can be partitioned by $\pi^{-1}(q)$ for $q \in Q$ and so does $\Sigma_{x_{i_2}}$. By a careful analysis, we show that there exists $q^* \in Q$ that, the partial sum of solution in $S \cap W_j$ having non-zero value in entries of $\pi^{-1}(q^*)$ and $\rho^{-1}(q^*)$ simultaneously. However, the target vector $\vec{t}$ has all-zero in these entries, so the partial sum of solution in $S \cap U$ must also have non-zero value in entries of $\pi^{-1}(q^*)$ and $\rho^{-1}(q^*)$ simultaneously, by the one-hot encoding structure, this indicates that the multi-assignment induced by $S \cap U$ satisfies φ_j. The above analysis holds for all $j \in [|\Phi|]$, so the multi-assignment satisfies all constraints. Since Π is not r-average-list satisfiable, $|S \cap U| > r|X|$.

Since $|S \cap U|$ could be significantly smaller than $|S \cap W|$, we employ a simple construction in [LLL24] to amplify the gap in $S \cap U$. This construction is basically repeating U for sufficiently many times such that its size dominates the whole instance. This finishes the reduction from Avg-r-Gap-2CSP with rectangular relations to γ-Gap-k-MLD problem. More details can be found in Section 4.

1.3 Discussions

The technique of constructing two parts of variables and arguing at least one part has a gap, is quite general in previous works of showing the W[1]-hardness of approximating k-SetCover [CL19, LRSW23], k-NCP [LLL24]. The idea of using collision number of error correcting codes to prove parameterized inapproximability was first introduced in [KN21] and was further developed in [LRSW23, LLL24]. We put forward the question of unifying previous techniques.

Question 5. *Is there a general framework for proving parameterized inapproximability of minimization problems?*

We also raise two questions on strengthening the Average Baby PIH, as it may serve as an important intermediate step for some further results. The first one asks whether the Average Baby PIH can be strengthened to require constant fraction of variables being assiged multiple values. Using the inapproximability of k-Clique [Lin21, KK22, CFLL23], it is not hard to show the following Question 6 holds for the case $r = 2$. We ask if a larger gap can be achieved.

Question 6. *Under* $\mathsf{W}[1] \neq \mathsf{FPT}$*, can we prove that for any constant* $r > 2$*, no* FPT *algorithm can decide whether a 2CSP instance is satisfiable, or any satisfying multi-assignment must have a constant fraction of variables being assigned* r *values?*

The second one asks whether the Average Baby PIH holds for "sparse" instances, i.e., those Avg-r-Gap-2CSP instances with smaller constraint set. This is motivated by the following fact.

Fact 7. *For any constants* $c, r > 1$ *with* $r < 2c$*, if there is no* FPT*-algorithm for* Avg-r-Gap-2CSP *instances with* k *variables and* ck *constraints, then the PIH holds.*

Proof Sketch. Let $\Pi = (X, \Sigma, \Phi)$ be a NO instance of Avg-r-Gap-2CSP. For any single-assignment $\sigma : X \to \Sigma$, assume that σ violates t constraints, then one can simply add at most $2t$ values to σ and obtain a satisfying multi-assignment $\hat{\sigma}$ with total value $k + 2t$. Since Π is a NO instance, we have $k + 2t \geq rk$. Thus, $t \geq \frac{r-1}{2}k$, which is a constant fraction of the constraint number ck. $\square$

Fact 7 gives an alternative way to prove the PIH. Note that our proof for the Average Baby PIH gives instances with $\Omega(k^2)$ constraints. This motivates the following question.

Question 8. *Can we prove that for any constant* $r > 1$*,* Avg-r-Gap-2CSP *is* $\mathsf{W}[1]$*-hard even on instances with* k *variables and sparse (say,* $o(k^2)$*) constraints?*

1.4 Organization

In Section 2 we introduce the main computational problems and complexity assumptions studied in this paper. As the central contribution, Section 3 explains our reduction from the Baby PIH to the average PIH. This in fact establishes the Average Baby PIH under $\mathsf{W}[1] \neq \mathsf{FPT}$. Due to space limitations, we discuss some applications of the Average Baby PIH in the appendices. In particular, Section 4 gives a reduction from Avg-r-Gap-2CSP to the constant approximation of k-MLD. And in Appendix A we present a reduction from Avg-r-Gap-2CSP to the constant approximation of k-ExactCover, which slightly differs from the construction in [GRS24].

2 Preliminaries

For a positive integer n, we use $[n]$ denote the set $\{1, 2, \cdots, n\}$. $S_1 \dot\cup \cdots \dot\cup S_k$ is the *disjoint union* of sets $S_1, \ldots, S_k$, where we tacitly assume that $S_1, \ldots, S_k$ are pairwise disjoint. We use log (without subscript) denote the logarithm number with base 2. For any prime number p, we write $\mathbb{F}_p$ for the (unique) finite field of size p. The reader is assumed to be familiar with basic notions in parameterized complexity theory, in particular FPT and $\mathsf{W}[1]$. Otherwise, the standard references are, e.g., [FG06, DF13, CFK+15].

2.1 Problems

Definition 9 (Parameterized 2CSP). *A* 2CSP *instance is defined as a triple* $\Pi = (X, \Sigma, \Phi)$ *where:*

- X *is a set of variable.*
- $\Sigma = \dot\bigcup_{x \in X} \Sigma_x$*, where each* Σ_x *has size* n*, containing values which the variable* $x \in X$ *can be assigned.*
- $\Phi = \{\varphi_1, \cdots, \varphi_{k'}\}$*, where each* $\varphi_j = (x_{i_1} x_{i_2}, C_j)$ *for some* $x_{i_1}, x_{i_2} \in X$*, and* C_i *is a subset of* $\Sigma_{x_{i_1}} \times \Sigma_{x_{i_2}}$*.*

The problem is to decide whether there exists an assignment $\sigma : X \to \Sigma$ *that satisfies:*

- *For all* $x \in X$*,* $\sigma(x) \in \Sigma_x$*.*
- *For all* $\varphi_j = (x_{i_1} x_{i_2}, C_j) \in \Phi$*,* $(\sigma(x_{i_1}), \sigma(x_{i_2})) \in C_j$*.*

The parameter for this problem is $k = |X|$*, the number of variables. Each pair of variables has at most one constraint, so* $|\Phi| \le \binom{k}{2}$*. Without loss of generality, each variable is related to some constraint in* Φ*. The size of instance* Π *is defined as* $|\Pi| = |\Sigma| + |\Phi|$*, where each* $|\varphi_j| = |C_j|$*.*

The approximation of parameterized 2CSP is referred to the following problem.

Definition 10 (ε-Gap-2CSP). *Given a 2CSP instance* $\Pi = (X, \Sigma, \Phi)$ *with parameter* $k = |X|$*, distinguish between:*

- Π *is satisfiable;*
- *any assignment can satisfy at most* ε*-fraction constraints in* Φ*.*

The notion of multi-assignment extends the usual assignment where each variable can be assigned multiple values.

Definition 11 (Multi-Assignment). *A multi-assignment of 2CSP instance* $\Pi = (X, \Sigma, \Phi)$ *is a function* $\hat\sigma : X \to 2^\Sigma$*, where* 2^Σ *is the power set of* Σ*, such that for all* $x \in X$*,* $\hat\sigma(x) \subseteq \Sigma_x$*. A multi-assignment* $\hat\sigma$ *satisfies the instance* Π *if:*

- *For all* $\varphi_j = (x_{i_1} x_{i_2}, C_j) \in \Phi$*, there exists* $c_1 \in \hat\sigma(x_{i_1}), c_2 \in \hat\sigma(x_{i_2})$ *that* $(c_1, c_2) \in C_j$*.*

For multi-assignment $\hat{\sigma}$, the individual value *of $\hat{\sigma}$ is defined as* $\max_{x\in X} |\hat{\sigma}(x)|$, *and the* total value *of $\hat{\sigma}$ is defined as* $\sum_{x\in X} |\hat{\sigma}(x)|$.

To ease notation, for $r \geq 1$, we say that a 2CSP instance $\Pi = (X, \Sigma, \Phi)$ is *r-list satisfiable* if there exists a multi-assignment $\hat{\sigma}$ with individual value no more than r which satisfies Π, and say that Π is *r-average list satisfiable* if there exists a multi-assignment $\hat{\sigma}$ with total value no more than $r|X|$ which satisfies Π. This motivates the following problem.

Definition 12 (Avg-r-Gap-2CSP). *Given a* 2CSP *instance Π, it is asked to distinguish between the following two cases:*

- Π *is satisfiable.*
- Π *is not r-average list satisfiable.*

We also consider the k-ExactCover problem (a.k.a. the k-UniqueSetCover problem) and k-MLD/k-NCP problem defined below.

