

# Parameterized Intractability of Even Set and Shortest Vector Problem from Gap-ETH

Arnab Bhattacharyya\*    Suprovat Ghoshal†    Karthik C. S.‡    Pasin Manurangsi§

March 26, 2018

## Abstract

The *k-Even Set* problem is a parameterized variant of the *Minimum Distance Problem* of linear codes over  $\mathbb{F}_2$ , which can be stated as follows: given a generator matrix  $\mathbf{A}$  and an integer  $k$ , determine whether the code generated by  $\mathbf{A}$  has distance at most  $k$ . Here,  $k$  is the parameter of the problem. The question of whether *k-Even Set* is fixed parameter tractable (FPT) has been repeatedly raised in literature and has earned its place in Downey and Fellows' book (2013) as one of the "most infamous" open problems in the field of Parameterized Complexity.

In this work, we show that *k-Even Set* does not admit FPT algorithms under the (randomized) Gap Exponential Time Hypothesis (Gap-ETH) [Dinur'16, Manurangsi-Raghavendra'16]. In fact, our result rules out not only exact FPT algorithms, but also any constant factor FPT approximation algorithms for the problem. Furthermore, our result holds even under the following weaker assumption, which is also known as the *Parameterized Inapproximability Hypothesis* (PIH) [Lokshtanov et al.'17]: no (randomized) FPT algorithm can distinguish a satisfiable 2CSP instance from one which is only 0.99-satisfiable (where the parameter is the number of variables).

We also consider the parameterized *k-Shortest Vector Problem* (SVP), in which we are given a lattice whose basis vectors are integral and an integer  $k$ , and the goal is to determine whether the norm of the shortest vector (in the  $\ell_p$  norm for some fixed  $p$ ) is at most  $k$ . Similar to *k-Even Set*, this problem is also a long-standing open problem in the field of Parameterized Complexity. We show that, for any  $p > 1$ , *k-SVP* is hard to approximate (in FPT time) to some constant factor, assuming PIH. Furthermore, for the case of  $p = 2$ , the inapproximability factor can be amplified to any constant.

**Keywords:** Parameterized Complexity, Inapproximability, Even Set, Minimum Distance Problem, Shortest Vector Problem

---

\*Indian Institute of Science, India. Email: [arnabb@iisc.ac.in](mailto:arnabb@iisc.ac.in). This work was partially supported by Ramanujan Fellowship DSTO 1358.

†Indian Institute of Science, India. Email: [suprovat@iisc.ac.in](mailto:suprovat@iisc.ac.in).

‡Weizmann Institute of Science, Israel. Email: [karthik.srikanta@weizmann.ac.il](mailto:karthik.srikanta@weizmann.ac.il). This work was supported by Irit Dinur's ERC-CoG grant 772839.

§University of California, Berkeley, USA. Email: [pasin@berkeley.edu](mailto:pasin@berkeley.edu). Some part of this work was done while the author was visiting Indian Institute of Science and supported by the Indo-US Joint Center for Pseudorandomness in Computer Science.

# 1 Introduction

The study of error-correcting codes gives rise to many interesting computational problems. One of the most fundamental among these is the problem of computing the distance of a linear code. In this problem, which is commonly referred to as the *Minimum Distance Problem (MDP)*, we are given as input a generator matrix  $\mathbf{A} \in \mathbb{F}_2^{m \times m}$  of a binary<sup>1</sup> linear code and an integer  $k$ . The goal is to determine whether the code has distance at most  $k$ . Recall that the distance of a linear code is  $\min_{\mathbf{0} \neq \mathbf{x} \in \mathbb{F}_2^m} \|\mathbf{A}\mathbf{x}\|_0$  where  $\|\cdot\|_0$  denote the 0-norm (aka the Hamming norm).

The study of this problem dates back to at least 1978 when Berlekamp et al. [BMvT78] conjectured that it is NP-hard. This conjecture remained open for almost two decades until it was positively resolved by Vardy [Var97a, Var97b]. Later, Dumer et al. [DMS03] strengthened this intractability result by showed that, even *approximately* computing the minimum distance of the code is hard. Specifically, they showed that, unless  $\text{NP} = \text{RP}$ , no polynomial time algorithm can distinguish between a code with distance at most  $k$  and one whose distance is greater than  $\gamma \cdot k$  for any constant  $\gamma \geq 1$ . Furthermore, under stronger assumptions, the ratio can be improved to superconstants and even almost polynomial. Dumer et al.'s result has been subsequently derandomized by Cheng and Wan [CW12] and further simplified by Austrin and Khot [AK14] and Micciancio [Mic14].

While the aforementioned intractability results rule out not only efficient algorithms but also efficient approximation algorithms for MDP, there is another popular technique in coping with NP-hardness of problems which is not yet ruled out by the known results: *parameterization*.

In parameterized problems, part of the input is an integer that is designated as the parameter of the problem, and the goal is now not to find a polynomial time algorithm but a *fixed parameter tractable (FPT)* algorithm. This is an algorithm whose running time can be upper bounded by some (computable) function of the parameter in addition to some polynomial in the input length. Specifically, for MDP, its parameterized variant<sup>2</sup>  $k$ -MDP has  $k$  as the parameter and the question is whether there exists an algorithm that can decide if the code generated by  $\mathbf{A}$  has distance at most  $k$  in time  $T(k) \cdot \text{poly}(mn)$  where  $T$  can be any computable function that depends only on  $k$ .

The parameterized complexity of  $k$ -MDP was first questioned by Downey et al. [DFVW99]<sup>3,4</sup> who showed that parameterized variants of several other coding-theoretic problems, including the Nearest Codeword Problem and the Nearest Vector Problem<sup>5</sup> which we will discuss in more details in Section 1.1.1, are  $W[1]$ -hard. Thereby, assuming the widely believed  $W[1] \neq \text{FPT}$  hypothesis, these problems are rendered intractable from the parameterized perspective. Unfortunately, Downey et al. fell short of proving such hardness for  $k$ -MDP and left it as an open problem:

**Open Question 1.1** *Is  $k$ -MDP fixed parameter tractable?*

Although almost two decades have passed, the above question remains unresolved to this day, de-

---

<sup>1</sup>Note that MDP can be defined over larger fields as well; we discuss more about this in Section 7.

<sup>2</sup>Throughout Sections 1 and 2, for a computational problem  $\Pi$ , we denote its parameterized variant by  $k$ - $\Pi$ , where  $k$  is the parameter of the problem.

<sup>3</sup> $k$ -MDP is formulated slightly differently in [DFVW99]. There, the input contains a parity-check matrix instead of the generator matrix, but, since we can efficiently compute one given the other, the two formulations are equivalent.

<sup>4</sup> $k$ -MDP is commonly referred to in the area of parameterized complexity as the  *$k$ -Even Set* problem due to its graph theoretic interpretation (see [DFVW99]).

<sup>5</sup>The Nearest Vector Problem is also referred to in the literature as the Closest Vector Problem.

spite receiving significant attention from the community. In particular, the problem was listed as an open question in the seminal book of Downey and Fellows [DF99] and has been reiterated numerous times over the years [DGMS07, FGMS12, GKS12, DF13, CFJ<sup>+</sup>14, CFK<sup>+</sup>15, BGG16, CFHW17, Maj17]. In fact, in their second book [DF13], Downey and Fellows even include this problem as one of the six<sup>6</sup> “most infamous” open questions in the area of Parameterized Complexity.

Another question posted in Downey et al.’s work [DFVW99] that remains open is the parameterized *Shortest Vector Problem* ( $k$ -SVP) in lattices. The input of  $k$ -SVP (in the  $\ell_p$  norm) is an integer  $k \in \mathbb{N}$  and a matrix  $\mathbf{A} \in \mathbb{Z}^{n \times m}$  representing the basis of a lattice, and we want to determine whether the shortest (non-zero) vector in the lattice has length at most  $k$ , i.e., whether  $\min_{\mathbf{0} \neq \mathbf{x} \in \mathbb{Z}^m} \|\mathbf{A}\mathbf{x}\|_p \leq k$ .

Again,  $k$  is the parameter of the problem. It should also be noted here that, similar to [DFVW99], we require the basis of the lattice to be integer-value, which is sometimes not enforced in literature (e.g. [vEB81, Ajt98]). This is because, if  $\mathbf{A}$  is allowed to be any matrix in  $\mathbb{R}^{n \times m}$ , then parameterization is meaningless because we can simply scale  $\mathbf{A}$  down by a large multiplicative factor.

The (non-parameterized) Shortest Vector Problem (SVP) has been intensively studied, motivated partly due to the fact that both algorithms and hardness results for the problem have numerous applications. Specifically, the celebrated LLL algorithm for SVP [LLL82] can be used to factor rational polynomials, and to solve integer programming (parameterized by the number of unknowns) [Len83] and many other computational number-theoretic problems (see e.g. [NV10]). Furthermore, the hardness of (approximating) SVP has been used as the basis of several cryptographic constructions [Ajt98, AD97, Reg03, Reg05]. Since these topics are out of scope of our paper, we refer the interested readers to the following surveys for more details: [Reg06, MR09, NV10, Reg10].

On the computational hardness side of the problem, van Emde-Boas [vEB81] was the first to show that SVP is NP-hard for the  $\ell_\infty$  norm, but left open the question of whether SVP on the  $\ell_p$  norm for  $1 \leq p < \infty$  is NP-hard. It was not until a decade and a half later that Ajtai [Ajt96] showed, under a randomized reduction, that SVP for the  $\ell_2$  norm is also NP-hard; in fact, Ajtai’s hardness result holds not only for exact algorithms but also for  $(1 + o(1))$ -approximation algorithms as well. The  $o(1)$  term in the inapproximability ratio was then improved in a subsequent work of Cai and Nerurkar [CN99]. Finally, Micciancio [Mic00] managed to achieve a factor that is bounded away from one. Specifically, Micciancio [Mic00] showed (again under randomized reductions) that SVP on the  $\ell_p$  norm is NP-hard to approximate to within a factor of  $\sqrt[p]{2}$  for every  $1 \leq p < \infty$ . Khot [Kho05] later improved the ratio to any constant, and even to  $2^{\log^{1/2-\epsilon}(nm)}$  under a stronger assumption. Haviv and Regev [HR07] subsequently simplified the gap amplification step of Khot and, in the process, improved the ratio to almost polynomial. We note that both Khot’s and Haviv-Regev reductions are also randomized and it is still open to find a deterministic NP-hardness reduction for SVP in the  $\ell_p$  norms for  $1 \leq p < \infty$  (see [Mic12]); we emphasize here that such a reduction is not known even for the exact (not approximate) version of the problem. For the  $\ell_\infty$  norm, the following stronger result due to Dinur is known [Din02]: SVP in the  $\ell_\infty$  norm is NP-hard to approximate to within  $n^{\Omega(1/\log \log n)}$  factor (under a *deterministic* reduction).

Very recently, fine-grained studies of SVP have been initiated [BGS17, AS18]. The authors of [BGS17, AS18] showed that SVP for any  $\ell_p$  norm cannot be solved (or even approximated to some constant

---

<sup>6</sup>So far, two of the six problems have been resolved: that of parameterized complexity of  $k$ -Biclique [Lin15] and that of parameterized approximability of  $k$ -Dominating Set [KLM18].

strictly greater than one) in subexponential time assuming the existence of a certain family of lattices<sup>7</sup> and the (randomized) *Gap Exponential Time Hypothesis (Gap-ETH)* [Din16, MR16], which states that no randomized subexponential time algorithm can distinguish between a satisfiable 3CNF formula and one which is only 0.99-satisfiable. (See Hypothesis 3.10.)

As with MDP, Downey et al. [DFVW99] were the first to question the parameterized tractability of  $k$ -SVP (for the  $\ell_2$  norm). Once again, Downey and Fellows included  $k$ -SVP as one of the open problems in both of their books [DF99, DF13], albeit, in their second book,  $k$ -SVP was in the “tough customers” list instead of the “most infamous” list that  $k$ -MDP belonged to. And again, as with Open Question 1.1, this question remains unresolved to this day:

**Open Question 1.2** *Is  $k$ -SVP fixed parameter tractable?*

## 1.1 Our Results

The main result of this paper is a resolution to the previously mentioned Open Question 1.1 and 1.2: more specifically, we prove that  $k$ -MDP and  $k$ -SVP (on  $\ell_p$  norm for any  $p > 1$ ) do not admit any FPT algorithm, assuming the aforementioned (randomized) Gap-ETH (Hypothesis 3.10). In fact, our result is slightly stronger than stated here in a couple of ways:

- (1) First, our result rules out not only exact FPT algorithms but also FPT approximation algorithms as well.
- (2) Second, our result works even under the so-called *Parameterized Inapproximability Hypothesis (PIH)* [LRSZ17], which asserts that no (randomized) FPT algorithm can distinguish between a satisfiable 2CSP instance and one which is only 0.99-satisfiable, where the parameter is the number of variables (See Hypothesis 3.9). It is known (and simple to see) that Gap-ETH implies PIH; please refer to Section 3.4, for more details regarding the two assumptions.

With this in mind, we can state our results starting with the parameterized intractability of  $k$ -MDP, more concretely (but still informally), as follows:

**Theorem 1.3 (Informal; see Theorem 5.1)** *Assuming PIH, for any  $\gamma \geq 1$  and any computable function  $T$ , no  $T(k) \cdot \text{poly}(nm)$ -time algorithm, on input  $(\mathbf{A}, k) \in \mathbb{F}_2^{n \times m} \times \mathbb{N}$ , can distinguish between*

- *the distance of the code generated by  $\mathbf{A}$  is at most  $k$  (i.e.,  $\min_{\mathbf{0} \neq \mathbf{x} \in \mathbb{F}_2^m} \|\mathbf{A}\mathbf{x}\|_0 \leq k$ ), and,*
- *the distance of the code generated by  $\mathbf{A}$  is more than  $\gamma \cdot k$  (i.e.,  $\min_{\mathbf{0} \neq \mathbf{x} \in \mathbb{F}_2^m} \|\mathbf{A}\mathbf{x}\|_0 > \gamma \cdot k$ ).*

Notice that our above result rules out FPT approximation algorithms with *any* constant approximation ratio for  $k$ -MDP. In contrast, we can only prove FPT inapproximability with *some* constant ratio for  $k$ -SVP in  $\ell_p$  norm for  $p > 1$ , with the exception of  $p = 2$  for which the inapproximability factor in our result can be amplified to any constant. These are stated more precisely below.

**Theorem 1.4 (Informal; see Theorem 6.1)** *For any  $p > 1$ , there exists a constant  $\gamma_p > 1$  such that, assuming PIH, for any computable function  $T$ , no  $T(k) \cdot \text{poly}(nm)$ -time algorithm, on input  $(\mathbf{A}, k) \in \mathbb{Z}^{n \times m} \times \mathbb{N}$ , can distinguish between*

---

<sup>7</sup>This additional assumption is only needed for  $1 \leq p \leq 2$ . For  $p > 2$ , their hardness is conditional only on Gap-ETH.

- the  $\ell_p$  norm of the shortest vector of the lattice generated by  $\mathbf{A}$  is  $\leq k$  (i.e.,  $\min_{\mathbf{0} \neq \mathbf{x} \in \mathbb{Z}^m} \|\mathbf{Ax}\|_p \leq k$ ), and,
- the  $\ell_p$  norm of the shortest vector of the lattice generated by  $\mathbf{A}$  is  $> \gamma_p \cdot k$  (i.e.,  $\min_{\mathbf{0} \neq \mathbf{x} \in \mathbb{Z}^m} \|\mathbf{Ax}\|_p > \gamma_p \cdot k$ ).

**Theorem 1.5 (Informal; see Theorem 6.2)** Assuming PIH, for any computable function  $T$  and constant  $\gamma \geq 1$ , no  $T(k) \cdot \text{poly}(nm)$ -time algorithm, on input  $(\mathbf{A}, k) \in \mathbb{Z}^{n \times m} \times \mathbb{N}$ , can distinguish between

- the  $\ell_2$  norm of the shortest vector of the lattice generated by  $\mathbf{A}$  is  $\leq k$  (i.e.,  $\min_{\mathbf{0} \neq \mathbf{x} \in \mathbb{Z}^m} \|\mathbf{Ax}\|_2 \leq k$ ), and,
- the  $\ell_2$  norm of the shortest vector of the lattice generated by  $\mathbf{A}$  is  $> \gamma \cdot k$  (i.e.,  $\min_{\mathbf{0} \neq \mathbf{x} \in \mathbb{Z}^m} \|\mathbf{Ax}\|_2 > \gamma \cdot k$ ).

We remark that our results do not yield hardness for SVP in the  $\ell_1$  norm and this remains an interesting open question. Section 7 contains discussion on this problem. We also note that, for Theorem 6.1 and onwards, we are only concerned with  $p \neq \infty$ ; this is because, for  $p = \infty$ , the problem is NP-hard to approximate even when  $k = 1$  [vEB81]!

### 1.1.1 Nearest Codeword Problem and Nearest Vector Problem

As we shall see in Section 2, our proof proceeds by first showing FPT hardness of approximation of the non-homogeneous variants of  $k$ -MDP and  $k$ -SVP called the  $k$ -Nearest Codeword Problem ( $k$ -NCP) and the  $k$ -Nearest Vector Problem ( $k$ -NVP) respectively. For both  $k$ -NCP and  $k$ -NVP, we are given a target vector  $\mathbf{y}$  (in  $\mathbb{F}_2^n$  and  $\mathbb{Z}^n$ , respectively) in addition to  $(\mathbf{A}, k)$ , and the goal is to find whether there is any  $\mathbf{x}$  (in  $\mathbb{F}_2^m$  and  $\mathbb{Z}^m$ , respectively) such that the (Hamming and  $\ell_p$ , respectively) norm of  $\mathbf{Ax} - \mathbf{y}$  is at most  $k$ .

As an intermediate step of our proof, we show that the  $k$ -NCP and  $k$ -NVP problems are hard to approximate<sup>8</sup> (see Theorem 4.1 and Theorem 6.3 respectively). This should be compared to [DFVW99], in which the authors show that both problems are  $W[1]$ -hard. The distinction here is that our result rules out not only exact algorithms but also approximation algorithms, at the expense of the stronger assumption than that of [DFVW99]. Indeed, if one could somehow show that  $k$ -NCP and  $k$ -NVP are  $W[1]$ -hard to approximate (to some constant strictly greater than one), then our reduction would imply  $W[1]$ -hardness of  $k$ -MDP and  $k$ -SVP (under randomized reductions). Unfortunately, no such  $W[1]$ -hardness of approximation of  $k$ -NCP and  $k$ -NVP is known yet.