Definition 13 (k-ExactCover). *Given a set U (which we call universe) and a collection of U's subsets $\mathcal{S}$, the k-*ExactCover *problem asks for distinguishing between:*

- *there exist at most k disjoint sets in $\mathcal{S}$ that form a partition of U,*
- *or U is not the union of any k sets in $\mathcal{S}$.*

Definition 14 (k-MLD). *For prime p, integer $d > 0$, given a (multi-)set V of vectors in $\mathbb{F}_p^d$, and a target vector $\vec{t} \in \mathbb{F}_p^d$, the k-MLD$_p$ problem asks for distinguishing between:*

- *$\vec{t}$ is in the linear space spanned by k vectors in V,*
- *or $\vec{t}$ is not in the linear space spanned by any k vectors in V.*

Without loss of the generality, we can assume V is partitioned into k parts and the YES case requires picking from each part a vector.

[LLL24] showed the condition of YES instance of k-MLD problem can be strengthened to requiring the existence of k vectors summing to $\vec{t}$.

Definition 15 (k-NCP). *For prime p, integer $d > 0$, given a (multi-)set V of vectors in $\mathbb{F}_p^d$, and a target vector $\vec{t} \in \mathbb{F}_p^d$, the k-NCP$_p$ problem asks for distinguishing between:*

- *the Hamming distance between $\vec{t}$ and the vector space spanned by V is at most k,*
- *or the Hamming distance between $\vec{t}$ and the vector space spanned by V is at least $k+1$.*

The k-NCP and k-MLD are equivalent (cf. [LLL24]), so we focus on the k-MLD problem. For simplicity, we omit the finite field size p when the result holds for all prime p.

2.2 Hypotheses

Hypothesis 16 (PIH [LRSZ20])**.** *For every constant $0 < \varepsilon < 1$, no FPT algorithm for ε-*Gap-*2CSP.*

The Baby PIH, a hypothesis implied by PIH, states the hardness of approximating individual value of satisfying multi-assignment. Formally,

Hypothesis 17 (Baby PIH [GRS24])**.** *For any constant $r > 0$, no* FPT *algorithm can on input a 2CSP instance, distinguish whether it is satisfiable, or cannot be satisfied by any multi-assignment with* individual *value less than r.*

The Baby PIH (Theorem 18) is a hardness hypothesis against a **local** property, i.e., the individual value of satisfying assignments. It is shown that the standard assumption $\mathsf{W}[1] \neq \mathsf{FPT}$ implies Baby PIH:

Theorem 18 ([GRS24])**.** *The Baby PIH (Hypothesis 17) holds under* $\mathsf{W}[1] \neq \mathsf{FPT}$.

To define the Average Baby PIH, we first introduce a useful property concerning relations in a 2CSP instance.

Definition 19 (Rectangular relation)**.** *A 2CSP instance $\Pi = (X, \Sigma, \Phi)$ is said to have* rectangular relations *if for each $\varphi_j = (x_{i_1}x_{i_2}, C_j) \in \Phi$, there exists a set Q_j and mappings $\pi_j, \rho_j : \Sigma \to Q_j$, such that $(a, b) \in C_j$ iff $\pi_j(a) = \rho_j(b)$. We call Q_j the* underlying set *of φ_j.*

The name "rectangular" comes from the following intuition. Recall that a subset $S \subseteq \Sigma^2$ is a *rectangle* if and only if there exist $A, B \subseteq \Sigma$ such that $S = A \times B$. It is easy to verify that $R \subseteq \Sigma^2$ is rectangular if and only if R is the union of a set of pairwise disjoint rectangles. [GRS24] raised the Average Baby PIH on hardness result against a **global** property, i.e., the total value of satisfying assignments.

Hypothesis 20 (Average Baby PIH)**.** *For any constant $r > 0$, there exists no* FPT *algorithm for the* Avg-r-Gap-2CSP *problem, even for instances with only rectangular relations.*

3 Average Baby PIH from Baby PIH

We show that the Average Baby PIH (Hypothesis 20) even with rectangular relations is implied by the Baby PIH. More precisely, we employ a local-to-global reduction developed in [LLL24] to amplify the local gap for one variable (Theorem 18) into a global gap for all variables.

Theorem 21 (Average Baby PIH)**.** *Under* $\mathsf{W}[1] \neq \mathsf{FPT}$*, for any constant $r > 0$, no* FPT *algorithm can distinguish a given* 2CSP *instance with rectangular relation is satisfiable, or cannot be satisfied by multi-assignment with total value no more than r.*

To show Theorem 21, we first introduce some gap-creating tools. The *collision number* of an error-correcting code characterizes the number of codewords needed to find "collision" on a constant fraction of coordinates. We use the definition in [LLL24]:

Definition 22 (ε-Collision Number)**.** *Let $m \geq 1$ and $x, y \in \Sigma^m$ with $x \neq y$. For every $i \in [m]$ we say that x and y* collide *on position i if $x[i] = y[i]$. Furthermore, a subset $S \subseteq \Sigma^m$* collides *on position i if there exist distinct $x, y \in S$ with $x[i] = y[i]$. We define the* collision set *of S as*

$$\mathrm{ColSet}(S) = \big\{ i \in [m] \;\big|\; S \text{ collides on position } i \big\}.$$

Observe that if $|S| \leq 1$, then $\mathrm{ColSet}(S) = \varnothing$.

Now for every $C \subseteq \Sigma^m$ and $0 < \varepsilon < 1$ the ε-collision number of C*, denoted by $\mathrm{Col}_\varepsilon(C)$, is the* maximum *$s \leq |C| + 1$ such that for* all *$S \in \binom{C}{s-1}$ we have*

$$\big|\mathrm{ColSet}(S)\big| \leq \varepsilon m.$$

The relationship between collision number and distance has been established. For Reed-Solomon codes we have:

Theorem 23 (Theorem 10 in [LRSW23], see also [KN21])**.** *For any $0 < \varepsilon < 1$, any Reed-Solomon code $\mathcal{C}^{RS} : \mathbb{F}_p^k \to \mathbb{F}_p^m$ with sufficiently large $k < m \leq p$, $\mathrm{Col}_\varepsilon(\mathcal{C}^{RS}) \geq \sqrt{\frac{2\varepsilon m}{k}}$.*

To prove the Average Baby PIH (Theorem 21), we first give an intermediate step. This step reduces any 2CSP instance to a new bipartite 2CSP instance whose variables are partitioned into two groups and obtains different gaps for average multi-assignments to both variable groups. Below, we introduce the notion *Average Multi-Assignment* for bipartite 2CSPs. In Lemma 27, we will show that the gaps for average multi-assignments to bipartite 2CSP can be transferred into that for one variable set, completing the proof for Theorem 21.

Definition 24 ((r, s)-Average Multi-Assignment)**.** *Let $\Pi = (X, \Sigma, \Phi)$ be a* bipartite 2CSP *instance, in particular $X = X_1 \,\dot\cup\, X_2$ and every $\varphi = (x_1 x_2, C) \in \Phi$ has $x_1 \in X_1$ and $x_2 \in X_2$. Then for $r_1, r_2 \geq 1$ an* (r_1, r_2)-average multi-assignment *of Π is a multi-assignment $\hat\sigma : X \to 2^\Sigma$ such that*

$$\frac{\sum_{x \in X_1} |\hat\sigma(x)|}{|X_1|} \leq r_1 \quad \text{and} \quad \frac{\sum_{x \in X_2} |\hat\sigma(x)|}{|X_2|} \leq r_2.$$

That is, the total value of $\hat\sigma$ restricted to X_1 is at most r_1, and the total value of $\hat\sigma$ restricted to X_2 is at most r_2 (cf. Definition 11). We say Π is (r_1, r_2)-average list satisfiable *if there is an (r_1, r_2)-average multi-assignment which satisfies Π.*

The intermediate step of our reduction is the following Lemma.

Lemma 25. *There is an algorithm $\mathcal{A}$ which on input a* 2CSP *instance $\Pi_0 = (X_0, \Sigma_0, \Phi_0)$, $\varepsilon > 0$, and $r \geq 1$ computes a bipartite* 2CSP *instance $\Pi = (X_1 \,\dot\cup\, X_2, \Sigma, \Phi)$ with the following properties.*

Completeness. *If Π_0 is satisfiable, then so is Π,*

Soundness. *For every $r \geq 1$ if Π_0 is not $2r$-list satisfiable, then Π is not (r_1, r_2)-average list satisfiable for every $r_1, r_2 \in \mathbb{N}$ with*

$$r_1 + r_2 \leq 2(1 - \varepsilon) r.$$

Rectangularity. *All constraints in Φ are rectangular.*

In addition, there exists a computable function f upper bounding the running time of $\mathcal{A}$ as

$$f(|X_0| + |\Phi_0| + 1/\varepsilon + r)|\Sigma_0|^{O(1)}. \tag{1}$$

And the number of variables $|X_1| + |X_2|$ and the number of constraints $|\Pi|$ in Π can also be upper bounded by $f(|X_0| + |\Phi_0| + 1/\varepsilon + r)$.

Proof. For the given 2CSP instance $\Pi_0 = (X_0, \Sigma_0, \Phi_0)$ we let

$$k = |X_0| \quad \text{and} \quad k' = |\Phi_0|.$$

Thereby we fix some enumerations of the variables in X_0 and the constraints in Φ_0 as

$$X_0 = \{x_1, \dots, x_k\} \quad \text{and} \quad \Phi_0 = \{\varphi_1, \dots, \phi_{k'}\}.$$

Let $\mathcal{C} : \mathbb{F}_p^k \to \mathbb{F}_p^{k''}$ be a Reed-Solomon code with

$$2|\Sigma_0|^{1/k} > p \geq |\Sigma_0|^{1/k} \quad \text{and} \quad k'' = \left\lfloor \frac{8(1-\varepsilon)^2 r^2}{\varepsilon} k(k')^2 \right\rfloor + 1.$$

Clearly $|\Sigma_0| \leq p^k$, and therefore we can assume without loss of generality

$$\Sigma_0 \subseteq \mathbb{F}_p^k.$$

Moreover, we only consider the case that

$$k'' \leq p \left(= \left|\mathbb{F}_p\right| \right) < 2|\Sigma_0|^{1/k},$$

i.e., Σ_0 is sufficiently larger than k and k'.[1] Hence we can invoke Theorem 23 on $\Sigma \leftarrow \mathbb{F}_p$, $k \leftarrow k$, $m \leftarrow k''$, and $\varepsilon \leftarrow \varepsilon$ to obtain

$$\mathrm{Col}_\varepsilon(\mathcal{C}(\mathbb{F}_p^k)) \geq \sqrt{\frac{2\varepsilon k''}{k}} > 4(1-\varepsilon)rk', \tag{2}$$

where the second inequality is by our choice of k''.

Now the algorithm $\mathcal{A}$ constructs the following bipartite 2CSP instance $\Pi = (X, \Sigma, \Phi)$.

Variables. $X = X_1 \dot\cup X_2$ with

$$X_1 = \{u_1, \dots, u_{k'}\} \quad \text{and} \quad X_2 = \{v_1, \dots, v_{k''}\}.$$

Alphabets. $\Sigma = \bigcup_{u \in X_1} \Sigma_u \cup \bigcup_{v \in X_2} \Sigma_v$ where:

[1] Otherwise, the original instance Π_0 can be solved in time of the form (1), and we can then output some predetermined Π depending on whether Π_0 is satisfiable.

- For every $j \in [k']$ the alphabet of the variable $u_j \in X_1$ is

$$\Sigma_{u_j} = \left\{ \left(\mathcal{C}(a_1), \mathcal{C}(a_2)\right) \;\middle|\; \varphi_j = (x_{i_1} x_{i_2}, C) \text{ and } (a_1, a_2) \in C \right\} \subseteq \left(\mathcal{C}(\mathbb{F}_p^k)\right)^2 \subseteq \left(\mathbb{F}_p^{k''}\right)^2. \quad (3)$$

  That is, Σ_{u_j} contains all the (partial) satisfying assignments of φ_j encoded by $\mathcal{C} : \mathbb{F}_p^k \to \mathbb{F}_p^{k''}$ as pairs of vectors in $\mathbb{F}_p^{k''}$. (Recall $\Sigma_0 \subseteq \mathbb{F}_p^k$.)

- For every $\ell \in [k'']$ we have $\Sigma_{v_\ell} = \mathbb{F}_p^k$. Since $p < 2|\Sigma_0|^{1/k}$, we have $|\Sigma_{v_\ell}| \le 2^k |\Sigma_0|$.

Constraints. Let $j \in [k']$ and $\varphi_j = (x_{i_1} x_{i_2}, C)$. Then for every $\ell \in [k'']$ we have a constraint between the variable $u_j \in X_1$ and $v_\ell \in X_2$ which checks whether u_j is assigned to $(w_1, w_2) \in \left(\mathbb{F}_p^{k''}\right)^2$ and v_ℓ to $s \in \mathbb{F}_q^\ell$ such that

$$w_1[\ell] = s[i_1] \quad \text{and} \quad w_2[\ell] = s[i_2]. \quad (4)$$

Consequently (4) implies that the constraint is rectangular.[2] Moreover, the number of constraints in Π is

$$k'k'' = k' \left\lfloor \frac{2(1-\varepsilon)^2 r^2}{\varepsilon} k(k')^2 \right\rfloor + k'.$$

The completeness of our reduction is straightforward. So we turn to the soundness. In particular, we assume that the given 2CSP instance Π_0 is *not $2r$-list satisfiable*. Furthermore, let $\hat{\sigma} : X \to 2^\Sigma$ be a satisfying multi-assignment for Π. We need to show that, for any $r_1, r_2 \in \mathbb{N}$ if there is a satisfying (r_1, r_2)-average multi-assignment $\hat{\sigma}$, then

$$r_1 + r_2 > 2(1-\varepsilon)r. \quad (5)$$

To that end, let

$$\mathsf{Word}_{\hat{\sigma}} = \bigcup_{u_j \in X_1} \bigcup_{(w_1, w_2) \in \hat{\sigma}(u_j)} \{w_1, w_2\} \subseteq \mathbb{F}_p^{k''}. \quad (6)$$

That is, $\mathsf{Word}_{\hat{\sigma}}$ is the set of all codewords in $\mathbb{F}_p^{k''}$ that $\hat{\sigma}$ uses for the variables in X_1.

Claim 26. *Let $\ell \in [k'']$ with $|\hat{\sigma}(v_\ell)| \le 2r$. Then $\mathsf{Word}_{\hat{\sigma}}$ collides on position ℓ.*

Proof of Claim 26. Let $\ell \in [k'']$ be fixed with $|\hat{\sigma}(v_\ell)| \le 2r$.

Consider an arbitrary constraint $\varphi_j = (x_{i_1} x_{i_2}, C) \in \Phi_0$ $\big($i.e., $j \in [k']\big)$. Since $\hat{\sigma}$ is a satisfying multi-assignment for Π, there exist

$$(w_1, w_2) \in \hat{\sigma}(u_j) \subseteq \Sigma_{u_j} \subseteq \left(\mathbb{F}_p^{k''}\right)^2 \quad \text{and} \quad s \in \hat{\sigma}(v_\ell) \subseteq \mathbb{F}_p^k$$

such that $u_j = (w_1, w_2)$ and $v_\ell = s$ satisfy the constraint between u_j and v_ℓ in Π. By $(w_1, w_2) \in \Sigma_{u_j}$ and (3) there are $a_1, a_2 \in \Sigma_0$ with $w_1 = \mathcal{C}(a_1)$ and $w_2 = \mathcal{C}(a_2)$ such that

$$x_{i_1} = a_1 \text{ and } x_{i_2} = a_2 \text{ satisfy } \varphi_j. \quad (7)$$

[2] To see this, we take $\pi(u_j) = \pi(w_1, w_2) = (w_1[\ell], w_2[\ell])$ and $\rho(v_\ell) = (v_\ell[i_1], v_\ell[i_2])$. Then equation (4) is precisely $\pi(u_j) = \rho(v_\ell)$ as in Definition 19.

Then we say that a_1 is $(\hat{\sigma}, \varphi_j)$-*suitable* for x_{i_1} *with respect to* s, and similarly a_2 is $(\hat{\sigma}, \varphi_j)$-*suitable* for x_{i_2} *with respect to* s.

In addition, by (4)

$$\mathcal{C}(a_1)[\ell] = s[i_1] \quad \text{and} \quad \mathcal{C}(a_2)[\ell] = s[i_2]. \tag{8}$$

Now we define a *multi-assignment* $\hat{\sigma}_0 : X_0 \to 2^{\Sigma_0}$ for the original instance $\Pi_0 = (X_0, \Sigma_0, \Phi_0)$ as follows. For every $x \in X_0$ let

$$\hat{\sigma}_0(x) = \bigcup_{s \in \hat{\sigma}(v_\ell)} \left\{ a \in \Sigma_0 \;\middle|\; j \in [k'] \text{ and } a \text{ is } (\hat{\sigma}, \varphi_j)\text{-suitable for } x \text{ with respect to } s \right\}. \tag{9}$$

(Recall that we have fixed an $\ell \in [k'']$ and hence $\hat{\sigma}(v_\ell)$.) Since every variable x must appear in at least one constraint $\varphi_j \in \Phi_0$ (cf. Definition 9), it is easy to see that $\hat{\sigma}_0$ is a satisfying multi-assignment for Π_0 by (7).