We end this section by remarking that the computational complexity of both (non-parameterized) NCP and NVP are also thoroughly studied (see e.g. [Mic01, DKRS03, Ste93, ABSS97, GMSS99] in addition to the references for MDP and SVP), and indeed the inapproximability results of these two problems form the basis of hardness of approximation for MDP and SVP.

## 1.2 Organization of the paper

In the next section, we give an overview of our reductions and proofs. After that, in Section 3, we define additional notations and preliminaries needed to fully formalize our proofs. In Section 4 we show the constant inapproximability of  $k$ -NCP. Next, in Section 5, we establish the constant

<sup>8</sup>While our  $k$ -MDP result only applies for  $\mathbb{F}_2$ , it is not hard to see that our intermediate reduction for  $k$ -NCP actually applies for any finite field  $\mathbb{F}_q$  too.

inapproximability of  $k$ -MDP. Section 6 provides the constant inapproximability of  $k$ -NVP and  $k$ -SVP. Finally, in Section 7, we conclude with a few open questions and research directions.

## 2 Proof Overview

In the non-parameterized setting, all the aforementioned inapproximability results for both MDP and SVP are shown in two steps: first, one proves the inapproximability of their inhomogeneous counterparts (i.e. NCP and NVP), and then reduces them to MDP and SVP. We follow this general outline. That is, we first show, via relatively simple reductions from PIH, that both  $k$ -NCP and  $k$ -NVP are hard to approximate. Then, we reduce  $k$ -NCP and  $k$ -NVP to  $k$ -MDP and  $k$ -SVP respectively. In this second step, we employ Dumer et al.’s reduction [DMS03] for  $k$ -MDP and Khot’s reduction [Kho05] for  $k$ -SVP. While the latter works almost immediately in the parameterized regime, there are several technical challenges in adapting Dumer et al.’s reduction to our setting. The remainder of this section is devoted to presenting all of our reductions and to highlight such technical challenges and changes in comparison with the non-parameterized settings.

The starting point of all the hardness results in this paper is Gap-ETH (Hypothesis 3.10). As mentioned earlier, it is well-known that Gap-ETH implies PIH (Hypothesis 3.9), i.e., PIH is weaker than Gap-ETH. Hence, for the rest of this section, we may start from PIH instead of Gap-ETH.

### 2.1 Parameterized Intractability of $k$ -MDP from PIH

We start this subsection by describing the Dumer et al.’s (henceforth DMS) reduction [DMS03]. The starting point of the DMS reduction is the NP-hardness of approximating NCP to any constant factor [ABSS97]. Let us recall that in NCP we are given a matrix  $\mathbf{A} \in \mathbb{F}_2^{n \times m}$ , an integer  $k$ , and a target vector  $\mathbf{y} \in \mathbb{F}_2^n$ , and the goal is to determine whether there is any  $\mathbf{x} \in \mathbb{F}_2^m$  such that  $\|\mathbf{Ax} - \mathbf{y}\|_0$  is at most  $k$ . Arora et al. [ABSS97] shows that for any constant  $\gamma \geq 1$ , it is NP-hard to distinguish the case when there exists  $\mathbf{x}$  such that  $\|\mathbf{Ax} - \mathbf{y}\|_0 \leq k$  from the case when for all  $\mathbf{x}$  we have that  $\|\mathbf{Ax} - \mathbf{y}\|_0 > \gamma k$ . Dumer et al. introduce the notion of “locally dense codes” to enable a gadget reduction from NCP to MDP. Informally, a locally dense code is a linear code  $\mathbf{L}$  with minimum distance  $d$  admitting a ball  $\mathcal{B}(\mathbf{s}, r)$  centered at  $\mathbf{s}$  of radius<sup>9</sup>  $r < d$  and containing a large (exponential in the dimension) number of codewords. Moreover, for the gadget reduction to MDP to go through, we require not only the knowledge of the code, but also the center  $\mathbf{s}$  and a linear transformation  $\mathbf{T}$  used to index the codewords in  $\mathcal{B}(\mathbf{s}, r)$ , i.e.,  $\mathbf{T}$  maps  $\mathcal{B}(\mathbf{s}, r) \cap \mathbf{L}$  onto a smaller subspace. Given an instance  $(\mathbf{A}, \mathbf{y}, k)$  of NCP, and a locally dense code  $(\mathbf{L}, \mathbf{T}, \mathbf{s})$  whose parameters (such as dimension and distance) we will fix later, Dumer et al. build the following matrix:

$$\mathbf{B} = \left[ \begin{array}{cc} \mathbf{ATL} & -\mathbf{y} \\ \vdots & \vdots \\ \mathbf{ATL} & -\mathbf{y} \\ \mathbf{L} & -\mathbf{s} \\ \vdots & \vdots \\ \mathbf{L} & -\mathbf{s} \end{array} \right] \left. \begin{array}{l} \vphantom{\left[ \begin{array}{cc} \mathbf{ATL} & -\mathbf{y} \\ \vdots & \vdots \\ \mathbf{ATL} & -\mathbf{y} \\ \mathbf{L} & -\mathbf{s} \\ \vdots & \vdots \\ \mathbf{L} & -\mathbf{s} \end{array} \right]} \vphantom{\left. \begin{array}{l} \\ \\ \\ \\ \\ \end{array} \right\}} \vphantom{\left. \begin{array}{l} \\ \\ \\ \\ \\ \end{array} \right\}} \left. \begin{array}{l} \vphantom{\left[ \begin{array}{cc} \mathbf{ATL} & -\mathbf{y} \\ \vdots & \vdots \\ \mathbf{ATL} & -\mathbf{y} \\ \mathbf{L} & -\mathbf{s} \\ \vdots & \vdots \\ \mathbf{L} & -\mathbf{s} \end{array} \right]} \vphantom{\left. \begin{array}{l} \\ \\ \\ \\ \\ \end{array} \right\}} \vphantom{\left. \begin{array}{l} \\ \\ \\ \\ \\ \end{array} \right\}} \left. \begin{array}{l} \\ \\ \\ \\ \\ \end{array} \right\} \begin{array}{l} a \text{ copies} \\ \\ b \text{ copies} \end{array} \right\} , \quad (1)$$

<sup>9</sup>Note that for the ball to contain more than a single codeword, we must have  $r \geq d/2$ .

where  $a, b$  are some appropriately chosen positive integers. If there exists  $\mathbf{x}$  such that  $\|\mathbf{Ax} - \mathbf{y}\|_0 \leq k$  then consider  $\mathbf{z}'$  such that  $\mathbf{TLz}' = \mathbf{x}$  (we choose the parameters of  $(\mathbf{L}, \mathbf{T}, \mathbf{s})$ , in particular the dimensions of  $\mathbf{L}$  and  $\mathbf{T}$  such that all these computations are valid). Let  $\mathbf{z} = \mathbf{z}' \circ \mathbf{1}$ , and note that  $\|\mathbf{Bz}\|_0 = a\|\mathbf{Ax} - \mathbf{y}\|_0 + b\|\mathbf{Lz} - \mathbf{s}\|_0 \leq ak + br$ . In other words, if  $(\mathbf{A}, \mathbf{y}, k)$  is a YES instance of NCP then  $(\mathbf{B}, ak + br)$  is a YES instance of MDP. On the other hand if we had that for all  $\mathbf{x}$ , the norm of  $\|\mathbf{Ax} - \mathbf{y}\|_0$  is more than  $\gamma k$  for some constant<sup>10</sup>  $\gamma > 2$ , then it is possible to show that for all  $\mathbf{z}$  we have that  $\|\mathbf{Bz}\|_0 > \gamma'(ak + br)$  for any  $\gamma' < \frac{2\gamma}{2+\gamma}$ . The proof is based on a case analysis of the last coordinate of  $\mathbf{z}$ . If that coordinate is 0, then, since  $\mathbf{L}$  is a code of distance  $d$ , we have  $\|\mathbf{Bz}\|_0 \geq bd > \gamma'(ak + br)$ ; if that coordinate is 1, then the assumption that  $(\mathbf{A}, \mathbf{y}, k)$  is a NO instance of NCP implies that  $\|\mathbf{Bz}\|_0 > a\gamma k > \gamma'(ak + br)$ . Note that this gives an inapproximability for MDP of ratio  $\gamma' < 2$ ; this gap is then further amplified by a simple tensoring procedure.

We note that Dumer et al. were not able to find a deterministic construction of locally dense code with all of the above described properties. Specifically, they gave an efficient deterministic construction of a code  $\mathbf{L}$ , but only gave a randomized algorithm that finds a linear transformation  $\mathbf{T}$  and a center  $\mathbf{s}$  w.h.p. Therefore, their hardness result relies on the assumption that  $\text{NP} \neq \text{RP}$ , instead of the more standard  $\text{NP} \neq \text{P}$  assumption. Later, Cheng and Wan [CW12] and Micciancio [Mic14] provided constructions for such (families of) locally dense codes with an explicit center, and thus showed the constant ratio inapproximability of MDP under the assumption of  $\text{NP} \neq \text{P}$ .

Trying to follow the DMS reduction in order to show the parameterized intractability of  $k$ -MDP, we face the following three immediate obstacles. First, there is no inapproximability result known for  $k$ -NCP, for any constant factor greater than 1. Note that to use the DMS reduction, we need the parameterized inapproximability of  $k$ -NCP, for an approximation factor which is greater than two. Second, the construction of locally dense codes of Dumer et al. only works when the distance is linear in the block length (which is a function of the size of the input). However, we need codes whose distance are bounded above by a function of the parameter of the problem (and not depend on the input size). This is because the DMS reduction converts an instance  $(\mathbf{A}, \mathbf{y}, k)$  of  $k$ -NCP to an instance  $(\mathbf{B}, ak + br)$  of  $(ak + br)$ -MDP, and for this reduction to be an FPT reduction, we need  $ak + br$  to be a function only depending on  $k$ , i.e.,  $d$ , the distance of the code  $\mathbf{L}$  (which is at most  $2r$ ), must be a function only of  $k$ . Third, recall that the DMS reduction needs to identify the vectors in the ball  $\mathcal{B}(\mathbf{s}, r) \cap \mathbf{L}$  with all the potential solutions of  $k$ -NCP. Notice that the number of vectors in the ball is at most  $(nm)^{O(r)}$  but the number of potential solutions of  $k$ -NCP is exponential in  $m$  (i.e. all  $\mathbf{x} \in \mathbb{F}_2^m$ ). However, this is impossible since  $r \leq d$  is bounded above by a function of  $k$ !

We overcome the first obstacle by proving the constant inapproximability of  $k$ -NCP under PIH. Specifically, assuming PIH, we first show the parameterized inapproximability of  $k$ -NCP for some constant factor greater than 1, and then boost the gap using a composition operator (self-recursively). Note that in order to follow the DMS reduction, we need the inapproximability of  $k$ -NCP for some constant factor greater than 2; in other words, the gap amplification for  $k$ -NCP is necessary, even if we are not interested in showing the inapproximability of  $k$ -NCP for all constant factors.

We overcome the third obstacle by introducing an intermediate problem in the DMS reduction, which we call the *sparse nearest codeword problem*. The sparse nearest codeword problem is a promise problem which differs from  $k$ -NCP in only one way: in the YES case, we want to find  $\mathbf{x} \in \mathcal{B}(\mathbf{0}, k)$

<sup>10</sup>Note that in the described reduction, we need the inapproximability of NCP to a factor greater than two, even to just reduce to the *exact* version of MDP.

(instead of the entire space  $\mathbb{F}_2^m$ ), such that  $\|\mathbf{Ax} - \mathbf{y}\|_0 \leq k$ . In other words, we only allow sparse  $\mathbf{x}$  as a solution. We show that  $k$ -NCP can be reduced to the sparse nearest codeword problem.

Finally, we overcome the second obstacle by introducing a variant of locally dense codes, which we call *sparse covering codes*. Roughly speaking, we show that any code which nears the sphere-packing bound (aka Hamming bound) in the high rate regime is a sparse covering code. Then we follow the DMS reduction with the new ingredient of sparse covering codes (replacing locally dense codes) to reduce the sparse nearest codeword problem to  $k$ -MDP.

We remark here that overcoming the second and third obstacles are our main technical contributions. In particular, our result on sparse covering codes might be of independent interest.

The full reduction goes through several intermediate steps, which we will describe in more detail in the coming paragraphs. The high-level summary of these steps is also provided in Figure 1. Throughout this section, for any gap problem, if we do not specify the gap in the subscript, then it implies that the gap can be any arbitrary constant. For every  $\varepsilon \geq 0$ , we denote by  $\text{GAP2CSP}_\varepsilon$  the gap problem where we have to determine if a given 2CSP instance  $\Gamma$ , i.e., a graph  $G = (V, E)$  and a set of constraints  $\{C_{uv}\}_{(u,v) \in E}$  over an alphabet set  $\Sigma$ , has an assignment to its vertices that satisfies all the constraints or if every assignment violates more than  $\varepsilon$  fraction of the constraints. Here each  $C_{uv}$  is simply the set of all  $(\sigma_u, \sigma_v) \in \Sigma \times \Sigma$  that satisfy the constraint. The parameter of the problem is  $|V|$ . PIH asserts that there exists some constant  $\varepsilon > 0$  such that no randomized FPT algorithm can solve  $\text{GAP2CSP}_\varepsilon$ . (See Hypothesis 3.9 for a more formal statement.)

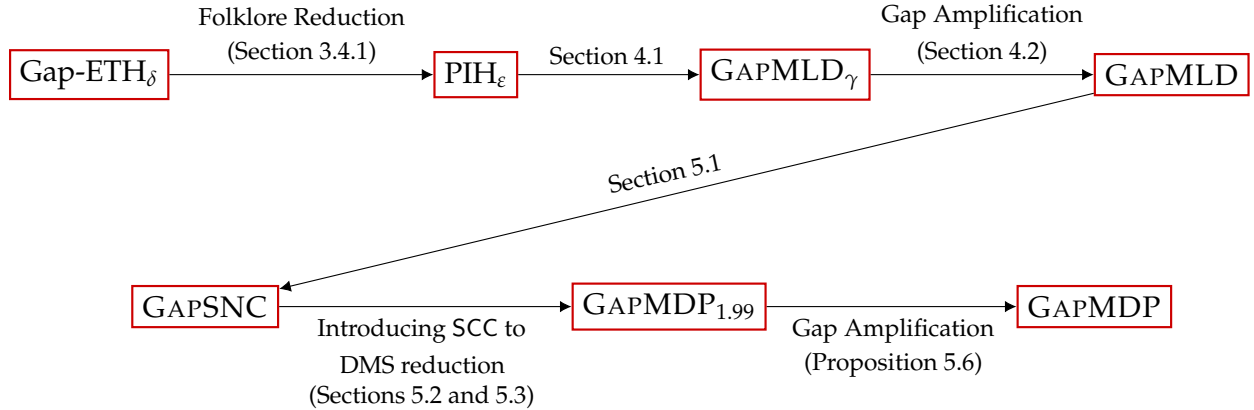


Figure 1: An overview of the reduction from Gap-ETH to the parameterized Minimum Distance problem. First, PIH follows from Gap-ETH due to known reductions (see Section 3.4.1). Next, we reduce  $\text{GAP2CSP}_\varepsilon$  to  $\text{GAPMLD}_\gamma$  in Section 4.1, and then amplify the gap using a composition operator in Section 4.2. Via a simple reduction from  $\text{GAPMLD}$ , in Section 5.1 we obtain the constant parameterized inapproximability of  $\text{GAPSNC}$ . In Section 5.2, we formally introduce sparse covering codes and show how to efficiently (but probabilistically) construct them. These codes are then used in Section 5 to obtain the parameterized innapproximability of  $\text{GAPMDP}_{1.99}$ . The final step is a known gap amplification by tensoring (Proposition 5.6).

**Reducing  $\text{GAP2CSP}_\varepsilon$  to  $\text{GAPMLD}_\gamma$ .** We start by showing the parameterized inapproximability of  $k$ -NCP for some constant ratio. Instead of working with  $k$ -NCP, we work with its equivalent formulation (by converting the generator matrix given as input into a parity-check matrix) which in



the literature is referred to as the *maximum likelihood decoding problem*<sup>11</sup>. We define the gap version of this problem (i.e., a promise problem), denoted by  $\text{GAPMLD}_\gamma$  (for some constant  $\gamma \geq 1$ ) as follows: on input  $(\mathbf{A}, \mathbf{y}, k)$ , distinguish between the YES case where there exists  $\mathbf{x} \in \mathcal{B}(\mathbf{0}, k)$  such that  $\mathbf{A}\mathbf{x} = \mathbf{y}$ , and the NO case where for all  $\mathbf{x} \in \mathcal{B}(\mathbf{0}, \gamma k)$  we have  $\mathbf{A}\mathbf{x} \neq \mathbf{y}$ . It is good to keep in mind that this is equivalent to asking whether there exist  $k$  columns of  $\mathbf{A}$  whose sum is equal to  $\mathbf{y}$  or whether any  $\leq \gamma k$  columns of  $\mathbf{A}$  do not sum up to  $\mathbf{y}$ .

Now, we will sketch the reduction from an instance  $\Gamma = (G = (V, E), \Sigma, \{C_{uv}\}_{(u,v) \in E})$  of  $\text{GAP2CSP}_\varepsilon$  to an instance  $(\mathbf{A}, \mathbf{y}, k)$  of  $\text{GAPMLD}_{1+\varepsilon/3}$ . The matrix  $\mathbf{A}$  will have  $|V||\Sigma| + \sum_{(u,v) \in E} |C_{uv}|$  columns and  $|V| + |E| + 2|E||\Sigma|$  rows. The first  $|V||\Sigma|$  columns of  $\mathbf{A}$  are labelled with  $(u, \sigma_u) \in V \times \Sigma$ , and the remaining columns of  $\mathbf{A}$  are labeled by  $(e, \sigma_u, \sigma_v)$  where  $e = (u, v) \in E$  and  $(\sigma_u, \sigma_v) \in C_{uv}$ .

Before we continue with our description of  $\mathbf{A}$ , let us note that, in the YES case where there is an assignment  $\phi : V \rightarrow \Sigma$  that satisfies every constraint, our intended solution for our  $\text{GAPMLD}$  instance is to pick the  $(u, \phi(u))$ -column for every  $u \in V$  and the  $((u, v), \phi(u), \phi(v))$ -column for every  $(u, v) \in E$ . Notice that  $|V| + |E|$  columns are picked, and indeed we set  $k = |V| + |E|$ . Moreover, we set the first  $|V| + |E|$  coordinates of  $\mathbf{y}$  to be one and the rest to be zero.