As Π_0 is *not* $2r$-list satisfiable, there is an $x_{i^*} \in X_0$ (i.e., $i^* \in [k]$) with

$$\left| \hat{\sigma}_0(x_{i^*}) \right| \geq 2r + 1.$$

We have assumed that

$$|\hat{\sigma}(v_\ell)| \leq 2r,$$

so by (9) there is an $s \in \hat{\sigma}(v_\ell)$ such that

$$\left| \left\{ a \in \Sigma_0 \;\middle|\; j \in [k'] \text{ and } a \text{ is } (\hat{\sigma}, \varphi_j)\text{-suitable for } x_{i^*} \text{ with respect to } s \right\} \right| \geq 2.$$

Hence there are $a_1, a_2 \in \Sigma_0$ with $a_1 \neq a_2$ and $j_1, j_2 \in [k']$ such that

- a_1 is $(\hat{\sigma}, \varphi_{j_1})$-suitable for x_{i^*} with respect to s,
- and a_2 is $(\hat{\sigma}, \varphi_{j_2})$-suitable for x_{i^*} with respect to s.

Then (8) implies that

$$\mathcal{C}(a_1)[\ell] = s[i^*] = \mathcal{C}(a_2)[\ell].$$

In other words, $\mathcal{C}(a_1)$ and $\mathcal{C}(a_2)$ collide on position ℓ. Clearly $\mathcal{C}(a_1), \mathcal{C}(a_2) \in \mathsf{Word}_{\hat{\sigma}}$, so this finishes the proof of the claim. ⊣

Let

$$r_1 = \frac{\sum_{x \in X_1} |\hat{\sigma}(x)|}{|X_1|} = \frac{\sum_{j \in [k']} |\hat{\sigma}(u_j)|}{k'} \quad \text{and} \quad r_2 = \frac{\sum_{x \in X_2} |\hat{\sigma}(x)|}{|X_2|} = \frac{\sum_{\ell \in [k'']} |\hat{\sigma}(v_\ell)|}{k''}.$$

Now we distinguish two cases.

- There are more than ε fraction of $\ell \in [k'']$ such that $|\hat{\sigma}(v_\ell)| \leq 2r$, then Claim 26 implies that $\mathsf{Word}_{\hat{\sigma}}$ collides on more than ε fraction of positions $\ell \in [k'']$. Recall (2), i.e.,

$$\mathrm{Col}_\varepsilon(\mathcal{C}(\mathbb{F}_p^k)) \geq \sqrt{\frac{2\varepsilon k''}{k}} > 4(1-\varepsilon) r k'.$$

Hence,

$$\left|\mathsf{Word}_{\hat{\sigma}}\right| \geq \mathrm{Col}_{\varepsilon}(\mathcal{C}(\mathbb{F}_p^k)) > 4(1-\varepsilon)rk'.$$

By the definition (6) of $\mathsf{Word}_{\hat{\sigma}}$ we deduce

$$\begin{aligned}\left|\mathsf{Word}_{\hat{\sigma}}\right| &= \left|\bigcup_{u_j \in X_1} \bigcup_{(w_1,w_2)\in\hat{\sigma}(u_i)} \{w_1, w_2\}\right| \\ &\leq \sum_{u_j\in X_1} \left|\bigcup_{(w_1,w_2)\in\hat{\sigma}(u_i)} \{w_1,w_2\}\right| \leq \sum_{u_j\in X_1} 2\left|\hat{\sigma}(u_i)\right|\end{aligned}$$

It follows that

$$r_1 = \frac{\sum_{j\in[k']} \left|\hat{\sigma}(u_j)\right|}{k'} > \frac{4(1-\varepsilon)rk'}{2k'} = 2(1-\varepsilon)r.$$

- There are at most ε fraction of $\ell \in [k'']$ with $|\hat{\sigma}(v_\ell)| \leq 2r$. Or equivalently, there are at least $(1-\varepsilon)$ fraction of $\ell \in [k'']$ with $|\hat{\sigma}(v_\ell)| \geq 2r+1$. Then

$$r_2 = \frac{\sum_{\ell\in[k'']} \left|\hat{\sigma}(v_\ell)\right|}{k''} \geq \frac{(1-\varepsilon)k''(2r+1)+\varepsilon k''}{k''} > 2(1-\varepsilon)r.$$

So both cases lead to (5) as desired. □

With some proper replication, the unbalanced (r_1, r_2)-gap can be turned into a balanced one, and yield the final r-average list unsatisfiability.

Lemma 27. *For any bipartite* 2CSP *instance* $\Pi = (X_1 \dot\cup X_2, \Sigma, \Phi)$ *and* $r > 1$ *we can compute in polynomial time a* 2CSP *instance* $\Pi' = (X', \Sigma', \Phi')$ *with*

$$|X| = 2|X_1||X_2|$$

such that

Completeness. *If* Π *is satisfiable, then so is* Π',

Soundness. *Let* $r \geq 1$. *If* Π *is not* (r_1, r_2)*-average list satisfiable for* every $r_1, r_2 \geq 1$ *with* $r_1 + r_2 \leq 2r$, *then* Π' *is not* r*-average list satisfiable. Or equivalently, if* Π' *is* r*-average list satisfiable, then for* some $r_1, r_2 \in \mathbb{N}$ *with* $r_1 + r_2 \leq 2r$ *the bipartite* Π *is* (r_1, r_2)*-average list satisfiable.*

Furthermore, if Π *is rectangular, then so is* Π'.

Proof. Let

$$k_1 = |X_1| \quad \text{and} \quad k_2 = |X_2|.$$

The desired $\Pi' = (X', \Sigma', \Phi')$ is constructed as below.

Variables. X' consists of k_2 copies of X_1 and k_1 copies of X_2, i.e., $X' = X'_1 \dot\cup X'_2$ where

$$X'_1 = \left\{x^{(i)} \;\middle|\; x \in X_1 \text{ and } i \in [k_2]\right\} \quad \text{and} \quad X'_2 = \left\{x^{(i)} \;\middle|\; x \in X_2 \text{ and } i \in [k_1]\right\}.$$

Note, $|X'_1| = |X'_2| = k_1k_2$, therefore Π' contains $2k_1k_2$ many variables.

Alphabets. $\Sigma' = \bigcup_{x\in X'} \Sigma'_x$ where:

- For every $x \in X_1$ and $i \in [k_2]$ let $\Sigma'_{x^{(i)}} = \Sigma_x$. Recall that $\Sigma_x \subseteq \Sigma$ is the alphabet for the variable x in the original 2CSP instance Π.
- Similarly, for every $x \in X_2$ and $i \in [k_1]$ let $\Sigma'_{x^{(i)}} = \Sigma_x$.

Constraints. For every constraint $\varphi = (x_1x_2, C) \in \Phi$ with $x_1 \in X_1$ and $x_2 \in X_2$, $i_1 \in [k_2]$, and $i_2 \in [k_1]$ we have a constraint

$$\varphi^{i_1,i_2} = \left(x_1^{(i_1)}x_2^{(i_2)}, C\right) \in \Phi'.$$

That is, φ^{i_1,i_2} is a copy of φ where the variable x_1 is replaced by its i_1-th copy $x_1^{(i_1)}$ and x_2 by its i_2-th copy $x_2^{(i_2)}$. It immediately implies that if Π is rectangular, then Π' is rectangular too.

Again the completeness is immediate. Towards the soundness, let $\hat\sigma' : X' \to 2^{\Sigma'}$ be a satisfying r-average multi-assignment for Π'. In particular,

$$r = \frac{\sum_{x\in X'} \left|\hat\sigma'(x)\right|}{|X'|} = \frac{\sum_{x\in X'_1} \left|\hat\sigma'(x)\right| + \sum_{x\in X'_2} \left|\hat\sigma'(x)\right|}{|X'_1| + |X'_2|} = \frac{\sum_{x\in X'_1} \left|\hat\sigma'(x)\right| + \sum_{x\in X'_2} \left|\hat\sigma'(x)\right|}{2k_1k_2}.$$

We set

$$r_1 = \frac{\sum_{x\in X'_1} \left|\hat\sigma'(x)\right|}{|X'_1|} = \frac{\sum_{x\in X'_1} \left|\hat\sigma'(x)\right|}{k_1k_2} \quad \text{and} \quad r_2 = \frac{\sum_{x\in X'_2} \left|\hat\sigma'(x)\right|}{|X'_2|} = \frac{\sum_{x\in X'_2} \left|\hat\sigma'(x)\right|}{k_1k_2} \tag{10}$$

It follows that

$$r_1 + r_2 = \frac{\sum_{x\in X'_1} \left|\hat\sigma'(x)\right| + \sum_{x\in X'_2} \left|\hat\sigma'(x)\right|}{k_1k_2} = 2r.$$

Note that

$$X'_1 = \dot{\bigcup_{i\in[k_2]}} \left\{x^{(i)} \mid x \in X_1\right\}. \tag{11}$$

Therefore,

$$\begin{aligned}
r_1k_1k_2 &= r_1|X'_1| && \left(\text{by } |X'_1| = k_1k_2\right)\\
&= \sum_{x\in X'_1} \left|\hat\sigma'(x)\right| && \left(\text{by } (10)\right)\\
&= \sum_{i\in[k_2]} \sum_{x\in X_1} \left|\hat\sigma'(x^{(i)})\right|. && \left(\text{by } (11)\right)
\end{aligned}$$

Hence, there exists an $i_1 \in [k_2]$ such that

$$\sum_{x \in X_1} \left|\hat{\sigma}'(x^{(i_1)})\right| \le r_1 k_1, \quad \text{or equivalently} \quad \frac{\sum_{x \in X_1} \left|\hat{\sigma}'(x^{(i_1)})\right|}{|X_1|} \le r_1$$

by $|X_1| = k_1$. Arguing similarly for X_2 we get an $i_2 \in [k_1]$ such that

$$\frac{\sum_{x \in X_2} \left|\hat{\sigma}'(x^{(i_2)})\right|}{|X_2|} \le r_2$$

Finally we define a multi-assignment $\hat{\sigma}$ for the original instance Π by

$$\hat{\sigma}(x) = \begin{cases} \hat{\sigma}'(x^{(i_1)}) & \text{if } x \in X_1 \\ \hat{\sigma}'(x^{(i_2)}) & \text{if } x \in X_2. \end{cases}$$

By the above argument, $\hat{\sigma}$ is (r_1, r_2)-average. Moreover, it satisfies Π, since $\hat{\sigma}'$ satisfies Π'. □

Putting all previous results together, we have Theorem 21.

Proof of Theorem 21. We give an FPT reduction from instances in the Baby PIH (Theorem 18) to Avg-r-Gap-2CSP. Then, since the Baby PIH holds under W[1] $\neq$ FPT, we deduce that the Average Baby PIH also holds under W[1] $\neq$ FPT.

For any 2CSP instance $\Pi_0 = (X_0, \Sigma_0, \Phi_0)$, we can construct a bipartite 2CSP instance $\Pi_1 = (X_1, \Sigma_1, \Phi_1)$ by Lemma 25, and then construct an Avg-r-Gap-2CSP instance $\Pi = (X, \Sigma, \Phi)$ from Π_1 by Lemma 27. Trivially, Π is satisfiable when Π_0 is satisfiable. When Π_0 is not r-list satisfiable, Π_1 is not (r_1, r_2)-average list satisfiable for every constants r_1, r_2 with $r_1 + r_2 \ge 2(1-\varepsilon)r$, and thus Π is not $(1-\varepsilon)r$-average list satisfiable. Furthermore, Π has rectangular relation because Π_1 has rectangular relation.

Moreover, the running time of this reduction can be bounded by

$$f(|X_0| + |\Phi_0| + 1/\varepsilon + r)|\Sigma_0|^{O(1)}$$

for a computable function f, and

$$|X| + |\Phi| \le f(|X_0| + |\Phi_0| + 1/\varepsilon + r)|\Sigma_0|^{O(1)}$$

as well, so the reduction is an FPT-reduction. □

4 From Average Baby PIH to Inapproximability of k-NCP

Since k-NCP and k-MLD are equivalent [LLL24, Appendix A], it suffices to give a gap-preserving reduction from Avg-r-Gap-2CSP to γ-Gap-k''-MLD$_p$ with $k'' = O(|\Phi|)$ and $\gamma < r$ for any constant γ to obtain the inapproximability of k-NCP. First we present a reduction to a "weighted" version of γ-Gap-k-MLD$_p$.

Theorem 28. *For any prime p, there is an* FPT *algorithm $\mathcal{A}_1$ that, given an instance (X, Σ, Φ) of* Avg-r-Gap-2CSP *with rectangular relations and $|X| = k, |\Phi| = k'$, $\mathcal{A}$ outputs a $(k+k')$-MLD$_p$ instance $(V, \vec{t})$ with $V = U \dot\cup W \subseteq \mathbb{F}_p^d$, where $U = U_1 \dot\cup \cdots \dot\cup U_k$ and $W = W_1 \dot\cup \cdots \dot\cup W_{k'}$, such that:*

- $|V| = O(k'|\Sigma|^2)$ *and* $d = 2k'|\Sigma| + k' + k$.
- (Completeness) *If the input is a YES instance, then there exists $\vec{u}_1 \in U_1, \cdots, \vec{u}_k \in U_k, \vec{w}_1 \in W_1, \cdots, \vec{w}_{k'} \in W_{k'}$ such that $\vec{u}_1 + \cdots + \vec{u}_k + \vec{w}_1 + \cdots + \vec{w}_{k'} = \vec{t}$.*
- (Soundness) *If the input is a NO instance, then any solution to $(V, \vec{t})$ must contain at least rk vectors with nonzero coefficients in U.*

Proof. For $j \in [k']$, Let $\varphi_j = (x_{i_1} x_{i_2}, C_j) \in \Phi$. For $i \in [k]$, let $\vec{e}_k(i) \in \mathbb{F}_p^k$ be the one-hot encoding of i, i.e., $\vec{e}_k(i)$ has 1 in the i-th entry and 0 in the rest entries. We omit the subscript $_k$ if it is clear from the context.

We divide the dimension of vectors in the $(k+k')$-MLD$_p$ instance into the following $(k'+1)$ parts:

- The j-th of the first k' parts, which we call the "φ_j part", is to guarantee the satisfaction of the constraint $\varphi_j = (x_{j_1} x_{j_2}, C_j)$ by one-hot encoding x_{j_1} and x_{j_2} in W_j and checking the consistency of each variable in U. This part is further divided into the "$(j,1)$ part" and "$(j,2)$ part", each of length $|\Sigma|$. We use $\vec{v}^{(j)} \in \mathbb{F}_p^{2|\Sigma|}$ to denote the φ_j part, and $\vec{v}^{(j,1)}, \vec{v}^{(j,2)} \in \mathbb{F}_p^{|\Sigma|}$ to denote the $(j,1)$ and $(j,2)$ part of $\vec{v} \in \mathbb{F}_p^d$ respectively.
- The last part of length $k'+k$, which we call "color part", is to ensure that at least one vector is chosen from each U_i and W_j. Let $\vec{v}^{(\text{color})} \in \mathbb{F}_p^{k'+k}$ denote the color part of $\vec{v} \in \mathbb{F}_p^d$.

The structure of the vectors is illustrated in Figure 1.

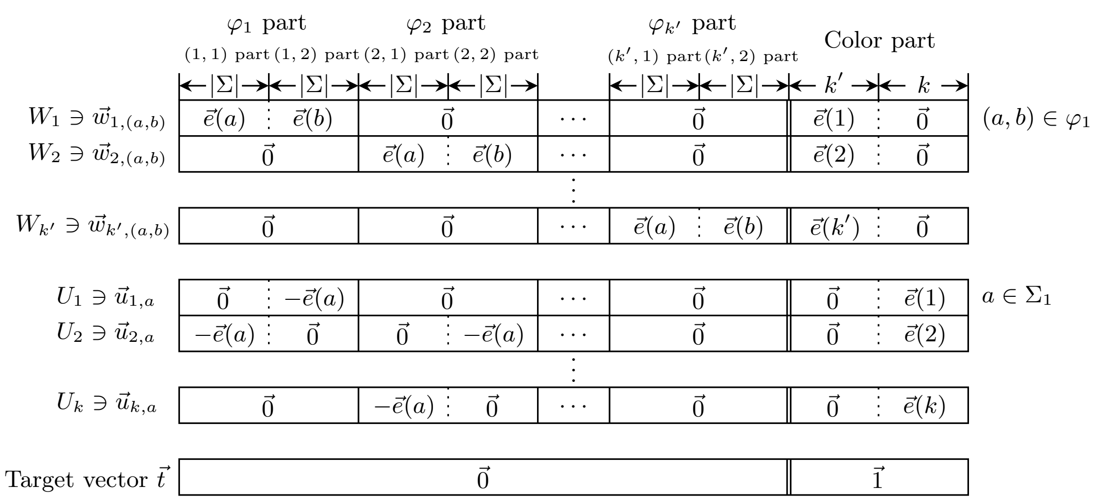

Figure 1: The structure of constructed vectors (assume that $\varphi_1 = (x_1 x_2, C_1)$, $\varphi_2 = (x_2 x_k, C_2)$ for example).