We also identify the first  $|V|$  rows of  $\mathbf{A}$  with  $u \in V$ , the next  $|E|$  rows of  $\mathbf{A}$  with  $e \in E$ , and the remaining  $2|E||\Sigma|$  rows of  $\mathbf{A}$  with  $(e, \sigma, b) \in E \times \Sigma \times \{0, 1\}$ . Figure 2 provides an illustration of the matrix  $\mathbf{A}$ . The rows of  $\mathbf{A}$  will be designed to serve the following purposes: the first  $|V|$  rows will ensure that, for each  $u \in V$ , at least one column of the form  $(u, \cdot)$  is picked, the next  $|E|$  rows will ensure that, for each  $e \in E$ , at least one column of the form  $(e, \cdot, \cdot)$  is picked, and finally the remaining  $2|E||\Sigma|$  rows will “check” that the constraint is indeed satisfied.

Specifically, each  $u$ -row for  $u \in V$  has only  $|\Sigma|$  non-zero entries: those in column  $(u, \sigma_u)$  for all  $\sigma_u \in \Sigma$ . Since our target vector  $\mathbf{y}$  has  $\mathbf{y}_u = 1$ , we indeed have that at least one column of the form  $(u, \cdot)$  must be selected for every  $u \in V$ . Similarly, each  $e$ -row for  $e = (u, v) \in E$  has  $|C_{uv}|$  non-zero entries: those in column  $(e, \sigma_u, \sigma_v)$  for all  $(\sigma_u, \sigma_v) \in C_{uv}$ . Again, these make sure that at least one column of the form  $(e, \cdot, \cdot)$  must be picked for every  $e \in E$ .

Finally, we will define the entries of the last  $2|E||\Sigma|$  rows. To do so, let us recall that, in the YES case, we pick the columns  $(u, \phi(u))$  for all  $u \in V$  and  $((u, v), \phi(u), \phi(v))$  for all  $(u, v) \in E$ . The goal of these remaining rows is to not only accept such a solution but also prevent any solution that picks columns  $(u, \sigma_u), (v, \sigma_v)$  and  $((u, v), \sigma'_u, \sigma'_v)$  where  $\sigma_u \neq \sigma'_u$  or  $\sigma_v \neq \sigma'_v$ . In other words, these rows serve as a “consistency checker” of the solution. Specifically, the  $|\Sigma|$  rows of the form  $((u, v), \cdot, 0)$  will force  $\sigma_u$  and  $\sigma'_u$  to be equal whereas the  $|\Sigma|$  rows of the form  $((u, v), \cdot, 1)$  will force  $\sigma_v$  and  $\sigma'_v$  to be equal. For convenience, we will only define the entries for the  $((u, v), \cdot, 0)$ -rows; the  $((u, v), \cdot, 1)$ -rows can be defined similarly. Each  $((u, v), \sigma, 0)$ -row has only one non-zero entry within the first  $|V||\Sigma|$  rows: the one in the  $(u, \sigma)$ -column. For the remaining columns, the entry in the  $((u, v), \sigma, 0)$ -row and the  $(e, \sigma_0, \sigma_1)$ -column is non-zero if and only if  $e = (u, v)$  and  $\sigma_0 = \sigma$ .

It should be clear from the definition that our intended solution for the YES case is indeed a valid solution because, for each  $((u, v), \phi(u), 0)$ -row, the two non-zero entries from the columns  $(u, \phi(u))$  and  $((u, v), \phi(u), \phi(v))$  cancel each other out. On the other hand, for the NO case, the

<sup>11</sup>The two formulations are equivalent but we use different names for them to avoid confusion when we use *Sparse Nearest Codeword Problem* later on.

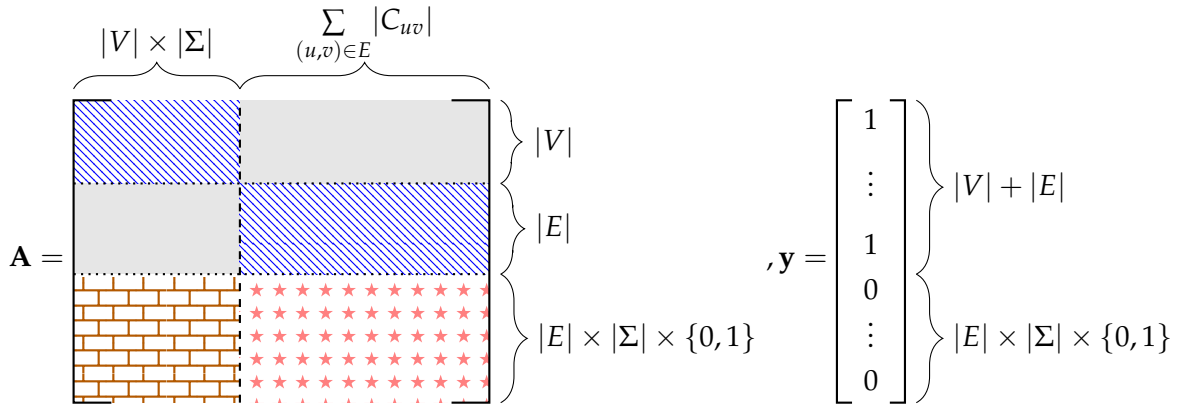


Figure 2: An illustration of  $\mathbf{A}$  and  $\mathbf{y}$ . All entries in the shaded area are zero. Each *row* in the brick pattern area has one non-zero entry in that area, and each *column* in the star pattern area has two non-zero entries in the area. Finally, each column has one non-zero entry in the lines pattern area.

main observation is that, for each edge  $(u, v) \in E$ , if only one column of the form  $(u, \cdot)$ , one of the form  $(v, \cdot)$  and one of the form  $((u, v), \cdot, \cdot)$  are picked, then the assignment corresponding to the picked columns satisfy the constraint  $C_{uv}$ . In particular, it is easy to argue that, if we can pick  $(1 + \varepsilon/3)(|V| + |E|)$  columns that sum up to  $\mathbf{y}$ , then all but  $\varepsilon$  fraction of all constraints fulfill the previous conditions, meaning that we can find an assignment that satisfies  $1 - \varepsilon$  fraction of the constraints. Thus, we have also proved the soundness of the reduction.

**Gap Amplification for  $\text{GAPMLD}_\gamma$ .** We have sketched the proof of the hardness of  $\text{GAPMLD}_\gamma$  for *some* constant  $\gamma \geq 1$ , assuming PIH. The next step is to amplify the gap and arrive at the hardness for  $\text{GAPMLD}_\gamma$  for *every* constant  $\gamma \geq 1$ . To do so, we define an operator  $\oplus$  over every pair of instances of  $\text{GAPMLD}_\gamma$  with the following property: if two instances  $(\mathbf{A}_1, \mathbf{y}_1, k_1)$  and  $(\mathbf{A}_2, \mathbf{y}_2, k_2)$  are both YES instances, then  $(\mathbf{A}, \mathbf{y}, k) := (\mathbf{A}_1, \mathbf{y}_1, k_1) \oplus (\mathbf{A}_2, \mathbf{y}_2, k_2)$  is a YES instance for  $\text{GAPMLD}_{\gamma'}$  where  $\gamma' \approx \gamma^2$ . On the other hand, if both  $(\mathbf{A}_1, \mathbf{y}_1, k_1)$  and  $(\mathbf{A}_2, \mathbf{y}_2, k_2)$  are NO instances, then  $(\mathbf{A}, \mathbf{y}, k)$  is a NO instance for  $\text{GAPMLD}_{\gamma'}$ . Hence, we can apply  $\oplus$  repeatedly to the  $\text{GAPMLD}_\gamma$  instance from the previous step (with itself) and amplify the gap to be any arbitrarily large constant. The definition of  $\oplus$ , while simple, is slightly tedious to formalized and we defer it to Section 4.2.

**Reducing  $\text{GAPMLD}$  to  $\text{GAPSNC}$ .** Now we introduce the *sparse nearest codeword problem* that we had briefly talked about. We define the gap version of this problem, denoted by  $\text{GAPSNC}_\gamma$  (for some constant  $\gamma \geq 1$ ) as follows: on input  $(\mathbf{A}', \mathbf{y}', k)$ , distinguish between the YES case where there exists  $\mathbf{x} \in \mathcal{B}(\mathbf{0}, k)$  such that  $\|\mathbf{A}'\mathbf{x} - \mathbf{y}'\|_0 \leq k$ , and the NO case where for all  $\mathbf{x}$  (in the entire space), we have  $\|\mathbf{A}'\mathbf{x} - \mathbf{y}'\|_0 > \gamma k$ . We highlight that the difference between  $k$ -NCP and  $\text{GAPSNC}_\gamma$  is that, in the YES case of the latter, we are promised that  $\mathbf{x} \in \mathcal{B}(\mathbf{0}, k)$ . We sketch below the reduction from

an instance  $(\mathbf{A}, \mathbf{y}, k)$  of  $\text{GAPMLD}_\gamma$  to an instance  $(\mathbf{A}', \mathbf{y}', k)$  of  $\text{GAPSN}_\gamma$ . Given  $\mathbf{A}, \mathbf{y}$ , let

$$\mathbf{A}' = \left. \begin{bmatrix} \mathbf{A} \\ \vdots \\ \mathbf{A} \\ \text{Id} \end{bmatrix} \right\} \gamma k + 1 \text{ copies}, \quad \mathbf{y}' = \left. \begin{bmatrix} \mathbf{y} \\ \vdots \\ \mathbf{y} \\ \mathbf{0} \end{bmatrix} \right\} \gamma k + 1 \text{ copies}.$$

Notice that for any  $\mathbf{x}$  (in the entire space), we have

$$\|\mathbf{A}'\mathbf{x} - \mathbf{y}'\|_0 = (\gamma k + 1)\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_0 + \|\mathbf{x}\|_0,$$

and thus both the completeness and soundness of the reduction easily follow.

**Sparse Covering Codes.** Before reducing  $\text{GAPSN}_\gamma$  to  $\text{GAPMDP}_{1.99}$  we need to introduce in more detail (but still informally) the notion of *sparse covering codes* that we previously mentioned.

A sparse covering code (SCC) is a linear code  $\mathbf{L}$  of block length  $h$  with minimum distance  $d$  admitting a ball  $\mathcal{B}(\mathbf{s}, r)$  centered at  $\mathbf{s}$  of radius  $r < d$  and containing a large (i.e., about  $h^k$ , where  $k = \Omega(d)$ ) number of codewords. Moreover, for the reduction to  $k$ -MDP to go through, we require not only the knowledge of the code, but also the center  $\mathbf{s}$  and a linear transformation  $\mathbf{T}$  used to index the codewords in  $\mathcal{B}(\mathbf{s}, r)$ , i.e.,  $\mathbf{T}(\mathcal{B}(\mathbf{s}, r) \cap \mathbf{L})$  needs to contain the ball of radius  $k$  centered at  $\mathbf{0}$ . Similar to how Dumer et al. only managed to show the probabilistic existence of the center, we too cannot find an explicit  $\mathbf{s}$  for the SCCs that we construct, but instead provide an efficiently samplable distribution such that, for any  $\mathbf{x} \in \mathcal{B}(\mathbf{0}, k)$ , the probability (over  $\mathbf{s}$  sampled from the distribution) that  $\mathbf{x} \in \mathbf{T}(\mathcal{B}(\mathbf{s}, r) \cap \mathbf{L})$  is non-negligible. This is what makes our reduction from  $\text{GAPSN}_\gamma$  to  $\text{GAPMDP}_{1.99}$  randomized. We will not elaborate more on this issue here, but focus on the (probabilistic) construction of such codes. For convenience, we will assume throughout this overview that  $k$  is much smaller than  $d$ , i.e.,  $k = 0.001d$ .

Recall that the sphere-packing bound (aka Hamming bound) states that a binary code of block length  $h$  and distance  $d$  can have at most  $2^h / |\mathcal{B}(\mathbf{0}, \lceil \frac{d-1}{2} \rceil)|$  codewords; this is simply because the balls of radius  $\lceil \frac{d-1}{2} \rceil$  at the codewords do not intersect. Our main theorem regarding the existence of sparse covering code is that any code that is “near” the sphere-packing bound is a sparse covering code with  $r = \lceil \frac{d-1}{2} \rceil + k \approx 0.501d$ . Here “near” means that the number of codewords must be at least  $2^h / |\mathcal{B}(\mathbf{0}, \lceil \frac{d-1}{2} \rceil)|$  divided by  $f(d) \cdot \text{poly}(h)$  for some function  $f$  that depends only on  $d$ . (Equivalently, this means that the message length must be at least  $h - (d/2 + O(1)) \log h$ .) The BCH code over binary alphabet is an example of a code satisfying such a condition.

While we will not sketch the proof of the existence theorem here, we note that the general idea is to set  $\mathbf{T}$  and the distribution over  $\mathbf{s}$  in such a way that the probability that  $\mathbf{x}$  lies in  $\mathbf{T}(\mathcal{B}(\mathbf{s}, r) \cap \mathbf{L})$  is at least the probability that a random point in  $\mathbb{F}_2^h$  is within distance  $r - k = \lceil \frac{d-1}{2} \rceil$  of some codeword. The latter is non-negligible from our assumption that  $\mathbf{L}$  nears the sphere-packing bound.

Finally, we remark that our proof here is completely different from the DMS proof of existence of locally dense codes. Specifically, DMS uses a group-theoretic argument to show that, when a code exceeds the Gilbert–Varshamov bound, there must be a center  $\mathbf{s}$  such that  $\mathcal{B}(\mathbf{s}, r)$  contains many

codewords. Then, they pick a random linear map  $\mathbf{T}$  and show that w.h.p.  $\mathbf{T}(\mathcal{B}(\mathbf{s}, r) \cap \mathbf{L})$  is the entire space. Note that this second step does not use any structure of  $\mathcal{B}(\mathbf{s}, r) \cap \mathbf{L}$ ; their argument is simply that, for any sufficiently large subset  $Y$ , a random linear map  $\mathbf{T}$  maps  $Y$  to an entire space w.h.p. However, such an argument fails for us, due to the fact that, in SCC, we want to cover a ball  $\mathcal{B}(\mathbf{0}, k)$  rather than the whole space, and it is not hard to see that there are very large subsets  $Y$  such that no linear map  $\mathbf{T}$  satisfies  $\mathbf{T}(Y) \supseteq \mathcal{B}(\mathbf{0}, k)$ . A simple example of this is when  $Y$  is a subspace of  $\mathbb{F}_2^h$ ; in this case, even when  $Y$  is as large as  $\exp(\text{poly}(h))$ , no desired linear map  $\mathbf{T}$  exists.

**Reducing  $\text{GAPSNC}_\gamma$  to  $\text{GAPMDP}_{1.99}$ .** Next, we prove the hardness of  $\text{GAPMDP}_{\gamma'}$  for all constant  $\gamma' \in [1, 2)$ , assuming PIH, using a gadget constructed from sparse covering codes.

Given an instance  $(\mathbf{A}, \mathbf{y}, k)$  of  $\text{GAPSNC}_\gamma$  for some  $\gamma > 2$  and a sparse covering code  $(\mathbf{L}, \mathbf{T}, \mathbf{s})$  we build an instance  $(\mathbf{B}, ak + br)$  of  $\text{GAPMDP}_{\gamma'}$  where  $\gamma' < \frac{2\gamma}{2+\gamma}$ , by following the DMS reduction (which was previously described, and in particular see (1)). If there exists  $\mathbf{x} \in \mathcal{B}(\mathbf{0}, k)$  such that  $\|\mathbf{Ax} - \mathbf{y}\|_0 \leq k$  then consider  $\mathbf{z}'$  such that  $\mathbf{TLz}' = \mathbf{x}$ . Note that the existence of such a  $\mathbf{z}'$  is guaranteed by the definition of SCC. Consider  $\mathbf{z} = \mathbf{z}' \circ \mathbf{1}$ , and note that  $\|\mathbf{Bz}\|_0 = a\|\mathbf{Ax} - \mathbf{y}\|_0 + b\|\mathbf{Lz} - \mathbf{s}\|_0 \leq ak + br$ . In other words, as in the DMS reduction, if  $(\mathbf{A}, \mathbf{y}, k)$  is a YES instance of NCP, then  $(\mathbf{B}, ak + br)$  is a YES instance of MDP. On the other hand, similar to the DMS reduction, if we had that  $\|\mathbf{Ax} - \mathbf{y}\|_0 > \gamma k$  for all  $\mathbf{x}$ , then  $\|\mathbf{Bz}\|_0 > \gamma'(ak + br)$  for all  $\mathbf{z}$ . The parameterized intractability of  $\text{GAPMDP}_{1.99}$  is obtained by setting  $\gamma = 400$  in the above reduction.

**Gap Amplification for  $\text{GAPMDP}_{1.99}$ .** It is well known that the distance of the tensor product of two linear codes is the product of the distances of the individual codes (see Proposition 5.6 for a formal statement). We can use this proposition to reduce  $\text{GAPMDP}_\gamma$  to  $\text{GAPMDP}_{\gamma^2}$  for any  $\gamma \geq 1$ . In particular, we can obtain, for any constant  $\gamma$ , the intractability of  $\text{GAPMDP}_\gamma$  starting from  $\text{GAPMDP}_{1.99}$  by just recursively tensoring the input code  $\lceil \log_{1.99} \gamma \rceil$  times.