Formally, the algorithm $\mathcal{A}_1$ constructs the $(k+k')$-MLD$_p$ instance $(U \dot\cup W, \vec{t})$ as follows:

- For every $i \in [k]$, let $U_i = \{\vec{u}_{i,a} \mid a \in \Sigma_{x_i}\}$, where
  - For every $j \in [k']$,
$$\vec{u}_{i,a}^{(j,1)} = \begin{cases} -\vec{e}_{|\Sigma|}(a), & \text{if } \varphi_j = (x_i x', C_j), \\ \vec{0}_{|\Sigma|}, & \text{otherwise.} \end{cases}$$
$$\vec{u}_{i,a}^{(j,2)} = \begin{cases} -\vec{e}_{|\Sigma|}(a), & \text{if } \varphi_j = (x' x_i, C_j), \\ \vec{0}_{|\Sigma|}, & \text{otherwise.} \end{cases}$$
  - $\vec{u}_{i,a}^{(\text{color})} = \vec{0} \circ \vec{e}_k(i)$.
- For every $j \in [k']$, $\varphi_j = (x_{j_1} x_{j_2}, C_j)$, let $W_j = \{\vec{w}_{j,(a,b)} \mid (a,b) \in C_j \ (\subseteq \Sigma_{x_{j_1}} \times \Sigma_{x_{j_2}})\}$, where
  - $\vec{w}_{j,(a,b)}^{(j)} = \vec{e}_{|\Sigma|}(a) \circ \vec{e}_{|\Sigma|}(b)$.
  - $\vec{w}_{j,(a,b)}^{(i)} = \vec{0}$ for every $i \neq j$.
  - $\vec{w}_{j,(a,b)}^{(\text{color})} = \vec{e}_{k'}(j) \circ \vec{0}$.
- Let the target vector be $\vec{t} = 0^{2k'|\Sigma|} \circ 1^{k'+k}$.

The algorithm clearly runs in FPT time. The completeness is also obvious: for any satisfying assignment $\sigma : X \to \Sigma$, pick $\vec{u}_{i,\sigma(x_i)} \in U_i$ for every $i \in [k]$ and $\vec{w}_{j,(\sigma(x_{j_1}),\sigma(x_{j_2}))} \in W_j$ for every $j \in [k']$ to obtain a solution to $(U \dot\cup W, \vec{t})$. To see the soundness, we first prove the following claim.

Claim 29. *Let $S \subseteq U \dot\cup W$ (with their coefficients $\lambda : S \to \mathbb{F}_p \setminus \{0\}$) be a solution to $(U \dot\cup W, \vec{t})$. Let $\vec{\alpha} = \sum_{\vec{w} \in S \cap W} \lambda(\vec{w})\vec{w}$. For every $\varphi_j = (x_{j_1} x_{j_2}, C_j) \in \Phi$ with rectangular relation, there exists $(a^*, b^*) \in C_j$ such that $\vec{\alpha}^{(j,1)}[a^*] \neq 0$ and $\vec{\alpha}^{(j,2)}[b^*] \neq 0$.*

Proof of Claim 29. Since φ_j is a rectangular relation, let Q be its underlying set and let π and ρ be the corresponding mappings from $\Sigma_{x_{j_1}}$ and $\Sigma_{x_{j_2}}$ to Q (recall Definition 19). Then, since every satisfying assignment $(a,b) \in C_j$ satisfies $\pi(a) = \rho(b) \in Q$, we can divide all satisfying assignments of φ_j into $|Q|$ disjoint parts according to the element in Q to which assignments a, b are mapped. More precisely, the assignments contained in group $q \in Q$ are $\varphi_j^{(q)} = \pi^{-1}(q) \times \rho^{-1}(q) \subseteq C_j$. Let

$$W_j^{(q)} = \{\vec{w}_{j,(a,b)} \mid (a,b) \in \varphi_j^{(q)}\} \subseteq W_j,$$

since $\{\varphi_j^{(q)}\}_{q \in Q}$ forms a partition of C_j, $\{W_j^{(q)}\}_{q \in Q}$ also forms a partition of W_j.

Since every vector outside W_j has 0 in the j-th entry of the color part, the sum of coefficients to $S \cap W_j$ is exactly the value in the j-th colored parts, i.e.,

$$\sum_{\vec{w} \in S \cap W_j} \lambda(\vec{w}) = \sum_{\vec{w} \in S \cap W_j} \lambda(\vec{w}) \cdot \vec{w}_j^{(\text{color})}[j] = \vec{t}^{(\text{color})}[j] = 1.$$

Meanwhile, the left hand side can be written as

$$\sum_{q \in Q} \sum_{\vec{w} \in S \cap W_j^{(q)}} \lambda(\vec{w}) = 1.$$

Hence there must exist some $q^* \in Q$ such that

$$\sum_{\vec{w} \in S \cap W_j^{(q^*)}} \lambda(\vec{w}) \neq 0.$$

Consider $\vec{\alpha}^{(j,1)}$, the first variable (x_{j_1}) of φ_j part in $\vec{\alpha}$. Recall that entries of $\vec{\alpha}^{(j,1)}$ are naturally indexed by elements in $\Sigma_{x_{j_1}}$, since $\vec{\alpha}^{(j,1)}$ is the linear combination of one-hot encodings, content in the entries corresponds to $\pi^{-1}(q^*)$ is exactly

$$\vec{\alpha}^{(j,1)}|_{\pi^{-1}(q^*)} = \sum_{\vec{w} \in S \cap W_j^{(q^*)}} \lambda(\vec{w}) \vec{w}^{(j,1)}|_{\pi^{-1}(q^*)}$$

(Here we use $\vec{w}^{(j,1)}|_{\pi^{-1}(q^*)}$ to denote the restriction of vector $w^{(j,1)}$ on entries $\pi^{-1}(q^*)$.) Since the sum of coefficients is not zero, and $\vec{w}^{(j,1)}|_{\pi^{-1}(q^*)}$'s are all one-hot vectors, the above sum is not $\vec{0}$. This indicates the existance of some $a^* \in \pi^{-1}(q^*)$ that

$$\vec{\alpha}^{(j,1)}[a^*] \neq 0.$$

Similarly, there exists $b^* \in \rho^{-1}(q^*)$ such that $\vec{\alpha}^{(j,2)}[b^*] \neq \vec{0}$, and $(a^*, b^*) \in C_j$ follows from $\pi(a^*) = \rho(b^*) = q^*$. This finishes the proof of Claim 29. ⊣

Claim 30 (Soundness of Reduction)**.** *Let $S \subseteq U \dot\cup W$ (with their coefficients $\lambda : S \to \mathbb{F}_p \setminus \{0\}$) be a solution to $(U \dot\cup W, \vec{t})$. Then*

$$\hat{\sigma}(x_i) = \{a \in \Sigma_{x_i} \mid \vec{u}_{i,a} \in S \cap U_i\}$$

is a satisfying multi-assignment to the original Avg-r-Gap-2CSP *instance with total value $|S \cap U|$.*

Proof of Claim 30. Following notation in Claim 29, let $\vec{\alpha} = \sum_{\vec{w} \in S \cap W} \lambda(\vec{w}) \vec{w}$, and $\vec{\beta} = \sum_{\vec{u} \in S \cap U} \lambda(\vec{u}) \cdot \vec{u}$. Since S is a solution, For every $\varphi_j = (x_{j_1} x_{j_2}, C_j) \in \Phi$,

$$\vec{t}^{(j,1)}[a^*] = (\vec{\alpha} + \vec{\beta})^{(j,1)}[a^*] = 0.$$

By Claim 29, there must exist $(a^*, b^*) \in C_j$ such that $\vec{\alpha}^{(j,1)}[a^*] \neq 0$ and $\vec{\alpha}^{(j,2)}[b^*] \neq 0$, hence

$$\vec{\beta}^{(j,1)}[a^*] \neq 0, \qquad \vec{\beta}^{(j,2)}[b^*] \neq 0.$$

Observe that $\vec{\beta}^{(j,1)}[a^*] \neq 0$ if and only if $\vec{u}_{j_1,a^*} \in S \cap U$, so $\vec{u}_{j_1,a^*} \in S$. Similarly, $\vec{u}_{j_2,b^*} \in S$. Therefore, $a^* \in \hat{\sigma}(x_{j_1})$ and $b^* \in \hat{\sigma}(x_{j_2})$, thus φ_j is satisfied by $\hat{\sigma}$. The total value of $\hat{\sigma}$ is

$$\sum_{x \in X} |\hat{\sigma}(x)| = \sum_{1 \le i \le k} |\{a \in \Sigma_{x_i} \mid \vec{u}_{i,a} \in S \cap U_i\}| = \sum_{1 \le i \le k} |S \cap U_i| = |S \cap U|.$$

⊣

From Claim 30, if the input 2CSP instance is not r-average-list satisfiable, then $|S \cap U| > rk$. Therefore, the soundness of Theorem 28 is shown. □

Theorem 28 gives a gap that only occurs in one side. We apply the construction in [LLL24, Theorem 20] as follows.