## 2.2 Parameterized Intractability of $k$ -SVP from PIH

We begin this subsection by briefly describing Khot's reduction. The starting point of Khot's reduction is the NP-hardness of approximating NVP in every  $\ell_p$  norm to any constant factor [ABSS97]. Let us recall that in NVP in the  $\ell_p$  norm, we are given a matrix  $\mathbf{A} \in \mathbb{Z}^{n \times m}$ , an integer  $k$ , and a target vector  $\mathbf{y} \in \mathbb{Z}^n$ , and the goal is to determine whether there is any  $\mathbf{x} \in \mathbb{Z}^m$  such that<sup>12</sup>  $\|\mathbf{Ax} - \mathbf{y}\|_p^p$  is at most  $k$ . The result of Arora et al. [ABSS97] states that for any constant  $\gamma \geq 1$ , it is NP-hard to distinguish the case when there exists  $\mathbf{x}$  such that  $\|\mathbf{Ax} - \mathbf{y}\|_p^p \leq k$  from the case when for all (integral)  $\mathbf{x}$  we have that  $\|\mathbf{Ax} - \mathbf{y}\|_p^p > \gamma k$ . Khot's reduction proceeds in four steps. First, he constructs a gadget lattice called the "BCH Lattice" using BCH Codes. Next, he reduces NVP in the  $\ell_p$  norm (where  $p \in (1, \infty)$ ) to an instance of SVP on an intermediate lattice by using the BCH Lattice. This intermediate lattice has the following property. For any YES instance of NVP the intermediate lattice contains multiple copies of the witness of the YES instance; For any NO instance of NVP there are also many "annoying vectors" (but far less than the total number of YES instance witnesses) which look like witnesses of a YES instance. However, since the annoying vectors are outnumbered, Khot reduces this intermediate lattice to a proper SVP instance, by randomly picking a sub-lattice via a random homogeneous linear constraint on the coordinates of the lattice vectors (this annihilates all the annoying vectors while retaining at least one witness for the YES

<sup>12</sup>Previously, we use  $\|\mathbf{Ax} - \mathbf{y}\|_p$  instead of  $\|\mathbf{Ax} - \mathbf{y}\|_p^p$ . However, from the fixed parameter perspective, these two versions are equivalent since the parameter  $k$  is only raised to the  $p$ -th power, and  $p$  is a constant in our setting.

instance). Thus he obtains some constant factor hardness for SVP. Finally, the gap is amplified via “Augmented Tensor Product”. It is important to note that Khot’s reduction is randomized, and thus his result of inapproximability of SVP is based on  $\text{NP} \neq \text{RP}$ .

Trying to follow Khot’s reduction, in order to show the parameterized intractability of  $k$ -SVP, we face only one obstacle: there is no known parameterized inapproximability of  $k$ -NVP for any constant factor greater than 1. Let us denote by  $\text{GAPNVP}_{p,\eta}$  for any constant  $\eta \geq 1$  the gap version of  $k$ -NVP in the  $\ell_p$  norm. Recall that in  $\text{GAPNVP}_{p,\eta}$  we are given a matrix  $\mathbf{A} \in \mathbb{Z}^{n \times m}$ , a target vector  $\mathbf{y} \in \mathbb{Z}^n$ , and a parameter  $k$ , and we would like to distinguish the case when there exists  $\mathbf{x} \in \mathbb{Z}^m$  such that  $\|\mathbf{Ax} - \mathbf{y}\|_p^p \leq k$  from the case when for all  $\mathbf{x} \in \mathbb{Z}^m$  we have that  $\|\mathbf{Ax} - \mathbf{y}\|_p^p > \eta k$ . As it turns out, our reduction from  $\text{GAP2CSP}_\varepsilon$  to  $\text{GAPSNC}$  (with arbitrary constant gap), having  $\text{GAPMLD}_\gamma$  and  $\text{GAPMLD}$  as intermediate steps, can be translated to show the constant inapproximability of  $\text{GAPNVP}_p$  (under PIH) in a straightforward manner. We will not elaborate on this part of the proof any further here and defer the detailed proof to Appendix B.

Once we have established the constant parameterized inapproximability of  $\text{GAPNVP}_p$ , we follow Khot’s reduction, and everything goes through as it is to establish the inapproximability for some factor of the gap version of  $k$ -SVP in the  $\ell_p$  norm (where  $p \in (1, \infty)$ ). We denote by  $\text{GAPSVP}_{p,\gamma}$  for some constant  $\gamma(p) \geq 1$  the the gap version of  $k$ -SVP (in the  $\ell_p$  norm) where we are given a matrix  $\mathbf{B} \in \mathbb{Z}^{n \times m}$  and a parameter  $k \in \mathbb{N}$ , and we would like to distinguish the case when there exists a non-zero  $\mathbf{x} \in \mathbb{Z}^m$  such that  $\|\mathbf{Bx}\|_p^p \leq k$  from the case when for all  $\mathbf{x} \in \mathbb{Z}^m \setminus \{\mathbf{0}\}$  we have that  $\|\mathbf{Bx}\|_p^p > \gamma k$ . Let  $\gamma^* := \frac{2^p}{2^{p-1}+1}$ . Following Khot’s reduction, we obtain the inapproximability of  $\text{GAPSVP}_{p,\gamma^*}$  (under PIH). To obtain inapproximability of  $\text{GAPSVP}_2$  for all constant ratios, we use the tensor product of lattices; the argument needed here is slightly more subtle than the similar step in MDP because, unlike distances of codes, the  $\ell_2$  norm of the shortest vector of the tensor product of two lattices is not necessarily equal to the product of the  $\ell_2$  norm of the shortest vector of each lattice. Fortunately, Khot’s construction is tailored so that the resulting lattice is “well-behaved” under tensoring [Kho05, HR07], and gap amplification is indeed possible for such instances.

We remark here that, for the (non-parameterized) inapproximability of SVP, the techniques of [Kho05, HR07] allow one to successfully amplify gaps for  $\ell_p$  norm where  $p \neq 2$  as well. Unfortunately, this does not work in our settings, as it requires the distance  $k$  to be dependent on  $nm$  which is not possible for us since  $k$  is the parameter of the problem.

Summarizing, in Figure 3, we provide the proof outline of our reduction from Gap-ETH to  $\text{GAPSVP}_p$  with some constant gap, for every  $p \in (1, \infty)$  (with the additional gap amplification to constant inapproximability for  $p = 2$ ).

### 3 Preliminaries

We use the following notations throughout the paper.

**Notations.** For  $p \in \mathbb{N}$ , we use  $\mathbf{1}_p$  (respectively,  $\mathbf{0}_p$ ) to denote the all ones (respectively, all zeros) vector of length  $p$ . We sometimes drop the subscript if the dimension is clear from the context.

For  $p, q \in \mathbb{N}$ , we use  $\mathbf{0}_{p \times q}$  to denote the all zeroes matrix of  $p$  rows and  $q$  columns. We use  $\text{Id}_q$  to denote the identity matrix of  $q$  rows and  $q$  columns.

For any vector  $x \in \mathbb{R}^d$ , the  $\ell_p$  norm of  $x$  is defined as  $\ell_p(x) = \|x\|_p = \left( \sum_{i=1}^d |x_i|^p \right)^{1/p}$ . Thus,

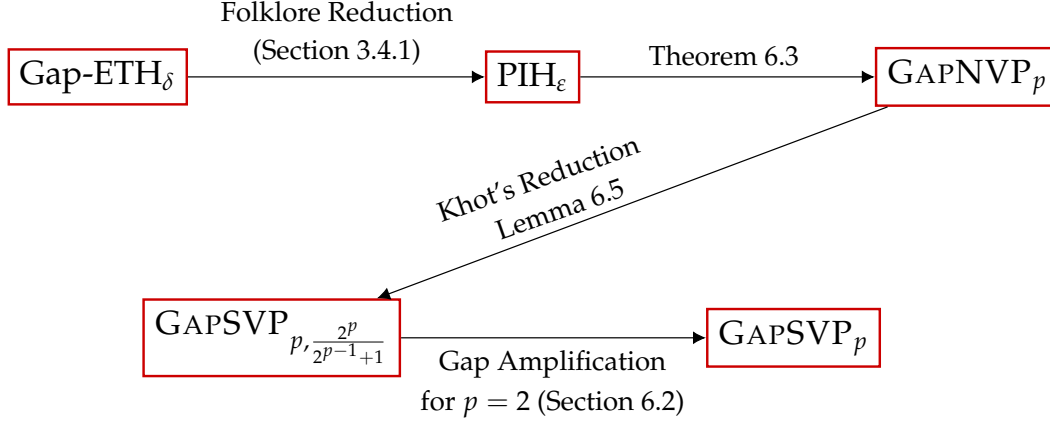


Figure 3: The figure provides an overview of the reduction from Gap-ETH to the parameterized Shortest Vector problem in the  $\ell_p$  norm, where  $p \in (1, \infty)$ . First, recall that Gap-ETH implies PIH (see Section 3.4.1). Next, we reduce GAP2CSP $_{\varepsilon}$  to GAPNVP $_p$  in Appendix B. Lemma 6.5 (i.e., Khot’s reduction) then implies the parameterized inapproximability of GAPSVP $_{p, \frac{2^p}{2^{p-1}+1}}$ . The final step (for  $p = 2$ ) is the Haviv-Regev gap amplification via tensor product, which is described in Section 6.2.

$\ell_{\infty}(x) = \|x\|_{\infty} = \max_{i \in [d]} \{|x_i|\}$ . The  $\ell_0$  norm of  $x$  is defined as  $\ell_0(x) = \|x\|_0 = |\{x_i \neq 0 : i \in [d]\}|$ , i.e., the number of non-zero entries of  $x$ . We note that the  $\ell_0$  norm is also referred to as the Hamming norm. For  $a \in \mathbb{N}$ ,  $t \in \mathbb{N} \cup \{0\}$ , and  $\mathbf{s} \in \{0, 1\}^a$ , we use  $\mathcal{B}_a(\mathbf{s}, t)$  to denote the Hamming ball of radius  $t$  centered at  $\mathbf{s}$ , i.e.,  $\mathcal{B}_a(\mathbf{s}, t) = \{\mathbf{x} \in \{0, 1\}^a \mid \|\mathbf{s} - \mathbf{x}\|_0 \leq t\}$ . Finally, given two vectors  $\mathbf{x} \in \mathbb{F}_2^m$  and  $\mathbf{y} \in \mathbb{F}_2^n$ , we use  $\mathbf{x} \circ \mathbf{y} \in \mathbb{F}_2^{m+n}$  to denote the concatenation of vectors  $\mathbf{x}$  and  $\mathbf{y}$ .

### 3.1 Parameterized Promise Problems and (Randomized) FPT Reductions

In this subsection, we briefly describe the various kinds of fixed-parameter reductions that are used in this paper. We start by defining the notion of promise problems in the fixed-parameter world, which is naturally analogous to promise problems in the NP world (see e.g. [Gol06]).

**Definition 3.1** A parameterized promise problem  $\Pi$  is a pair of parameterized languages  $(\Pi_{\text{YES}}, \Pi_{\text{NO}})$  such that  $\Pi_{\text{YES}} \cap \Pi_{\text{NO}} = \emptyset$ .

Next, we formalize the notion of algorithms for these parameterized promise problems:

**Definition 3.2** A deterministic algorithm  $\mathcal{A}$  is said to be an FPT algorithm for  $\Pi$  if the following holds:

- On any input  $(x, k)$ ,  $\mathcal{A}$  runs in time  $f(k)|x|^c$  for some computable function  $f$  and constant  $c$ .
- (YES) For all  $(x, k) \in \Pi_{\text{YES}}$ ,  $\mathcal{A}(x, k) = 1$ .
- (NO) For all  $(x, k) \in \Pi_{\text{NO}}$ ,  $\mathcal{A}(x, k) = 0$ .

**Definition 3.3** A Monte Carlo algorithm  $\mathcal{A}$  is said to be a randomized FPT algorithm for  $\Pi$  if the following holds:

- $\mathcal{A}$  runs in time  $f(k)|x|^c$  for some computable function  $f$  and constant  $c$  (on every randomness).
- (YES) For all  $(x, k) \in \Pi_{\text{YES}}$ ,  $\Pr[\mathcal{A}(x, k) = 1] \geq 2/3$ .

- (NO) For all  $(x, k) \in \Pi_{NO}$ ,  $\Pr[\mathcal{A}(x, k) = 0] \geq 2/3$ .

Finally, we define deterministic and randomized reductions between these problems.

**Definition 3.4** A (deterministic) FPT reduction from a parameterized promise problem  $\Pi$  to a parameterized promise problem  $\Pi'$  is a (deterministic) procedure that transforms  $(x, k)$  to  $(x', k')$  that satisfies the following:

- The procedure runs in  $f(k)|x|^c$  for some computable function  $f$  and constant  $c$ .
- There exists a computable function  $g$  such that  $k' \leq g(k)$  for every input  $(x, k)$ .
- For all  $(x, k) \in \Pi_{YES}$ ,  $(x', k') \in \Pi'_{YES}$ .
- For all  $(x, k) \in \Pi_{NO}$ ,  $(x', k') \in \Pi'_{NO}$ .

**Definition 3.5** A randomized (one sided error) FPT reduction from a parameterized promise problem  $\Pi$  to a parameterized promise problem  $\Pi'$  is a randomized procedure that transforms  $(x, k)$  to  $(x', k')$  that satisfies the following:

- The procedure runs in  $f(k)|x|^c$  for some computable function  $f$  and constant  $c$  (on every randomness).
- There exists a computable function  $g$  such that  $k' \leq g(k)$  for every input  $(x, k)$ .
- For all  $(x, k) \in \Pi_{YES}$ ,  $\Pr[(x', k') \in \Pi'_{YES}] \geq 1/(f'(k)|x|^{c'})$  for some computable function  $f'$  and constant  $c'$ .
- For all  $(x, k) \in \Pi_{NO}$ ,  $\Pr[(x', k') \in \Pi'_{NO}] = 1$ .

Note that the above definition corresponds to the notion of *Reverse Unfaithful Random (RUR) reductions* in the classical world [Joh90]. The only difference (besides the allowed FPT running time) is that the above definition allows the probability that the YES case gets map to the YES case to be as small as  $1/(f'(k)\text{poly}(|x|))$ , whereas in the RUR reductions this can only be  $1/\text{poly}(|x|)$ . The reason is that, as we will see in Lemma 3.7 below, FPT algorithms can afford to repeat the reduction  $f'(k)\text{poly}(|x|)$  times, whereas polynomial time algorithms can only repeat  $\text{poly}(|x|)$  times.

We also consider randomized two-sided error FPT reductions, which are defined as follows.

**Definition 3.6** A randomized two sided error FPT reduction from a parameterized promise problem  $\Pi$  to a parameterized promise problem  $\Pi'$  is a randomized procedure that transforms  $(x, k)$  to  $(x', k')$  that satisfies the following:

- The procedure runs in  $f(k)|x|^c$  for some computable function  $f$  and constant  $c$  (on every randomness).
- There exists a computable function  $g$  such that  $k' \leq g(k)$  for every input  $(x, k)$ .
- For all  $(x, k) \in \Pi_{YES}$ ,  $\Pr[(x', k') \in \Pi'_{YES}] \geq 2/3$ .
- For all  $(x, k) \in \Pi_{NO}$ ,  $\Pr[(x', k') \in \Pi'_{NO}] \geq 2/3$ .

Note that this is not a generalization of the standard randomized FPT reduction (as defined in Definition 3.5), since the definition requires the success probabilities for the YES and NO cases to be constants independent of the parameter. In both cases, using standard techniques randomized FPT reductions, can be used to transform randomized FPT algorithms for  $\Pi'$  to randomized FPT algorithm for  $\Pi$ , as stated by the following lemma:

**Lemma 3.7** Suppose there exists a randomized (one sided/ two sided) error FPT reduction from a parameterized promise problem  $\Pi$  to a parameterized promise problem  $\Pi'$ . If there exists a randomized FPT algorithm  $\mathcal{A}$  for  $\Pi'$ , there there also exists a randomized FPT algorithm for  $\Pi$ .

*Proof.* We prove this for one sided error reductions, the other case follows using similar arguments.

Suppose there exists a randomized one sided error reduction from  $\Pi$  to  $\Pi'$ . Let  $f'(\cdot), c'$  be as in Definition 3.5. We consider the following subroutine. Given instance  $(x, k)$  of promise problem  $\Pi$ , we apply the randomized reduction on  $(x, k)$  to get instance  $(x', k')$  of promise problem  $\Pi'$ . We run  $\mathcal{A}$  on  $(x', k')$  repeatedly  $100 \log(f'(k)|x|^{c'})$  times, and output the majority of the outcomes.

If  $(x, k)$  is a YES instance, then with probability at least  $1/(f'(k)|x|^{c'})$ ,  $(x', k')$  is also a YES instance for  $\Pi'$ . Using Chernoff bound, conditioned on  $(x', k')$  being a YES instance, the majority of the outcomes is YES with probability at least  $1 - e^{-10 \log(f'(k)|x|^{c'})}$ . Therefore using union bound, the output of the above algorithm is YES with probability at least  $1/(f'(k)|x|^{c'}) - e^{-10 \log(f'(k)|x|^{c'})} \geq 1/2(f'(k)|x|^{c'})$ . Similarly, if  $(x, k)$  is a NO instance, then the subroutine outputs YES with probability at most  $e^{-10 \log(f'(k)|x|^{c'})}$ .

Equipped with the above subroutine, our algorithm is simply the following: given  $(x, k)$ , it runs the subroutine  $10f'(k)|x|^{c'}$  times. If at least one of the outcomes is YES, then the algorithm outputs YES, otherwise it outputs NO. Again we can analyze this using elementary probability. If  $(x, k)$  is a YES instance, then the algorithm outputs NO only if outcomes of all the trials is NO. Therefore, the algorithm outputs YES with probability at least  $1 - (1 - 1/2(f'(k)|x|^{c'}))^{10f'(k)|x|^{c'}} \geq 0.9$ . Conversely, if  $(x, k)$  is a NO instance, then by union bound, the algorithm outputs NO with probability at least  $1 - 10f'(k)|x|^{c'}e^{-10 \log(f'(k)|x|^{c'})} \geq 0.9$ . Finally, if  $\mathcal{A}$  is FPT, then the running time of the proposed algorithm is also FPT. Hence the claim follows<sup>13</sup>.  $\square$

Since the conclusion of the above proposition holds for both types of randomized reductions, we will not be distinguishing between the two types in the rest of the paper.

### 3.2 Minimum Distance Problem

In this subsection, we define the fixed-parameter variant of the minimum distance problem and other relevant parameterized problems. We actually define them as gap problems – as later in the paper, we show the constant inapproximability of these problems.

For every  $\gamma \geq 1$ , we define the  $\gamma$ -gap minimum distance problem<sup>14</sup> as follows:

<p><math>\gamma</math>-Gap Minimum Distance Problem (GAPMDP<math>_{\gamma}</math>)</p> <p><b>Input:</b> A matrix <math>\mathbf{A} \in \mathbb{F}_2^{n \times m}</math> and a positive integer <math>k \in \mathbb{N}</math></p> <p><b>Parameter:</b> <math>k</math></p> <p><b>Question:</b> Distinguish between the following two cases:</p> <ul style="list-style-type: none"> <li>• (YES) there exists <math>\mathbf{x} \in \mathbb{F}_2^m \setminus \{\mathbf{0}\}</math> such that <math>\ \mathbf{Ax}\ _0 \leq k</math></li> <li>• (NO) for all <math>\mathbf{x} \in \mathbb{F}_2^m \setminus \{\mathbf{0}\}</math>, <math>\ \mathbf{Ax}\ _0 &gt; \gamma \cdot k</math></li> </ul>
---

<sup>13</sup>For the case of 2-sided error, we change the final step of the algorithm as follows; we invoke the subroutine  $O(\log 1/\delta)$ -times (where  $\delta$  is a constant) and again output the majority of the outcomes. The guarantees again follow by a Chernoff bound argument.

<sup>14</sup>In the parameterized complexity literature, this problem is referred to as the  $k$ -Even set problem [DFVW99] and the input to the problem is (equivalently) given through the parity-check matrix, instead of the generator matrix as described in this paper.



Next, for every  $\gamma \geq 1$ , we define the  $\gamma$ -gap maximum likelihood decoding problem<sup>15</sup> as follows:

$\gamma$ -Gap Maximum Likelihood Decoding Problem (GAPMLD $_{\gamma}$ )

**Input:** A matrix  $\mathbf{A} \in \mathbb{F}_2^{n \times m}$ , a vector  $\mathbf{y} \in \mathbb{F}_2^n$  and a positive integer  $k \in \mathbb{N}$

**Parameter:**  $k$

**Question:** Distinguish between the following two cases:

- (YES) there exists  $\mathbf{x} \in \mathcal{B}_m(\mathbf{0}, k)$  such that  $\mathbf{Ax} = \mathbf{y}$
- (NO) for all  $\mathbf{x} \in \mathcal{B}_m(\mathbf{0}, \gamma k)$ ,  $\mathbf{Ax} \neq \mathbf{y}$

For brevity, we shall denote the exact version (i.e., GAPMLD $_1$ ) of the problem as MLD. Finally, we introduce a “sparse” version of the GAPMLD problem called the sparsest nearest codeword problem, and later in the paper we show a reduction from GAPMLD to this problem, followed by a reduction from this problem to GAPMDP. Formally, for every  $\gamma \geq 1$ , we define the  $\gamma$ -gap sparsest nearest codeword problem as follows:

$\gamma$ -Gap Sparse Nearest Codeword Problem (GAPSNC $_{\gamma}$ )

**Input:** A matrix  $\mathbf{A} \in \mathbb{F}_2^{n \times m}$ , a vector  $\mathbf{y} \in \mathbb{F}_2^n$  and a positive integer  $k \in \mathbb{N}$

**Parameter:**  $k$

**Question:** Distinguish between the following two cases:

- (YES) there exists  $\mathbf{x} \in \mathcal{B}_m(\mathbf{0}, k)$  such that  $\|\mathbf{Ax} - \mathbf{y}\|_0 \leq k$
- (NO) for all  $\mathbf{x} \in \mathbb{F}_2^m$ ,  $\|\mathbf{Ax} - \mathbf{y}\|_0 > \gamma \cdot k$

### 3.3 Shortest Vector Problem and Nearest Vector Problem

In this subsection, we define the fixed-parameter variants of the shortest vector and nearest vector problems. As in the previous subsection, we define them as gap problems, for the same reason that later in the paper, we show the constant inapproximability of these two problems.

Fix  $p \in \mathbb{R}_{\geq 1}$ . For every  $\gamma \geq 1$ , we define the  $\gamma$ -gap shortest vector problem in the  $\ell_p$ -norm<sup>16</sup> as follows:

$\gamma$ -Gap Shortest Vector Problem (GAPSVP $_{p,\gamma}$ )

**Input:** A matrix  $\mathbf{A} \in \mathbb{Z}^{n \times m}$  and a positive integer  $k \in \mathbb{N}$

**Parameter:**  $k$

<sup>15</sup>The maximum likelihood decoding problem is also equivalently known in the literature as the nearest codeword problem.

<sup>16</sup>Note that we define GAPNVP and GAPSVP problems in terms of  $\ell_p^p$ , whereas traditionally, it is defined in terms of  $\ell_p$ . However, it is sufficient for us to work with the  $\ell_p^p$  variant, since an  $\alpha$ -factor inapproximability in  $\ell_p^p$  translates to an  $\alpha^{1/p}$ -factor inapproximability in the  $\ell_p$  norm, for any  $\alpha \geq 1$

**Question:** Distinguish between the following two cases:

- (YES) there exists  $\mathbf{x} \in \mathbb{Z}^m \setminus \{\mathbf{0}\}$  such that  $\|\mathbf{Ax}\|_p^p \leq k$
- (NO) for all  $\mathbf{x} \in \mathbb{Z}^m \setminus \{\mathbf{0}\}$ ,  $\|\mathbf{Ax}\|_p^p > \gamma \cdot k$

For every  $\gamma \geq 1$ , we define the  $\gamma$ -gap nearest vector problem in the  $\ell_p$ -norm as follows:

$\gamma$ -Gap Nearest Vector Problem (GAPNVP <sub>$p,\gamma$</sub> )

**Input:** A matrix  $\mathbf{A} \in \mathbb{Z}^{n \times m}$ , vector  $\mathbf{y} \in \mathbb{Z}^n$  and a positive integer  $k \in \mathbb{N}$

**Parameter:**  $k$

**Question:** Distinguish between the following two cases:

- (YES) there exists  $\mathbf{x} \in \mathbb{Z}^m$  such that  $\|\mathbf{Ax} - \mathbf{y}\|_p^p \leq k$
- (NO) for all  $\mathbf{x} \in \mathbb{Z}^m$ ,  $\|\mathbf{Ax} - \mathbf{y}\|_p^p > \gamma \cdot k$

### 3.4 CSPs and Parameterized Inapproximability Hypothesis

In this section, we will formally state the Parameterized Inapproximability Hypothesis (PIH). To do so, we first have to define 2CSP and its corresponding gap problem, starting with the former:

**Definition 3.8 (2CSP)** An instance  $\Gamma$  of 2CSP consists of

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

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

The gap problem for 2CSP can then be defined as follows:

$\varepsilon$ -Gap 2CSP (GAP2CSP <sub>$\varepsilon$</sub> )

**Input:** A 2CSP instance  $\Gamma = (G = (V, E), \Sigma, \{C_{uv}\}_{(u,v) \in E})$ .

**Parameter:**  $|V|$

**Question:** Distinguish between the following two cases:

- (YES)  $\text{val}(\psi) = 1$ .
- (NO)  $\text{val}(\psi) < 1 - \varepsilon$ .

Note that, when  $\varepsilon = 0$ , GAP2CSP<sub>0</sub> is simply asking whether the input instance  $\Gamma$  is fully satisfiable. It is easy to see that this problem generalizes the  $k$ -Clique problem, which is well known to be

$W[1]$ -hard; in other words,  $\text{GAP2CSP}_0$  is  $W[1]$ -hard, and hence does not admit FPT algorithm unless  $W[1] = \text{FPT}$ . It is believed that  $\text{GAP2CSP}_\varepsilon$  remains  $W[1]$ -hard even for some constant  $\varepsilon > 0$ ; this belief has recently been formalized by Lokshtanov et al. [LRSZ17] as the *Parameterized Inapproximability Hypothesis (PIH)*. For the purpose of this work, we will use an even weaker version of the hypothesis than Lokshtanov et al.’s. Namely, we will only assume that  $\text{GAP2CSP}_\varepsilon$  is not in FPT for some  $\varepsilon > 0$  (rather than assuming that it is  $W[1]$ -hard), as stated below.

**Hypothesis 3.9 (Parameterized Inapproximability Hypothesis (PIH) [LRSZ17])** *There exists  $\varepsilon > 0$  such that there is no randomized FPT algorithm for  $\text{GAP2CSP}_\varepsilon$ .*

### 3.4.1 Relation to Gap Exponential Time Hypothesis

There are several supporting evidences for PIH. One such evidence is that it follows from the Gap Exponential Time Hypothesis hypothesis, which can be stated as follows.

**Hypothesis 3.10 (Randomized Gap Exponential Time Hypothesis (Gap-ETH) [Din16, MR16])** *There exist constants  $\varepsilon, \delta > 0$  such that any randomized algorithm that, on input a 3CNF formula  $\varphi$  on  $n$  variables and  $O(n)$  clauses, can distinguish between  $\text{SAT}(\varphi) = 1$  and  $\text{SAT}(\varphi) < 1 - \varepsilon$ , with probability at least  $2/3$ , must run in time at least  $2^{\delta n}$ .*

Gap-ETH itself is a strengthening of the Exponential Time Hypothesis (ETH) [IP01, IPZ01], which states that deciding whether a 3CNF formula is satisfiable cannot be done in subexponential time. We remark here that Gap-ETH would follow from ETH if a linear-size PCP exists; unfortunately, no such PCP is known yet, with the shortest known PCP having quasi-linear size [BS08, Din07]. It has also recently been shown that Gap-ETH would follow from ETH if we assume certain “smoothness” conditions on instances from ETH [App17]. We will not discuss the evidences supporting Gap-ETH in details here; we refer interested readers to [Din16].

As stated earlier, it is known that PIH follows from Gap-ETH (see e.g. [DM18, CFM17]<sup>17</sup>). Unfortunately, the proofs in literature so far have been somewhat complicated, since the previous works on the topic (e.g. [DM18, CCK<sup>+</sup>17, Man17]) put emphasis on achieving as large a factor hardness of approximation for 2CSP as possible. On the other hand, some constant inapproximability factor strictly greater than zero suffices to show PIH. For this regime, there is a (folklore) proof that is much simpler than those in the literature. Since we are not aware of this proof being fully written down anywhere, we provide it in Appendix A. Note that this proof also yields a running time lower bound of  $T(k) \cdot |\Sigma|^{\Omega(k)}$  of solving  $\text{GAP2CSP}_\varepsilon$  for some  $\varepsilon > 0$ ; this running time lower bound is better than those provided by the aforementioned previous proofs and is essentially optimal, since one can solve 2CSP (even exactly) in time  $|\Sigma|^{O(k)}$ .

## 3.5 Error-Correcting Codes

An error correcting code  $C$  over alphabet  $\Sigma$  is a function  $C : \Sigma^m \rightarrow \Sigma^h$  where  $m$  and  $h$  are positive integers which are referred to as the *message length* and *block length* of  $C$  respectively. Intuitively, the function  $C$  encodes an original message of length  $m$  to an encoded message of length  $h$ . The *distance* of a code, denoted by  $d(C)$ , is defined as  $\min_{x \neq y \in \Sigma^m} \|C(x) - C(y)\|_0$ , i.e., the number of coordinates on

<sup>17</sup>Note that, in [CFM17], 2CSP is referred to as *Maximum Colored Subgraph Isomorphism*. Note also that the hardness proof of [CFM17] relies heavily on the parameterized inapproximability of Densest  $k$ -Subgraph from [CCK<sup>+</sup>17], which in turns relies on the reduction and the main lemma from [Man17].

which  $C(x)$  and  $C(y)$  disagree. We also define the systematicity of a code as follows: Given  $s \in \mathbb{N}$ , a code  $C : \Sigma^m \rightarrow \Sigma^h$  is  $s$ -systematic if there exists a size- $s$  subset of  $[h]$ , which for convenience we identify with  $[s]$ , such that for every  $x \in \Sigma^s$  there exists  $w \in \Sigma^m$  in which  $x = C(w) \upharpoonright_{[s]}$ . We use the shorthand  $[h, m, d]_{|\Sigma|}$  to denote a code of message length  $m$ , block length  $h$ , and distance  $d$ .

Additionally, we will need the following existence and efficient construction of BCH codes for every message length and distance parameter.

**Theorem 3.11 (BCH Code [Hoc59, BR60])** *For any choice of  $h, d \in \mathbb{N}$  such that  $h + 1$  is a power of two and that  $d \leq h$ , there exists a linear code over  $\mathbb{F}_2$  with block length  $h$ , message length  $h - \left\lceil \frac{d-1}{2} \right\rceil \cdot \log(h + 1)$  and distance  $d$ . Moreover, the generator matrix of this code can be computed in  $\text{poly}(h)$  time.*

Finally, we define the tensor product of codes which will be used later in the paper. Consider two linear codes  $C_1 \subseteq \mathbb{F}_2^m$  (generated by  $\mathbf{G}_1 \in \mathbb{F}_2^{m \times m'}$ ) and  $C_2 \subseteq \mathbb{F}_2^n$  (generated by  $\mathbf{G}_2 \in \mathbb{F}_2^{n \times n'}$ ). Then the tensor product of the two codes  $C_1 \otimes C_2 \subseteq \mathbb{F}_2^{m \times n}$  is defined as

$$C_1 \otimes C_2 = \{ \mathbf{G}_1 \mathbf{X} \mathbf{G}_2^\top \mid \mathbf{X} \in \mathbb{F}_2^{m' \times n'} \}.$$

## 4 Parameterized Intractability of GAPMLD

In this section, we will show the parameterized intractibility of GAPMLD as stated below.

**Theorem 4.1** *Assuming PIH, there is no randomized FPT algorithm for  $\text{GAPMLD}_\gamma$  for any  $\gamma \geq 1$ .*

The proof proceeds in two steps. First, we reduce  $\text{GAP2CSP}$  to  $\text{GAPMLD}_\gamma$  for some  $\gamma > 1$  in Section 4.1. Then, we boost the gap in the GAPMLD problem in Section 4.2.

### 4.1 Reducing GAP2CSP to GAPMLD

In this subsection, we show an FPT reduction from  $\text{GAP2CSP}$  to  $\text{GAPMLD}$ , as stated below:

**Lemma 4.2** *For any  $\varepsilon \geq 0$ , there is an FPT reduction from  $\text{GAP2CSP}_\varepsilon$  to  $\text{GAPMLD}_\gamma$  where  $\gamma = 1 + \varepsilon/3$ .*

Combining the reduction with PIH, we obtain the following hardness of GAPMLD:

**Theorem 4.3** *Assuming PIH, for some  $\gamma > 1$ , there is no randomized FPT algorithm for  $\text{GAPMLD}_\gamma$ .*

*Proof of Lemma 4.2.* Let  $\Gamma = (G = (V, E), \Sigma, \{C_{uv}\}_{(u,v) \in E})$  be the input for  $\text{GAP2CSP}_\varepsilon$ . We produce an instance  $(\mathbf{A}, \mathbf{y}, k)$  of  $\text{GAPMLD}_\gamma$  as follows. Let  $n = |V| + |E| + 2|E||\Sigma|$  and  $m = |V||\Sigma| + \sum_{(u,v) \in E} |C_{uv}|$ ; our matrix  $\mathbf{A}$  will be of dimension  $(n \times m)$ . For convenience, let us label the first  $|V|$  rows of  $\mathbf{A}$  by the vertices  $u \in V$ , the next  $|E|$  rows by the edges  $e \in E$ , and the last  $2|E||\Sigma|$  rows by a tuple  $(e, \sigma, b) \in E \times \Sigma \times \{0, 1\}$ . Furthermore, we label the first  $|V||\Sigma|$  columns of  $\mathbf{A}$  by  $(u, \sigma_u)$  where  $u \in V$  and  $\sigma_u \in \Sigma$ , and the rest of the columns by  $(e, \sigma_0, \sigma_1)$  where  $e \in E$  and  $(\sigma_0, \sigma_1) \in C_e$ . The entries of our matrix  $\mathbf{A} \in \mathbb{F}_2^{n \times m}$  can now be defined as follows.

- For each column of the form  $(u, \sigma_u)$ , let the following entries be one:
  - $\mathbf{A}_{u, (u, \sigma_u)}$ .
  - $\mathbf{A}_{((u, v), \sigma_u, 0), (u, \sigma_u)}$  for every  $v \in V$  such that  $(u, v) \in E$ .
  - $\mathbf{A}_{((v, u), \sigma_u, 1), (u, \sigma_u)}$  for every  $v \in V$  such that<sup>18</sup>  $(v, u) \in E$ .

<sup>18</sup>For the clarity of presentation, we assume here that  $G$  is directed.

The rest of the entries are set to zero.

- For each column of the form  $(e, \sigma_0, \sigma_1)$ , let the following three entries be one:  $\mathbf{A}_{e, (e, \sigma_0, \sigma_1)}$ ,  $\mathbf{A}_{(e, \sigma_0, 0), (e, \sigma_0, \sigma_1)}$  and  $\mathbf{A}_{(e, \sigma_1, 1), (e, \sigma_0, \sigma_1)}$ . The rest of the entries are set to zero.

Finally, we set  $\mathbf{y} = \mathbf{1}_{|V|+|E|} \circ \mathbf{0}_{2|E||\Sigma|}$  and  $k = |V| + |E|$ .

**Parameter Dependency.** The new parameter  $k$  is simply  $|V| + |E| = O(|V|^2)$ .

Before we move on to prove the completeness and soundness of the reduction, let us state an observation that will be useful in the analysis:

**Observation 4.4** For any  $\mathbf{x} \in \mathbb{F}_2^m$ , the following properties hold.

- For every row of the form  $u \in V$ , we have

$$(\mathbf{Ax})_u = \sum_{\sigma \in \Sigma} \mathbf{x}_{(u, \sigma)}. \quad (2)$$

- For every row of the form  $e \in E$ , we have

$$(\mathbf{Ax})_e = \sum_{(\sigma_0, \sigma_1) \in C_e} \mathbf{x}_{(e, \sigma_0, \sigma_1)}. \quad (3)$$

- For every row of the form  $(e = (u_0, u_1), \sigma, b)$ , we have

$$(\mathbf{Ax})_{(e, \sigma, b)} = \mathbf{x}_{(u_b, \sigma)} + \sum_{\substack{(\sigma_0, \sigma_1) \in C_e \\ \sigma_b = \sigma}} \mathbf{x}_{(e, \sigma_0, \sigma_1)}. \quad (4)$$

Note that the above observation follows trivially from our definition of  $\mathbf{A}$ .

**Completeness.** Suppose that  $\text{val}(\Gamma) = 1$ . That is, there exists an assignment  $\psi : V \rightarrow \Sigma$  that satisfies all the edges. We define the vector  $\mathbf{x} \in \mathbb{F}_2^m$  as follows.

- For each  $(u, \sigma)$ , let  $\mathbf{x}_{(u, \sigma)} = \mathbb{1}[\psi(u) = \sigma]$
- For each  $(e = (u, v), \sigma_0, \sigma_1)$ , let  $\mathbf{x}_{(e, \sigma_0, \sigma_1)} = \mathbb{1}[(\psi(u) = \sigma_0) \wedge (\psi(v) = \sigma_1)]$ .