Lemma 31 (cf. Theorem 20 in [LLL24])**.** *There is an* FPT *algorithm* $\mathcal{A}_2$ *that, for any constants* $r > \gamma > 1$*,* $\mathcal{A}_2$ *take an instance of* $(k+k')$-MLD$_p$ *instance* $(V, \vec{t})$ *with the following properties as input:*

- $V = U \dot\cup W \subseteq \mathbb{F}_p^d$*,*
- *either there exists* k *vectors in* U *and* k' *vectors in* W *whose sum being* $\vec{t}$*,*
- *or any solution to* $(V, \vec{t})$ *must contain at least* rk *vectors in* U*.*

$\mathcal{A}_2$ *output a new* γ-Gap-k''-MLD$_p$ *instance* $(V', \vec{t'})$ *satisfying:*

- $V' = U' \dot\cup W' \subseteq \mathbb{F}_p^{d'}$*,*
- $k'' = O(k+k')$*,* $d' = \lceil \frac{\gamma}{r-\gamma} \cdot \frac{k'}{k} \rceil d$*,*

Proof Sketch. Let $\ell = \lceil \frac{\gamma}{r-\gamma} \cdot \frac{k'}{k} \rceil$ and $k'' = \lceil \frac{r}{r-\gamma} \rceil k'$. $\mathcal{A}_2$ produces vector set $W' \subseteq \mathbb{F}_p^{\ell d}$ and target $\vec{t'} \in \mathbb{F}_p^{\ell d}$ as

$$W' = \{\underbrace{\vec{w} \circ \vec{w} \circ \cdots \circ \vec{w}}_{\ell \text{ times}} : \vec{w} \in W\}, \quad \vec{t'} = \underbrace{\vec{t} \circ \vec{t} \circ \cdots \circ \vec{t}}_{\ell \text{ times}}.$$

For each $\vec{u} \in U (\subseteq \mathbb{F}_p^d)$, $\mathcal{A}_2$ produces ℓ "copies" $\vec{u}_1, \cdots, \vec{u}_\ell \in \mathbb{F}_p^{\ell d}$ as

$$\vec{u}_i = 0^{(i-1)d} \circ \vec{u} \circ 0^{(\ell-i)d}, \text{ for } 1 \le i \le \ell.$$

And the set U' collects all of these "copies" as

$$U' = \dot\bigcup_{\vec{u} \in U} \{u_1, \cdots, \vec{u}_\ell\}.$$

Finally, $V' = U' \dot\cup W'$. The completeness is clear from the construction. For the soundness. It is not hard to see (cf. [LLL24, Theorem 20]) that any solution to $(V', \vec{t'})$ need to contain at least rk vectors in each "copy" of U, which means in total at least $r\ell k$ vectors in U'. Hence the solution must have size at least

$$r\ell k + k' = \lceil \frac{r\gamma}{r-\gamma} + 1 \rceil k' > \lceil \frac{r\gamma}{r-\gamma} \rceil k' = (1 - o(1))\gamma \cdot k''.$$

The reduction clearly runs in FPT time. □

Putting all things together, we have:

Theorem 32. *For any prime* p*, any constants* $r > \gamma > 1$*, there is an* FPT *reduction from* Avg-r-Gap-2CSP *instance* $\Pi = (X, \Sigma, \Phi)$ *to* γ-Gap-k'-MLD$_p$ *instance* $(V, \vec{t})$ *that satisfies* $k' = \lceil \frac{r}{r-\gamma} \rceil |\Phi|$*.*

Proof. Combine the reductions in Theorem 28 and Lemma 31. □

Acknowledgements

The authors want to thank Guohang Liu, Mingjun Liu, Yangluo Zheng for their discussion in the early stage of this work.

References

[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. 1

[AMS06] Noga Alon, Dana Moshkovitz, and Shmuel Safra. Algorithmic construction of sets for k-restrictions. *ACM Trans. Algorithms*, 2(2):153–177, 2006. 1, 2, 23

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

[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. 3

[BGKM18] Arnab Bhattacharyya, Suprovat Ghoshal, Karthik C. S., and Pasin Manurangsi. Parameterized intractability of even set and shortest vector problem from Gap-ETH. In Ioannis Chatzigiannakis, Christos Kaklamanis, Dániel Marx, and Donald Sannella, editors, *45th International Colloquium on Automata, Languages, and Programming, ICALP 2018, July 9-13, 2018, Prague, Czech Republic*, volume 107 of *LIPIcs*, pages 17:1–17:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018. 3, 4

[BK22] Libor Barto and Marcin Kozik. Combinatorial gap theorem and reductions between promise CSPs. In Joseph (Seffi) Naor and Niv Buchbinder, editors, *Proceedings of the 2022 ACM-SIAM Symposium on Discrete Algorithms, SODA 2022, Virtual Conference / Alexandria, VA, USA, January 9 - 12, 2022*, pages 1204–1220. SIAM, 2022. 2

[CFK+15] Marek Cygan, Fedor V. Fomin, Lukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michal Pilipczuk, and Saket Saurabh. *Parameterized Algorithms*. Springer, 2015. 7

[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. 6

[CL19] Yijia Chen and Bingkai Lin. The constant inapproximability of the parameterized dominating set problem. *SIAM J. Comput.*, 48(2):513–533, 2019. 3, 6

[DF13] Rodney G. Downey and Michael R. Fellows. *Fundamentals of Parameterized Complexity*. Texts in Computer Science. Springer, 2013. 7

[Din07] Irit Dinur. The PCP theorem by gap amplification. *J. ACM*, 54(3):12, 2007. 1

[FG06] Jörg Flum and Martin Grohe. *Parameterized Complexity Theory*. Texts in Theoretical Computer Science. An EATCS Series. Springer, 2006. 7

[GLR+24a] Venkatesan Guruswami, Bingkai Lin, Xuandi Ren, Yican Sun, and Kewen Wu. Almost optimal time lower bound for approximating parameterized clique, CSP, and more, under ETH. *CoRR*, abs/2404.08870, 2024. 2, 3

[GLR+24b] Venkatesan Guruswami, Bingkai Lin, Xuandi Ren, Yican Sun, and Kewen Wu. Parameterized inapproximability hypothesis under exponential time hypothesis. In Bojan Mohar, Igor Shinkar, and Ryan O'Donnell, editors, *Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC 2024, Vancouver, BC, Canada, June 24-28, 2024*, pages 24–35. ACM, 2024. 2

[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. 2, 3, 4, 6, 9, 23

[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. 6

[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. 3

[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. 6, 10

[Lin19] Bingkai Lin. A simple gap-producing reduction for the parameterized set cover problem. In Christel Baier, Ioannis Chatzigiannakis, Paola Flocchini, and Stefano Leonardi, editors, *46th International Colloquium on Automata, Languages, and Programming, ICALP 2019, July 9-12, 2019, Patras, Greece*, volume 132 of *LIPIcs*, pages 81:1–81:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. 3

[Lin21] Bingkai Lin. Constant approximating k-clique is W[1]-hard. In Samir Khuller and Virginia Vassilevska Williams, editors, *STOC '21: 53rd Annual ACM SIGACT Symposium on Theory of Computing, Virtual Event, Italy, June 21-25, 2021*, pages 1749–1756. ACM, 2021. 6

[LLL24] Shuangle Li, Bingkai Lin, and Yuwei Liu. Improved lower bounds for approximating parameterized nearest codeword and related problems under ETH. In Karl Bringmann, Martin Grohe, Gabriele Puppis, and Ola Svensson, editors, *51st International Colloquium on Automata, Languages, and Programming, ICALP 2024, July 8-12, 2024,* *Tallinn, Estonia*, volume 297 of *LIPIcs*, pages 107:1–107:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2024. 3, 4, 5, 6, 8, 9, 16, 20

[LRSW23] 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. 3, 6, 10

[LRSZ20] Daniel Lokshtanov, M. S. Ramanujan, Saket Saurabh, and Meirav Zehavi. Parameterized complexity and approximability of directed odd cycle transversal. 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 2181–2200. SIAM, 2020. 2, 9

[LY94] Carsten Lund and Mihalis Yannakakis. On the hardness of approximating minimization problems. *J. ACM*, 41(5):960–981, 1994. 23

[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. 3, 4

A From Average Baby PIH to Inapproximability of k-ExactCover

We present a proof relying on a construction that slightly differs from the one in [GRS24]. Their proof make use of the (T, m)-set gadget [LY94, AMS06] that previously used to show hardness of approximating SetCover problem. On the other hand, our proof develops a novel composition of collision number of ECCs (recall Definition 22) with the following well-known combinatorial object.