We claim that  $\mathbf{Ax} = \mathbf{y}$ . To see that this is the case, consider the following three cases of rows.

- For each row of the form  $u \in V$ , we have

$$(\mathbf{Ax})_u \stackrel{(2)}{=} \sum_{\sigma \in \Sigma} \mathbf{x}_{(u, \sigma)} = \sum_{\sigma \in \Sigma} \mathbb{1}[\psi(u) = \sigma] = 1.$$

- For each row of the form  $e = (u, v) \in E$ , we have

$$(\mathbf{Ax})_e \stackrel{(3)}{=} \sum_{(\sigma_0, \sigma_1) \in C_e} \mathbf{x}_{(e, \sigma_0, \sigma_1)} = \sum_{(\sigma_0, \sigma_1) \in C_e} \mathbb{1}[(\psi(u) = \sigma_0) \wedge (\psi(v) = \sigma_1)] = 1.$$

Note that, in the last equality, we use the fact that the edge  $e$  is satisfied (i.e.  $(\psi(u), \psi(v)) \in C_e$ ).

- For each row  $(e = (u_0, u_1), \sigma, b) \in E \times \Sigma \times \{0, 1\}$ , we have

$$(\mathbf{Ax})_{(e,\sigma,b)} \stackrel{(4)}{=} \mathbf{x}_{(u_b,\sigma)} + \sum_{\substack{(\sigma_0,\sigma_1) \in C_e \\ \sigma_b = \sigma}} \mathbf{x}_{(e,\sigma_0,\sigma_1)} = \mathbb{1}[\psi(u_b) = \sigma] + \mathbb{1}[\psi(u_b) = \sigma] = 0,$$

where the second equality uses the fact that exactly one of  $\mathbf{x}_{(e,\sigma_0,\sigma_1)}$  is not zero among  $(\sigma_0, \sigma_1) \in C_e$ , which is  $\mathbf{x}_{(e,\psi(u_0),\psi(u_1))}$ .

Hence,  $\mathbf{Ax}$  is indeed equal to  $\mathbf{y}$ . Finally, observe that  $\|\mathbf{x}\|_0 = |V| + |E|$  as desired.

**Soundness.** We will prove this by contrapositive. Suppose that the constructed instance  $(\mathbf{A}, \mathbf{y}, k)$  is not a NO instance of  $\text{GAPMLD}_\gamma$ , i.e., for some  $x \in \mathcal{B}_m(\mathbf{0}, \gamma k)$ ,  $\mathbf{Ax} = \mathbf{y}$ . We will show that  $\Gamma$  is not a NO instance of  $\text{GAP2CSP}_\varepsilon$ , i.e., that  $\text{val}(\Gamma) \geq 1 - \varepsilon$ .

First, for each vertex  $u \in V$ , let  $S_u = \{\sigma \in \Sigma \mid \mathbf{x}_{(u,\sigma)} = 1\}$  and, for each  $e \in E$ , let  $T_e = \{(\sigma_0, \sigma_1) \in C_e \mid \mathbf{x}_{(e,\sigma_0,\sigma_1)} = 1\}$ . From (2) and from  $(\mathbf{Ax})_u = \mathbf{y}_u = 1$ , we can conclude that  $|S_u| \geq 1$  for all  $u \in V$ . Similarly, from (3) and from  $(\mathbf{Ax})_e = \mathbf{y}_e = 1$ , we have  $|T_e| \geq 1$  for all  $e \in E$ .

We define an assignment  $\psi : V \rightarrow \Sigma$  of  $\Gamma$  by setting  $\psi(u)$  to be an arbitrary element of  $S_u$  for all  $u \in V$ . We will show that  $\text{val}(\psi) \geq 1 - \varepsilon$ , which indeed implies  $\text{val}(\Gamma) \geq 1 - \varepsilon$ .

To do so, let  $E_{\text{unique}}$  denote the set of  $e \in E$  such that  $|T_e| = 1$ . Notice that

$$\begin{aligned} (\gamma - 1)k &\geq \|\mathbf{x}\|_0 - k & (5) \\ &= \sum_{u \in V} |S_u| + \sum_{e \in E} |T_e| - k \\ &= \sum_{u \in V} (|S_u| - 1) + \sum_{e \in E} (|T_e| - 1) \\ &\geq \sum_{e \in E \setminus E_{\text{unique}}} (|T_e| - 1) \\ &\geq |E \setminus E_{\text{unique}}|, & (6) \end{aligned}$$

which implies that  $|E_{\text{unique}}| \geq |E| - (\gamma - 1)k$ , which, from our choice of  $\gamma$ , is at least  $|E| - \frac{\varepsilon k}{3} \geq (1 - \varepsilon)|E|$ . Note that the last inequality follows from  $|V| \leq 2|E|$ , which can be assumed w.l.o.g.

Since  $|E_{\text{unique}}| \geq (1 - \varepsilon)|E|$ , to show that  $\text{val}(\psi) \geq 1 - \varepsilon$ , it suffices to show that  $\psi$  satisfies every edge in  $E_{\text{unique}}$ . To see that this is the case, let  $e = (u, v)$  be any edge in  $E_{\text{unique}}$ . Let  $(\sigma_0^*, \sigma_1^*)$  be the only element of  $T_e$ . Observe that, from (4) with  $\sigma = \psi(u)$  and  $b = 0$  and from  $(\mathbf{Ax})_{e,\sigma,b} = \mathbf{y}_{e,\sigma,b} = 0$ , we can conclude that  $\sigma_0^* = \psi(u)$ . Similarly, we can conclude that  $\sigma_1^* = \psi(v)$ . As a result,  $(\psi(u), \psi(v))$  must be in  $C_e$ , meaning that  $\psi$  satisfies  $e$ , which completes our proof.  $\square$

## 4.2 Gap Amplification

In this subsection, we describe the gap amplification step for  $\text{GAPMLD}$ . Towards that end, we define a *composition* operation  $\oplus$  on  $\text{GAPMLD}$  instances, which can be used to efficiently amplify the gap to any constant factor with only a polynomial blowup in the instance size.

**The Composition Operator:** Consider two  $\text{GAPMLD}$  instances given by coefficient matrices  $\mathbf{A} \in \mathbb{F}_2^{u \times v}$ ,  $\mathbf{B} \in \mathbb{F}_2^{u' \times v'}$  and non-zero target vectors  $\mathbf{z} \in \mathbb{F}_2^u$ ,  $\mathbf{z}' \in \mathbb{F}_2^{u'}$ . Their composition is a  $\text{GAPMLD}$  instance  $(\mathbf{C}, \mathbf{w}, k_2(k_1 + 1)) = (\mathbf{A}, \mathbf{z}, k_1) \oplus (\mathbf{B}, \mathbf{z}', k_2)$ , given by coefficient matrix  $\mathbf{C} \in \mathbb{F}_2^{(u'+uv') \times (v'+vv')}$ ,

and target vector  $\mathbf{w} \in \mathbb{F}_2^{u'+uv'}$ . They are constructed as follows. Consider the following partition of  $[u' + uv']$  into blocks  $S_0, S_1, \dots, S_{v'}$ , where  $S_0$  consists of the first  $u'$  coordinates, and for every  $i \in [v']$  the block  $S_i$  consists of coordinates  $u' + (i-1)v + 1, \dots, u' + iv$ . Similarly, consider the partition  $T_0, T_1, \dots, T_{v'}$  of  $[v' + vv']$ , where  $T_0$  consists of the first  $v'$  coordinates, and blocks  $T_1, \dots, T_{v'}$  is the contiguous equipartition of the remaining coordinates. Furthermore, for any choice of  $S_i, T_j$ , let  $\mathbf{C}_{S_i, T_j}$  be the sub-matrix of  $\mathbf{C}$  consisting of rows and columns indexed by  $S_i$  and  $T_j$  respectively. Extending the notation to vectors  $\mathbf{x} \in \mathbb{F}_2^{v'+vv'}$ , we use  $\mathbf{x}^i$  to denote the restriction of the vectors along the coordinates in  $T_i$ . Now we describe the construction of  $\mathbf{C}$  and  $\mathbf{w}$  in a row block wise order.

1. **Row block  $S_0$ :** We set  $\mathbf{C}_{S_0, T_0} = \mathbf{B}$ , and  $\mathbf{C}_{S_0, T_i} = \mathbf{0}_{u' \times v}$  for every  $i \in [v']$ . Additionally, we set the corresponding sub-vector  $\mathbf{w}_{S_0} = \mathbf{z}'$ .
2. **Row block  $S_i$  (for  $i \geq 1$ ):** Here  $\mathbf{C}_{S_i, T_0}$  contains  $\mathbf{z}$  in the  $i^{\text{th}}$  column and is zero everywhere else. The sub-matrix  $\mathbf{C}_{S_i, T_i}$  is set to the coefficient matrix  $\mathbf{A}$ , and for all blocks  $j \notin \{0, i\}$ , we set the sub-matrices  $\mathbf{C}_{S_i, T_j} = \mathbf{0}_{u \times v}$ . Finally, for the target vector we set the corresponding block  $\mathbf{w}_{S_i} = \mathbf{0}_u$ .

The intuition underlying the above construction is as follows. For simplicity, let  $k = k_1 = k_2$ . Consider the row block  $S_0$  in  $(\mathbf{C}, \mathbf{w}, k^2 + k)$ . By construction, the set of indices of non-zero column of  $\mathbf{C}$  along  $S_0$  is exactly  $T_0$ , and by construction  $\mathbf{C}_{S_0, T_0} = \mathbf{B}$ . In particular, for any vector  $\mathbf{x} \in \mathbb{F}^{v'+vv'}$  satisfying  $\mathbf{C}\mathbf{x} = \mathbf{w}$ , the constraints along rows in  $S_0$  enforce  $\mathbf{B}\mathbf{x}^0 = \mathbf{z}'$ , and hence  $\|\mathbf{x}^0\|_0 \geq k$ . Similarly, for any  $i \in [v']$ , the non-zero columns blocks corresponding to the row block  $S_0$  are  $T_0$  and  $T_i$ . Since  $\mathbf{C}_{S_i, T_i} = \mathbf{A}$ , the row block  $S_i$  forces the constraints  $\mathbf{A}\mathbf{x}^i = x_i^0 \mathbf{z}$ . Therefore, for any  $i \in [v]$  such that  $x_i^0 \neq 0$ , the sub-vector  $\mathbf{x}^i$  must satisfy  $\mathbf{A}\mathbf{x}^i = \mathbf{z}$ , and hence  $\|\mathbf{x}^i\|_0 \geq k$ . Since this must happen for every  $i \in [v']$  such that  $x_i^0 \neq 0$ , the overall Hamming weight of the vector  $\mathbf{x}$  must be at least  $k + k^2$ . These observations are made formal in the following lemma.

**Lemma 4.5** Consider coefficient matrices  $\mathbf{A} \in \mathbb{F}_2^{u \times v}$ ,  $\mathbf{B} \in \mathbb{F}_2^{u' \times v'}$ , and target vectors  $\mathbf{z} \in \mathbb{F}_2^v$ ,  $\mathbf{z}' \in \mathbb{F}_2^{v'}$  corresponding to GAPMLD $_\gamma$  instances  $(\mathbf{A}, \mathbf{z}, k_1)$  and  $(\mathbf{B}, \mathbf{z}', k_2)$ . Then  $(\mathbf{C}, \mathbf{w}, k_2(k_1 + 1)) = (\mathbf{A}, \mathbf{z}, k_1) \oplus (\mathbf{B}, \mathbf{z}', k_2)$  as constructed above satisfies the following properties:

- If  $(\mathbf{A}, \mathbf{z}, k_1)$  and  $(\mathbf{B}, \mathbf{z}', k_2)$  are YES instances, then there exists  $\mathbf{x} \in \mathcal{B}_{v'+vv'}(\mathbf{0}, k_2 + k_1 k_2)$  such that  $\mathbf{C}\mathbf{x} = \mathbf{w}$ .
- If  $(\mathbf{A}, \mathbf{z}, k_1)$  and  $(\mathbf{B}, \mathbf{z}', k_2)$  are NO instances, then for every  $\mathbf{x} \in \mathcal{B}_{v'+vv'}(\mathbf{0}, \gamma k_2 + \gamma^2 k_1 k_2)$  we have  $\mathbf{C}\mathbf{x} \neq \mathbf{w}$ .

*Proof.* Suppose  $(\mathbf{A}, \mathbf{z}, k_1)$  and  $(\mathbf{B}, \mathbf{z}', k_2)$  are NO instances. Consider any vector  $\mathbf{x} \in \mathbb{F}_2^{v'+vv'}$  satisfying  $\mathbf{C}\mathbf{x} = \mathbf{w}$ . By construction, it must satisfy the constraints corresponding to row-block  $S_0$  i.e.,  $\mathbf{B}\mathbf{x}^0 = \mathbf{z}'$ , and therefore  $\|\mathbf{x}^0\|_0 > \gamma k_2$ . Let  $\Omega = \{i \in [v'] : x_i^0 \neq 0\}$  be the set of non-zero indices in  $\mathbf{x}^0$ , and fix any  $i \in \Omega$ . It is easy to see that along row-block  $S_i$  the only variables acting on non-zeros blocks are the variable  $x_i^0$  (acting on the sub-matrix  $\mathbf{C}_{S_i, T_0}$ ) and the vector  $\mathbf{x}^i$  (acting on the sub-matrix  $\mathbf{C}_{S_i, T_i} = \mathbf{A}$ ). All together, this induces constraints  $\mathbf{A}\mathbf{x}^i = x_i^0 \cdot \mathbf{z}$ . Since  $x_i^0 \neq 0$ , the vector  $\mathbf{x}^i$  must satisfy  $\mathbf{A}\mathbf{x}^i = \mathbf{z}$  and therefore  $\|\mathbf{x}^i\|_0 > \gamma k_1$ . Moreover, since this must happen for every  $i \in \Omega$  and  $|\Omega| > \gamma k_2$ , we have  $\|\mathbf{x}\|_0 > \gamma k_2 + \gamma^2 k_1 k_2$  as desired.

It remains to be shown if  $(\mathbf{A}, \mathbf{z}, k_1)$  and  $(\mathbf{B}, \mathbf{z}', k_2)$  are YES instances, then there exists a satisfying solution with Hamming weight at most  $k_2(k_1 + 1)$ . Let  $\mathbf{a} \in \mathcal{B}_v(\mathbf{0}, k_1)$  and  $\mathbf{b} \in \mathcal{B}_{v'}(\mathbf{0}, k_2)$  be such

that  $\mathbf{A}\mathbf{a} = \mathbf{z}$  and  $\mathbf{B}\mathbf{b} = \mathbf{z}'$ . We construct  $\mathbf{x} \in \mathbb{F}_2^{v'+vv'}$  as follows. We set  $\mathbf{x}^0 = \mathbf{b}$  and for every  $i \in [v]$ , we set  $\mathbf{x}^i = b_i \cdot \mathbf{a}$ . It can be easily verified that the vector  $\mathbf{x}$  is a solution to  $(\mathbf{C}, \mathbf{w})$ . Indeed, along row-block  $S_0$ , we have  $\mathbf{B}\mathbf{x}^0 = \mathbf{B}\mathbf{b} = \mathbf{z}$ . Furthermore, if  $x_i^0 = 1$  the corresponding row-block  $S_i$  satisfies  $\mathbf{A}\mathbf{x}^i = \mathbf{A}\mathbf{a} = \mathbf{z}$ . Finally, if  $x_i^0 = 0$ , the corresponding row-block evaluates to  $\mathbf{A}\mathbf{x}^i = \mathbf{0}_u$  which is trivially satisfied by the given construction. Since the vector  $\mathbf{x}$  has Hamming weight at most  $k_2(k_1 + 1)$ , the claim follows.  $\square$

The above lemma can be used to amplify the gap for  $\text{GAPMLD}_\gamma$  instances, as shown by the following corollary.

**Corollary 4.6** *For all choices of  $\gamma > 1, \eta > 0$  and  $k \geq (\gamma^\eta - 1)^{-1}$ , there exists an FPT reduction from  $k$ - $\text{GAPMLD}_\gamma$  to  $(k^2 + k)$ - $\text{GAPMLD}_{\gamma^{2-\eta}}$ .*

*Proof.* Let  $(\mathbf{A}, \mathbf{z}, k)$  be a  $\text{GAPMLD}_\gamma$  instance, where  $\mathbf{A} \in \mathbb{F}_2^{u \times v}$ . Let  $(\mathbf{C}, \mathbf{w}, k^2 + k) = (\mathbf{A}, \mathbf{z}, k) \oplus (\mathbf{A}, \mathbf{z}, k)$  be constructed as above. If  $(\mathbf{A}, \mathbf{z}, k)$  is a YES instance, then there exists  $\mathbf{x} \in \mathcal{B}_{v+uv}(\mathbf{0}, k + k^2)$  of Hamming such that  $\mathbf{C}\mathbf{x} = \mathbf{w}$ . Conversely, if  $(\mathbf{A}, \mathbf{z}, k)$  is a NO instance, then every solution to  $\mathbf{C}\mathbf{x} = \mathbf{w}$  has Hamming weight at least  $\gamma^2 k^2 + \gamma k \geq \gamma^{2-\eta}(k^2 + k)$  (by our choice of  $k$ ). Hence the claim follows.  $\square$

Theorem 4.1 is a direct consequence of the above lemma:

*Proof of Theorem 4.1.* Consider the hardness of  $\text{GAPMLD}$  as given in Theorem 4.3. For any  $\gamma \geq 1 + \varepsilon/3$ , applying Corollary 4.6  $i \geq \lceil \log(\gamma) / \log(1 + \varepsilon/3) \rceil$  times on  $\text{GAPMLD}_{(1+\varepsilon/3)}$  instance can boost the gap to  $\text{GAPMLD}_\gamma$ . Since all the steps in the reduction are FPT, the claim follows.  $\square$

## 5 Parameterized Intractability of Minimum Distance Problem

Next, we will prove our main theorem regarding parameterized intractability of  $\text{GAPMDP}$ :

**Theorem 5.1** *Assuming PIH, there is no randomized FPT algorithm for  $\text{GAPMDP}_\gamma$  for any  $\gamma \geq 1$ .*

This again proceeds in two steps. First, we provide an FPT reduction from  $\text{GAPMLD}$  to  $\text{GAPSNC}$  in Section 5.1. Next, we formalize the definition of Sparse Covering Codes (SCC), prove their existence, and show how to use them to reduce  $\text{GAPSNC}$  to  $\text{GAPMDP}$  in Section 5.2.

### 5.1 Reducing $\text{GAPMLD}$ to $\text{GAPSNC}$

In this subsection, we show an FPT reduction  $\text{GAPMLD}$  to  $\text{GAPSNC}$ :

**Lemma 5.2** *For any constant  $\gamma > 1$ , there is an FPT reduction from  $\text{GAPMLD}_\gamma$  to  $\text{GAPSNC}_\gamma$ .*

*Proof.* Given an instance  $(\mathbf{A}, \mathbf{y}, k)$  of  $\text{GAPMLD}_\gamma$  where  $\mathbf{A} \in \mathbb{F}_2^{n \times m}$ . We create an instance  $(\mathbf{A}', \mathbf{y}', k')$  of  $\text{GAPSNC}_\gamma$  by letting

$$\mathbf{A}' = \begin{bmatrix} \mathbf{1}_{\lceil \gamma k + 1 \rceil} \otimes \mathbf{A} \\ \text{Id}_m \end{bmatrix} \in \mathbb{F}_2^{\lceil \gamma k + 1 \rceil (n+m) \times m}, \mathbf{y}' = \begin{bmatrix} \mathbf{1}_{\lceil \gamma k + 1 \rceil} \otimes \mathbf{y} \\ \mathbf{0}_m \end{bmatrix} \in \mathbb{F}_2^{\lceil \gamma k + 1 \rceil (n+m)}$$



and  $k' = k$ .

**Completeness** Suppose that  $(\mathbf{A}, \mathbf{y}, k)$  is a YES instance of  $\text{GAPMLD}_\gamma$ , i.e., there exists  $\mathbf{x} \in \mathcal{B}_m(\mathbf{0}, k)$  such that  $\mathbf{A}\mathbf{x} = \mathbf{y}$ . Consider  $\mathbf{A}'\mathbf{x} - \mathbf{y}'$ . We have

$$\mathbf{A}'\mathbf{x} - \mathbf{y}' = \begin{bmatrix} \mathbf{1}_{\lceil \gamma k + 1 \rceil} \otimes (\mathbf{A}\mathbf{x} - \mathbf{y}) \\ \mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{\lceil \gamma k + 1 \rceil n} \\ \mathbf{x} \end{bmatrix}.$$

Hence,  $\mathbf{x}$  is a vector in the ball  $\mathcal{B}_m(\mathbf{0}, k)$  such that  $\|\mathbf{A}'\mathbf{x} - \mathbf{y}'\|_0 = \|\mathbf{x}\|_0 \leq k$ , meaning that  $(\mathbf{A}', \mathbf{y}', k')$  is a YES instance of  $\text{GAPSNC}_\gamma$ .

**Soundness** Suppose that  $(\mathbf{A}, \mathbf{y}, k)$  is a NO instance of  $\text{GAPMLD}_\gamma$ . Consider any  $\mathbf{x} \in \mathcal{B}_m(\mathbf{0}, \gamma k)$ . To show that  $\|\mathbf{A}'\mathbf{x} - \mathbf{y}'\|_0 > \gamma k$ , let us consider two cases.

- Case 1:  $\|\mathbf{x}\|_0 > \gamma k$ . In this case, we have  $\|\mathbf{A}'\mathbf{x} - \mathbf{y}'\|_0 \geq \|\mathbf{x}\|_0 > \gamma k$ .
- Case 2:  $\|\mathbf{x}\|_0 \leq \gamma k$ . In this case, we have  $\|\mathbf{A}'\mathbf{x} - \mathbf{y}'\|_0 \geq \lceil \gamma k + 1 \rceil \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_0 \geq \lceil \gamma k + 1 \rceil > \gamma k$ , where the third inequality comes from the fact that  $(\mathbf{A}, \mathbf{y}, k)$  is a NO instance of  $\text{GAPMLD}_\gamma$  and hence  $\mathbf{A}\mathbf{x} \neq \mathbf{y}$ .

As a result, we can conclude that  $(\mathbf{A}', \mathbf{y}', k')$  is a NO instance of  $\text{GAPSNC}_\gamma$ .  $\square$

## 5.2 Sparse Covering Codes

Below, we provide the definition of Sparse Covering Codes<sup>19</sup> over  $\mathbb{F}_2$ .

**Definition 5.3** A Sparse Covering Code (SCC) over  $\mathbb{F}_2$  with parameters<sup>20</sup>  $(m, q, t, d, r, \delta)$  is a tuple  $(\mathbf{L}, \mathbf{T}, \mathcal{D})$  where  $\mathbf{L} \in \mathbb{F}_2^{h \times m}$  is the basis of an  $m$  dimensional linear code with minimum distance (at least)  $d$ ,  $\mathbf{T} \in \mathbb{F}_2^{q \times h}$  is a linear transformation, and  $\mathcal{D}$  is a poly( $h$ ) time samplable distribution over  $\mathbb{F}_2^h$  such that for any  $\mathbf{x} \in \mathcal{B}_q(\mathbf{0}, t)$ , the following holds:

$$\Pr_{\mathbf{s} \sim \mathcal{D}} \left[ \mathbf{x} \in \mathbf{T} \left( \mathcal{B}_h(\mathbf{s}, r) \cap \mathbf{L}\mathbb{F}_2^m \right) \right] \geq \delta. \quad (7)$$

We would like to note that SCC are closely related to Locally Dense Codes (see [DMS03] or [Mic14] for the definition). A key difference in our definition is that the code is tailored towards covering sparse vectors which allows us greater flexibility in the parameters. This is in direct contrast to previous constructions which were designed to cover entire subspaces, with parameters which end up having strong dependencies on the ambient dimension. Consequently, they are not directly applicable to the parameterized setting.

Next, we show the existence of SCC for a certain range of parameters. Informally, the existence of SCC follows from the existence of codes near the sphere-packing bound, such as BCH codes.

**Lemma 5.4** For any  $q, t \in \mathbb{N}$  and any  $\varepsilon > 0$ , there exist  $d, r, h, m \in \mathbb{N}$ , two linear maps  $\mathbf{L} \in \mathbb{F}_2^{h \times m}$  and  $\mathbf{T} \in \mathbb{F}_2^{q \times h}$ , and a poly( $h$ ) time samplable distribution  $\mathcal{D}$  over  $\mathbb{F}_2^h$  such that  $(\mathbf{L}, \mathbf{T}, \mathcal{D})$  is a SCC with parameters  $(m, q, t, d, r, \frac{1}{d^{1/2}})$ . Additionally, the following holds:

<sup>19</sup>The definition can be naturally extended to fields of larger size.

<sup>20</sup>We remark that the parameter  $h$  is implicit in specifying SCC.

- $r \leq (1/2 + \varepsilon)d$ ,
- $r, d \leq O(t/\varepsilon)$  and  $h, m \leq \text{poly}(q, t, 1/\varepsilon)$ ,
- $\mathbf{L}$  and  $\mathbf{T}$  can be computed in  $\text{poly}(q, t, 1/\varepsilon)$  time for all  $\varepsilon > 0$ .

*Proof.* Let  $d = 2\lceil t/\varepsilon \rceil + 1$  and let  $r = \left(\frac{d-1}{2}\right) + t$ . Let  $h$  be the smallest positive integer such that  $h+1$  is a power of two and that  $h \geq \max\{2q, d\}$ . Finally, we set  $m = h - \left(\frac{d-1}{2}\right) \log(h+1)$ . It is clear that the chosen parameters satisfy the first two conditions, i.e.,  $r \leq (1/2 + \varepsilon)d$  and  $r, d \leq O(t/\varepsilon)$ , and  $h, m \leq \text{poly}(q, t, 1/\varepsilon)$ .

Let  $\mathbf{L}$  be the generator matrix of the  $[h, m, d]_2$  linear code as given by Theorem 3.11. Without loss of generality, we assume that the code is systematic on the first  $m$  coordinates. The linear map  $\mathbf{T}$  is defined as  $\mathbf{T} \stackrel{\text{def}}{=} \begin{bmatrix} \text{Id}_q & \mathbf{0}_{q \times h-q} \end{bmatrix}$  i.e., it is the matrix which projects onto the first  $q$  coordinates. The distribution  $\mathcal{D}$  is given as follows: we set the first  $q$  coordinates of the vector to 0 and each of the remaining  $(h-q)$  coordinates is sampled uniformly and independently at random from  $\mathbb{F}_2$ . It is clear that  $\mathbf{T}$  is computed in  $\text{poly}(q, t, 1/\varepsilon)$  time and  $\mathcal{D}$  is a  $\text{poly}(h)$  time samplable distribution over  $\mathbb{F}_2^h$ . From Theorem 3.11, we also have that  $\mathbf{L}$  can be computed in  $\text{poly}(h) = \text{poly}(q, t, 1/\varepsilon)$  time.

It remains to show that for our choices of matrices  $\mathbf{L}$ ,  $\mathbf{T}$  and distribution  $\mathcal{D}$ , equation 7 holds for any fixed choice of  $\mathbf{x} \in \mathcal{B}_q(\mathbf{0}, t)$ . Fix a vector  $\mathbf{x} \in \mathcal{B}_q(\mathbf{0}, t)$  and define the set  $\mathcal{C} = \left\{ \mathbf{z} \in \mathbb{F}^{h-q} \mid \mathbf{x} \circ \mathbf{z} \in \mathbf{L}\mathbb{F}^m \right\}$ . Since the code generated by  $\mathbf{L}$  is systematic on the first  $m$  coordinates, we have that  $|\mathcal{C}| = 2^{m-q}$ . Moreover, since the code generated by  $\mathbf{L}$  has distance  $d$ , we have that every distinct pair of vectors  $\mathbf{z}_1, \mathbf{z}_2 \in \mathcal{C}$  are at least  $d$ -far from each other (i.e.  $\|\mathbf{z}_1 - \mathbf{z}_2\|_0 \geq d$ ).

Recall that  $r - t = \frac{d-1}{2}$ . Therefore, for any distinct pair of vectors  $\mathbf{z}_1, \mathbf{z}_2 \in \mathcal{C}$ , the sets  $\mathcal{B}_{h-q}(\mathbf{z}_1, r-t)$  and  $\mathcal{B}_{h-q}(\mathbf{z}_2, r-t)$  are disjoint. Hence the number of vectors in the union of  $(r-t)$ -radius Hamming balls around every  $\mathbf{z} \in \mathcal{C}$  is

$$2^{m-q} \left| \mathcal{B}_{h-q} \left( \mathbf{0}, \frac{d-1}{2} \right) \right| \geq 2^{m-q} \binom{h-q}{\frac{d-1}{2}} \geq 2^{m-q} \binom{h/2}{\frac{d-1}{2}} \geq 2^{m-q} \left( \frac{h}{d-1} \right)^{\frac{d-1}{2}}$$

On the other hand,  $|\mathbb{F}_2^{h-q}| = 2^{h-q} = 2^{m-q}(h+1)^{\frac{d-1}{2}}$ . Hence, with probability at least  $\left( \frac{h}{(d-1)(h+1)} \right)^{\frac{d-1}{2}} \geq \frac{1}{d^{d/2}}$ , a vector  $\mathbf{s}'$  sampled uniformly from  $\mathbb{F}^{h-q}$  lies in  $\mathcal{B}_{h-q}(\mathbf{p}', r-t)$  for some vector  $\mathbf{p}' \in \mathcal{C}$ . Fix such a vector  $\mathbf{s}'$  and consider the vectors  $\mathbf{s} = \mathbf{0}_q \circ \mathbf{s}'$  and  $\mathbf{p} = \mathbf{x} \circ \mathbf{p}'$ . Then by construction,

$$\|\mathbf{s} - \mathbf{p}\|_0 = \|\mathbf{s}' - \mathbf{p}'\|_0 + \|\mathbf{x}\|_0 \leq (r-t) + t = r,$$

and  $\mathbf{T}\mathbf{p} = \mathbf{x}$ . To complete the proof, we observe that the distribution of the vector  $\mathbf{s}$  (as constructed from  $\mathbf{s}'$ ) is identical to  $\mathcal{D}$ , and hence the claim follows.  $\square$

### 5.3 Reducing GAPSNC to GAPMDP

In this subsection, we state and prove the FPT reduction from the GAPSNC problem to the GAPMDP problem. It uses the basic template of the reduction from [DMS03], which is modified to work in combination with SCC.

**Lemma 5.5** For any constants  $\gamma' > 2$  and  $\gamma \geq 1$  such that  $\gamma < \frac{2\gamma'}{2+\gamma'}$ , there is a randomized FPT reduction from  $\text{GAPSNC}_{\gamma'}$  to  $\text{GAPMDP}_{\gamma}$ .

*Proof.* Let  $(\mathbf{B}, \mathbf{y}, t)$  be the input for  $\text{GAPSNC}_{\gamma}$  where  $\mathbf{B} \in \mathbb{F}_2^{n \times q}$ ,  $\mathbf{y} \in \mathbb{F}_2^n$ , and  $t$  is the parameter. Moreover, let  $\varepsilon > 0$  be a sufficiently small constant such that  $\gamma < \frac{2\gamma'}{2+(1+2\varepsilon)\gamma'}$ . Let  $d, r, h, m \in \mathbb{N}$ ,  $\mathbf{L} \in \mathbb{F}_2^{h \times m}$ ,  $\mathbf{T} \in \mathbb{F}_2^{q \times h}$  and  $\mathcal{D}$  be as in Lemma 5.4. Pick  $a', b' \in \mathbb{N}$  such that

$$\frac{\gamma}{\gamma' - \gamma} < \frac{a'}{b'} < \frac{(d/r) - \gamma}{\gamma}. \quad (8)$$

To see that such  $a'$  and  $b'$  exists, observe the following:

$$\zeta = \frac{(d/r) - \gamma}{\gamma} - \frac{\gamma}{\gamma' - \gamma} \geq \frac{2\gamma' - \gamma(2 + (1 + 2\varepsilon)\gamma' + 2\varepsilon\gamma)}{(1 + 2\varepsilon)\gamma(\gamma' - \gamma)} > \frac{2\varepsilon\gamma}{(1 + 2\varepsilon) \cdot (\gamma' - \gamma)} > 0.$$

Moreover, we can always choose  $a'$  and  $b'$  so that they are at most  $2/\zeta = O(1)$ .

Let  $a = a'r$  and  $b = b't$ . Note that, condition (8) implies that  $\gamma(at + br) < \min\{\gamma'at, bd\}$ . We produce an instance  $(\mathbf{A}, k)$  for  $\text{GAPMDP}_{\gamma}$  by first sampling  $\mathbf{s} \sim \mathcal{D}$ . Then, we set  $k = at + br$  and

$$\mathbf{A} = \begin{bmatrix} \mathbf{1}_a \otimes \mathbf{BTL} & -\mathbf{1}_a \otimes \mathbf{y} \\ \mathbf{1}_b \otimes \mathbf{L} & -\mathbf{1}_b \otimes \mathbf{s} \end{bmatrix} \in \mathbb{F}_2^{(an+bh) \times (m+1)}.$$

**Parameter Dependencies.** Notice that  $k = at + br = a'rt + b'tt = O(rt)$ . Moreover, from Lemma 5.4, we have that  $r \leq O(t)$ . Hence, we can conclude that  $k = O(t^2)$  (note that  $t$  is the parameter of the input instance  $(\mathbf{B}, \mathbf{y}, t)$  of  $\text{GAPSNC}_{\gamma'}$ ).

**Completeness.** Suppose that  $(\mathbf{B}, \mathbf{y}, t)$  is a YES instance of  $\text{GAPSNC}_{\gamma'}$ . That is, there exists  $\mathbf{x} \in \mathcal{B}_q(\mathbf{0}, t)$  such that  $\|\mathbf{Bx} - \mathbf{y}\|_0 \leq t$ . Now, from Lemma 5.4, with probability at least  $1/d^{d/2}$ , we have  $\mathbf{x} \in \mathbf{T}(\mathcal{B}_m(\mathbf{s}, r) \cap \text{LF}_2^m)$ . This is equivalent to the following: there exists  $\mathbf{z}' \in \mathbb{F}_2^m$  such that  $\mathbf{TLz}' = \mathbf{x}$  and  $\|\mathbf{Lz}' - \mathbf{s}\|_0 \leq r$ . Conditioned on this event, we can pick  $\mathbf{z} = \mathbf{z}' \circ \mathbf{1} \in \mathbb{F}_2^{m+1}$ , which yields

$$\|\mathbf{Az}\|_0 = a\|\mathbf{BTLz}' - \mathbf{y}\|_0 + b\|\mathbf{Lz}' - \mathbf{s}\|_0 = a\|\mathbf{Bx} - \mathbf{y}\|_0 + b\|\mathbf{Lz}' - \mathbf{s}\|_0 \leq at + br.$$

In other words, with probability at least  $1/d^{d/2}$ ,  $(\mathcal{A}, k)$  is a YES instance of  $\text{GAPMDP}_{\gamma}$  as desired.

**Soundness.** Suppose that  $(\mathbf{B}, \mathbf{y}, t)$  is a NO instance of  $\text{GAPSNC}_{\gamma'}$ . We will show that, for all non-zero  $\mathbf{z} \in \mathbb{F}_2^{m+1}$ ,  $\|\mathbf{Az}\|_0 > \gamma(at + br)$ ; this implies that  $(\mathbf{A}, k)$  is a NO instance of  $\text{GAPMDP}_{\gamma}$ .

To show this let us consider two cases, based on the last coordinate  $\mathbf{z}_{m+1}$  of  $\mathbf{z}$ . Let  $\mathbf{z}' = [\mathbf{z}_1 \cdots \mathbf{z}_m] \in \mathbb{F}_2^m$  be the vector consisting of the first  $m$  coordinates of  $\mathbf{z}$ .

If  $\mathbf{z}_{m+1} = 0$ , then  $\|\mathbf{Az}\|_0 = a\|\mathbf{BTLz}'\|_0 + b\|\mathbf{Lz}'\|_0 \geq b\|\mathbf{Lz}'\|_0 \geq bd$ , where the last inequality comes from the fact that  $\mathbf{L}$  is a generator matrix of a code of distance  $d$  (and that  $\mathbf{z} \neq \mathbf{0}$ ). Finally, recall that our parameter selection implies that  $bd > \gamma(at + br)$ , which yields the desired result for this case.