Definition 33 (Hypercube Partition System)**.** *Let A, B be two sets. Then the (A, B)-*hypercube partition system *is defined by*

- *the universe $\mathcal{M} = A^B \left(= \{z \mid \text{a function } z : B \to A\}\right)$, and*
- *a collection of subsets $\{P_{x,y}\}_{x \in B, y \in A}$ where each $P_{x,y} = \{z \in \mathcal{M} \mid z(x) = y\}$.*

Theorem 34 (cf. Theorem 21 in [GRS24])**.** *Assume that the Average Baby PIH holds on all 2CSP instances with rectangular relations. Then k-ExactCover cannot be approximated in* FPT *time within any constant factor. More precisely, for every constant $r > 1$ no* FPT *algorithm, on a given k-SetCover instance $\Pi = (S, U)$ with size n and $k \geq 1$, can distinguish between the following two cases:*

- *We can choose k* disjoint *sets in S whose union is U.*
- *U is not the union of any rk sets in S.*

Proof. Let $\Pi = (X, \Sigma, \Phi)$ be an AVG-r-GAP-2CSP instance with rectangular relations. We set $k = |X|$. Moreover, for each rectangular constraint $\varphi_j = (x_{i_1} x_{i_2}, C_j) \in \Phi$ we use Q_j to denote the underlying set and $\pi_j, \rho_j : \Sigma \to Q_j$ the associated mappings as in Definition 19. That is, for every $a, b \in \Sigma$, it holds that $(a, b) \in C_j$ if and only if $\pi_j(a) = \rho_j(b)$. Then we set

$$t = \max_{\varphi_j \in \Phi} |Q_j|. \tag{12}$$

Clearly, we can assume without loss of generality

$$t \leq |\Pi|.$$

Now we reduce Π to a k-EXACTCOVER instance. To that end, we choose a further alphabet Δ whose size is a prime and satifies

$$\max\left\{\lceil \log t \rceil, 2^{2^{r^2 k^2}}\right\} \leq |\Delta| \leq 2 \max\left\{\lceil \log t \rceil, 2^{2^{r^2 k^2}}\right\}.$$

Moreover,let

$$d = \left\lceil \frac{2r^2 k^2 \log t}{\log |\Delta|} \right\rceil.$$

This leads to a Reed-Solomon code

$$\text{Enc} : \Delta^{\left\lceil \frac{\log t}{\log |\Delta|} \right\rceil} \to \Delta^d.$$

Plugging

$$k \leftarrow \left\lfloor \frac{\log t}{\log |\Delta|} \right\rfloor, \ m \leftarrow d, \ p \leftarrow |\Delta|,[3] \text{ and } \varepsilon \leftarrow 1/2$$

in Theorem 23 we conclude that the 1/2-collision number of Enc is

$$\text{Col}_{1/2}(\text{Enc}) \geq \sqrt{\frac{d}{\log t / \log |\Delta|}} > rk.$$

Observe that (12) implies that every tuple in C_i can be identified with a string in $\Delta^{\left\lceil \frac{\log m}{\log |\Delta|} \right\rceil}$, i.e., the domain of Enc.

Then, for each variable $x \in X$ and every its possible value $a \in \Sigma$, we define a set $S_{x,a}$ as follows. For each constraint $\varphi_j = (x_{i_1} x_{i_2}, C_j) \in \Phi$ with associated set T_j and mappings $\pi_j, \rho_j : \Sigma \to Q_j$, and for each $\ell \in [d]$, we construct a $([2], \Delta)$-hypercube partition system

$$\left(\mathcal{M}^{(j,\ell)}, \{P_{u,v}^{(j,\ell)}\}_{u \in \Delta, v \in [2]}\right). \tag{13}$$

[3] Observe that

$$\left\lfloor \frac{\log t}{\log |\Delta|} \right\rfloor < \left\lceil \frac{2r^2 k^2 \log t}{\log |\Delta|} \right\rceil \leq \lceil \log t \rceil \leq |\Delta|,$$

hence, the condition $k < m \leq p$ in Theorem 23 is satisfied.

Then for each $(a,b) \in C_j$ we add $P^{(j,\ell)}_{\mathrm{Enc}(\pi_j(a))[\ell],1}$ to $S_{x_{i_1},a}$ and similarly $P^{(j,\ell)}_{\mathrm{Enc}(\rho_j(b))[\ell],2}$ to $S_{x_{i_2},b}$. Finally, let the universe be

$$U = \dot{\bigcup}_{\varphi_j \in \Phi, \ell \in [d]} \mathcal{M}^{(j,\ell)}, \quad \text{and} \quad \mathcal{S} = \{ S_{x,a} \mid x \in X \text{ and } a \in \Sigma \}.$$

For the completeness, let $\sigma : X \to \Sigma$ be a satisfying assignment of Π, it is routine to check that $\{S_{x,\sigma(x)}\}_{x \in X}$ is a partition of U.

For the soundness, assume that every satisfying multi-assignment of Π has total value at least rk (cf. Definition 11). Let $\mathcal{S}' \subseteq \mathcal{S}$ be a cover of U. Consider the multi-assignment that maps every variable $x \in X$ to $\{ a \in \Sigma \mid S_{x,a} \in \mathcal{S}' \}$. If this multi-assignment satisfies Π, the our assumption implies $|\mathcal{S}'| \geq rk$. Otherwise, assume that there exists some constraint $\varphi_j = (x_{i_1} x_{i_2}, C_j) \in \Phi$ which is not satisfied. Note that the above multi-assignment assigns x_{i_1} to $E_1 = \{ a \in \Sigma \mid S_{x_{i_1},a} \in \mathcal{S}' \}$ and x_{i_2} to $E_2 = \{ b \in \Sigma \mid S_{x_{i_2},b} \in \mathcal{S}' \}$. Since φ_j is not satisfied, for all $(a,b) \in E_1 \times E_2$ we have $\mathrm{Enc}(\pi_j(a)) \neq \mathrm{Enc}(\rho_j(b))$. However, for each $\ell \in [d]$, since $\mathcal{M}^{(j,\ell)}$ is covered by $\mathcal{S}'$, there must exist $a \in E_1$ and $b \in E_2$ with $\mathrm{Enc}(\pi_j(a))[\ell] = \mathrm{Enc}(\rho_j(b))[\ell]$. Therefore, the set $\{\pi_j(a)\}_{a \in E_1} \cup \{\rho_j(b)\}_{b \in E_2}$ collides on all coordinates $\ell \in [d]$, hence it must have size at least $\mathrm{Col}_{1/2}(\mathrm{Enc})$. We deduce

$$|\mathcal{S}'| \geq |E_1| + |E_2| \geq \left| \{\pi_j(a)\}_{a \in E_1} \cup \{\rho_j(b)\}_{b \in E_2} \right| \geq \mathrm{Col}_{1/2}(\mathrm{Enc}) > rk.$$

Finally, in each hypercube partition system (13) it holds that

$$\left| \mathcal{M}^{(j,\ell)} \right| = 2^{|\Delta|} \leq 4^{\lceil \log t \rceil} + 4^{2^{2r^2k^2}} \leq |\Pi|^2 + 4^{2^{2r^2k^2}},$$

and there are at most $\binom{k}{2} d \leq k^2 r^2 k^2 \log t \leq r^2 k^4 \log |\Pi|$ such systems. The size of the universe U is thus at most $g(r,k)|\Pi|^3$ for some appropriate computable function $g : \mathbb{N}^2 \to \mathbb{N}$, while the parameter of the k-ExactCover instance remains $k = |X|$. It follows easily that the running time of this reduction is FPT. □

Combining Theorem 34 and Theorem 21, we obtain:

Theorem 35. *For any constant $r > 1$, r-approximating k-ExactCover is* W[1]*-hard.*