On the other hand, if  $\mathbf{z}_{m+1} = 1$ , then  $\|\mathbf{Az}\|_0 = a\|\mathbf{BTLz}' - \mathbf{y}\|_0 + b\|\mathbf{Lz}' - \mathbf{s}\|_0 \geq a\|\mathbf{BTLz}' - \mathbf{y}\|_0 \geq \gamma'at$ , where the last inequality simply follows from our assumption that  $(\mathbf{B}, \mathbf{y}, t)$  is a NO instance of  $\text{GAPSNC}_{\gamma'}$ . Finally, recall that our parameter selection guarantees that  $\gamma'at > \gamma(at + br)$ . This concludes the proof.  $\square$

**Gap Amplification.** Finally, the above gap hardness can be boosted to any constant using the now standard technique of tensoring the code (c.f. [DMS03],[AK14]) using the following lemma:

**Proposition 5.6 (E.g. [DMS03])** *Given two linear codes  $C_1 \subseteq \mathbb{F}_2^m$  and  $C_2 \subseteq \mathbb{F}_2^n$ , let  $C_1 \otimes C_2 \subseteq \mathbb{F}_2^{m \times n}$  be the tensor product of  $C_1$  and  $C_2$ . Then  $d(C_1 \otimes C_2) = d(C_1)d(C_2)$ .*

We briefly show how the above proposition can be used to amplify the gap. Consider a  $\text{GAPMDP}_\gamma$  instance  $(\mathbf{A}, k)$  where  $\mathbf{A} \in \mathbb{F}_2^{m \times n}$ . Let  $C \subseteq \mathbb{F}_2^m$  be the linear code generated by it. Let  $C^{\otimes 2} = C \otimes C$  be the tensor product of the code with itself, and let  $\mathbf{A}^{\otimes 2}$  be its generator matrix. By the above proposition, if  $(\mathbf{A}, k)$  is a YES instance, then  $d(C^{\otimes 2}) \leq k^2$ . Conversely, if  $(\mathbf{A}, k)$  is a NO instance, then  $d(C^{\otimes 2}) \geq \gamma^2 k^2$ . Therefore  $(\mathbf{A}^{\otimes 2}, k^2)$  is a  $\text{GAPMDP}_{\gamma^2}$  instance. Hence, for any  $\alpha \in \mathbb{R}_+$ , repeating this argument  $\lceil \log_\gamma \alpha \rceil$ -number of times gives us an FPT reduction from  $k$ - $\text{GAPMDP}_\gamma$  to  $k^{2^{\lceil \log_\gamma \alpha \rceil}}$ - $\text{GAPMDP}_\alpha$ . We have thereby completed our proof of Theorem 5.1.

## 6 Parameterized Intractability of Shortest Vector Problem

We begin by stating the intractability of  $\text{GAPSVP}$ , which is the main result of this section:

**Theorem 6.1 (FPT Inapproximability of  $\text{GAPSVP}$ )** *Assuming PIH, any  $p > 1$ , there exists constant  $\gamma_p > 1$  (where  $\gamma_p$  depends on  $p$ ), such that there is no randomized FPT algorithm for  $\text{GAPSVP}_{p, \gamma_p}$ .*

For  $p = 2$ , the gap can be boosted to any constant factor, as stated by the following:

**Theorem 6.2 (Constant Inapproximability of  $\text{GAPSVP}_{2, \gamma}$ )** *Assuming PIH, for all constants  $\gamma \geq 1$ , there is no randomized FPT algorithm for  $\text{GAPSVP}_{2, \gamma}$ .*

The proof of Theorem 6.1 is essentially the same as that of Khot [Kho05], with only a small change in parameter selection. We devote Subsection 6.1 to this proof. Theorem 6.2 follows by gap amplification via tensor product of lattices. Unlike tensor products of codes, the length of the shortest vector of the tensor product of two lattices is not necessarily equal to the product of the length of the shortest vector of each lattice. Fortunately, Haviv and Regev [HR07] showed that a strengthened notion of soundness from [Kho05], which will be described soon, allows gap amplification via tensoring in the  $\ell_2$  norm. This is explained in more detail in Section 6.2.

### 6.1 Following Khot's Reduction: Proof of Theorem 6.1

The proof of Theorem 6.1 goes through the following FPT inapproximability of  $\text{GAPNVP}_{p, \eta}$ .

**Theorem 6.3 (FPT Inapproximability of  $\text{GAPNVP}$ )** *Assuming PIH, for any  $\gamma, p \geq 1$ , there is no FPT algorithm which on input  $(\mathbf{B}, \mathbf{y}, t)$  where  $\mathbf{B} \in \mathbb{Z}^{n \times q}$ ,  $\mathbf{y} \in \mathbb{Z}^n$  and  $t \in \mathbb{N}$ , can distinguish between*

- (YES) there exists  $\mathbf{x} \in \mathbb{Z}^q$  such that  $\mathbf{B}\mathbf{x} - \mathbf{y} \in \{0, 1\}^n$  and  $\|\mathbf{B}\mathbf{x} - \mathbf{y}\|_p^p \leq t$ .
- (NO) for all choices of  $\mathbf{x} \in \mathbb{Z}^q$  and  $w \in \mathbb{Z} \setminus \{0\}$ , we have  $\|\mathbf{B}\mathbf{x} - w \cdot \mathbf{y}\|_p^p > \gamma \cdot t$

Since the proof of the above theorem is identical to the FPT reduction of  $\text{GAP2CSP}$  to  $\text{GAPSNC}$ , we defer it to Appendix B. Note that the reduction has stronger guarantees than a usual reduction from  $\text{GAP2CSP}_\varepsilon$  to  $\text{GAPNVP}_{p, \eta}$  in both the YES and NO cases. In the YES case, the witness of the  $\text{GAPNVP}$  instance is a  $\{0, 1\}$ -vector. Furthermore, in the NO case of the above lemma, the lower bound guarantee is in terms of the Hamming norm. These properties are used crucially in the analysis for the reduction from  $\text{GAPNVP}_{p, \eta}$  to  $\text{GAPSVP}_{p, \gamma}$ . The proof of Theorem 6.1 also requires the following definition of annoying vectors from [Kho05].

**Definition 6.4 (Annoying Vectors)** For a parameter  $l \in \mathbb{N}$ , we say that a vector  $\mathbf{z} \in \mathbb{Z}^u$  for some  $u \in \mathbb{N}$  is an Annoying Vector in the  $\ell_p$  norm, if it violates all of the following conditions:

1. The Hamming norm of  $\mathbf{z}$  is at least  $l$ .
2. All coordinates of  $\mathbf{z}$  are even, and its Hamming norm is at least  $l/2^p$ .
3. All coordinates of  $\mathbf{z}$  are even, and it has at least one coordinate of magnitude at least  $l^{10l}$ .

As shown in [Kho05, HR07], lattices without annoying vectors (for an appropriate value of  $l$ ) are “well behaved” with respect to tensor products of lattices. This notion will be useful later, when the lattice tensor product is used to amplify the gap to any constant for  $p = 2$  (Section 6.2).

The second step is a randomized FPT reduction from  $\text{GAPNVP}_{p,\eta}$  to  $\text{GAPSV}_{p,\gamma}$ , as stated by the following lemma. For succinctness, we define a couple of additional notations: let  $\mathcal{L}(\mathbf{A})$  denote the lattice generated by the matrix  $\mathbf{A} \in \mathbb{Z}^{n \times m}$ , i.e.,  $\mathcal{L}(\mathbf{A}) = \{\mathbf{A}\mathbf{x} \mid \mathbf{x} \in \mathbb{Z}^m\}$ , and let  $\lambda_p(\mathcal{L})$  denote the length (in the  $\ell_p$  norm) of the shortest vector of the lattice  $\mathcal{L}$ , i.e.,  $\lambda_p(\mathcal{L}) = \min_{\mathbf{0} \neq \mathbf{y} \in \mathcal{L}} \|\mathbf{y}\|_p$ .

**Lemma 6.5** Fix  $p > 1$ , and let  $\eta \geq 1$  be such that  $\frac{1}{2} + \frac{1}{2^p} + \frac{(2^p+1)}{\eta} < 1$ . Let  $(\mathbf{B}, \mathbf{y}, t)$  be a  $\text{GAPNVP}_{p,\eta}$  instance, as given by Theorem 6.3. Then, there is a randomized FPT reduction from  $\text{GAPNVP}_{p,\eta}$  instance  $(\mathbf{B}, \mathbf{y}, t)$  to  $\text{GAPSV}_{p,\gamma}$  instance  $(\mathbf{B}_{\text{svp}}, \gamma_p^{-1}l)$  with  $l = \eta \cdot t$  such that

- $\mathbf{B}_{\text{svp}} \in \mathbb{Z}^{(n+h+g+1) \times (q+h+g+2)}$  where  $m = l^{O(l)} \cdot \text{poly}(n, q)$  and  $h = \text{poly}(l, n, q)$ .
- If  $(\mathbf{B}, \mathbf{y}, t)$  is a YES instance (from Theorem 6.3), then with probability 0.8,  $\lambda_p(\mathcal{L}(\mathbf{B}_{\text{svp}}))^p \leq \gamma_p^{-1}l$ .
- If  $(\mathbf{B}, \mathbf{y}, t)$  is a NO instance, then with probability 0.9, for all choices of  $\mathbf{x} \in \mathbb{Z}^{q+h+g+2} \setminus \{\mathbf{0}\}$ , the vector  $\mathbf{B}_{\text{svp}}\mathbf{x}$  is not annoying. In particular, we have  $\lambda_p(\mathcal{L}(\mathbf{B}_{\text{svp}}))^p \geq l$ .

Here  $\gamma_p := \frac{1}{\frac{1}{2} + (2^p+1)/\eta + 1/2^p}$  is strictly greater than 1 by our choice of  $\eta$ .

Combining the above lemma with Theorem 6.3 gives us Theorem 6.1.

The reduction in the above lemma is nearly identical to the one from [Kho05]. We devote the rest of this subsection to describing this reduction and proving Lemma 6.5. The rest of this section is organized as follows. In Section 6.1.1, we define the BCH lattice, which is the key gadget used in the reduction. Using the BCH lattice and the  $\text{GAPNVP}_{p,\eta}$  instance, we construct the intermediate lattice in Section 6.1.2. The intermediate lattice serves to blow up the number of “good vectors” for the YES case, while controlling the number of annoying vectors for the NO case. In particular, this step ensures that the number of good vectors (Lemma 6.7) far outnumber the number of annoying vectors (Lemma 6.8). Finally, in Section 6.1.3 we compose the intermediate lattice with a random homogeneous constraint (sampled from an appropriate distribution), to give the final  $\text{GAPSV}_{p,\gamma}$  instance. The additional random constraint is used to annihilate all annoying vectors in the NO case, while retaining at least one good vector in the YES case.

**Setting:** For the rest of the section, we fix  $(\mathbf{B}, \mathbf{y}, t)$  to be a  $\text{GAPNVP}_{p,\eta}$  instance (as given by Theorem 6.3) and we set  $l := \eta \cdot t$ .

### 6.1.1 The BCH Lattice gadget

We begin by defining the BCH lattices which is the key gadget used in the reduction. Given parameters  $l, h \in \mathbb{N}$  where  $h$  is a power of 2, the BCH lattice is defined by the generator  $\mathbf{B}_{\text{BCH}}$

defined as

$$\mathbf{B}_{\text{BCH}} = \begin{bmatrix} \text{Id}_h & \mathbf{0}_{h \times g} \\ Q \cdot \mathbf{P}_{\text{BCH}} & 2Q \cdot \text{Id}_g \end{bmatrix} \in \mathbb{Z}^{(h+g) \times (h+g)}.$$

where  $\mathbf{P}_{\text{BCH}} \in \{0, 1\}^{g \times h}$  is the parity check matrix of a  $[h, h - g, l]_2$ -BCH code (i.e.,  $g$  is the codimension of the code whose parity check matrix is  $\mathbf{P}_{\text{BCH}}$ ), and  $Q = l^{10l}$ . By our choice of parameters it follows that  $g \leq \frac{1}{2} \log h$  (see Theorem 3.11). The following lemma summarizes the key properties of BCH lattices, as defined above.

**Lemma 6.6** *Let  $\mathbf{B}_{\text{BCH}} \in \mathbb{Z}^{(h+g) \times (h+g)}$  be the BCH lattice constructed as above. Let  $r = \left(\frac{1}{2} + \frac{1}{2^p} + \frac{1}{\eta}\right)\eta t$ . Then there exists a randomized polynomial time algorithm, which with probability at least 0.99, returns vector  $\mathbf{s} \in \mathbb{Z}^{h+g}$  such that the following holds. There exists at least  $\frac{1}{100} 2^{-g} \binom{h}{r}$  distinct coefficient vectors  $\mathbf{z} \in \mathbb{Z}^{h+g}$  such that the vector  $(\mathbf{B}_{\text{BCH}}\mathbf{z} - \mathbf{s})$  is a  $\{0, 1\}$ -vector of Hamming weight exactly  $r$ .*

We skip the proof of the above lemma, since it is identical to proof of Lemma 4.3 in [Kho05]. Note that this is in fact even weaker than Khot's lemma, since we do not impose a bound on  $\|\mathbf{z}\|_p$ .

### 6.1.2 The Intermediate Lattice

We now define the intermediate lattice. Let  $(\mathbf{B}, \mathbf{y}, t)$  be an instance of  $\text{GAPNVP}_{p, \eta}$ , where  $\mathbf{B} \in \mathbb{Z}^{n \times q}$ . The intermediate lattice  $\mathbf{B}_{\text{int}}$  is constructed as follows. Let  $l = \eta t$ . Let  $h$  be the smallest power of 2 such that  $h \geq \max\{2^{10p+10}n, (100l)^{100\eta l}\}$ , and let  $\mathbf{B}_{\text{BCH}}$  be constructed as above. Then

$$\mathbf{B}_{\text{int}} = \begin{bmatrix} 2\mathbf{B} & \mathbf{0}_{n \times (h+g)} & 2\mathbf{y} \\ \mathbf{0}_{(h+g) \times q} & \mathbf{B}_{\text{BCH}} & \mathbf{s} \end{bmatrix} \in \mathbb{Z}^{(n+h+g) \times (q+h+g+1)}.$$

where  $\mathbf{s} \in \mathbb{Z}^{h+g}$  is the vector given by Lemma 6.6. The following lemmas bound the good and annoying vectors in the YES and NO cases respectively.

**Lemma 6.7** *Let  $(\mathbf{B}, \mathbf{y}, t)$  be a YES instance, and let  $\mathbf{B}_{\text{int}}$  be the corresponding intermediate lattice. Then, with probability at least 0.99, there exists at least  $h^r \left(100h^{l/2}l^l\right)^{-1}$  distinct vectors  $\mathbf{x} \in \mathbb{Z}^{q+h+g+1}$  such that  $\mathbf{B}_{\text{int}}\mathbf{x}$  are distinct  $\{0, 1, 2\}$ -vectors with  $\ell_p$  norm at most  $(2^p t + r)^{1/p}$ .*

*Proof.* Let  $\tilde{\mathbf{x}} \in \mathbb{Z}^q$  be vector such that  $\mathbf{B}\tilde{\mathbf{x}} - \mathbf{y}$  is a  $\{0, 1\}$ -vector of Hamming weight at most  $t$ . From Lemma 6.6, there exists at least  $2^{-g} \binom{h}{r} / 100$  distinct vectors  $\mathbf{z} \in \mathbb{Z}^{h+g}$  such that  $\mathbf{B}_{\text{BCH}}\mathbf{z} - \mathbf{s}$  is a  $\{0, 1\}$ -vector of Hamming weight exactly  $r$ . For each such distinct coefficient vector, consider the vector  $\mathbf{x} = \tilde{\mathbf{x}} \circ \mathbf{z} \circ -1$ . By construction, it follows that  $\mathbf{B}_{\text{int}}\mathbf{x} = (2\mathbf{B}\tilde{\mathbf{x}} - 2\mathbf{y}) \circ (\mathbf{B}_{\text{BCH}}\mathbf{z} - \mathbf{s})$  is a  $\{0, 1, 2\}$ -vector and  $\|\mathbf{B}_{\text{int}}\mathbf{x}\|_p^p = 2^p \|\mathbf{B}\tilde{\mathbf{x}} - \mathbf{y}\|_p^p + \|\mathbf{B}_{\text{BCH}}\mathbf{z} - \mathbf{s}\|_p^p \leq 2^p t + r$ . Since the number of such vectors is at least the number of distinct coefficient vectors  $\mathbf{z}$ , it can be lower bounded by

$$\frac{1}{100} \cdot 2^{-g} \binom{h}{r} \geq \frac{1}{100} \cdot 2^{-\frac{1}{2} \log h} \binom{h}{r} \geq \frac{1}{100} \cdot \frac{h^r}{r^r h^{l/2}} \geq \frac{1}{100} \cdot \frac{h^r}{l^l h^{l/2}}.$$

Finally, observe that each  $\mathbf{z}$  produces different  $\mathbf{B}_{\text{BCH}}\mathbf{z}$  and hence all  $\mathbf{B}_{\text{int}}\mathbf{v}$ 's are distinct.  $\square$

**Lemma 6.8** *Let  $(\mathbf{B}, \mathbf{y}, t)$  be a NO instance, and let  $\mathbf{B}_{\text{int}}$  be the corresponding intermediate lattice. Then the number of annoying vectors for  $\mathbf{B}_{\text{int}}$  is at most  $10^{-5}h^r \left(100h^{l/2}l\right)^{-1}$ .*

*Proof.* Consider  $\mathbf{x} \in \mathbb{Z}^{q+h+g+1}$  such that  $\mathbf{B}_{\text{int}}\mathbf{x}$  is an annoying vector. We partition the vector and write it as  $\mathbf{x} = \mathbf{x}_1 \circ \mathbf{x}_2 \circ x$  where  $\mathbf{x}_1 \in \mathbb{Z}^q$ ,  $\mathbf{x}_2 \in \mathbb{Z}^{m+h}$  and  $x \in \mathbb{Z}$ . Using this, we can express  $\mathbf{B}_{\text{int}}\mathbf{x}$  as  $\mathbf{B}_{\text{int}}\mathbf{x} = (2\mathbf{B}\mathbf{x}_1 - 2x \cdot \mathbf{y}) \circ (\mathbf{B}_{\text{BCH}}\mathbf{x}_2 - x \cdot \mathbf{s})$ .

Suppose  $x \neq 0$ . Since  $(\mathbf{B}, \mathbf{y}, t)$  is a NO instance, the vector  $[2\mathbf{B} \ \mathbf{0} \ 2\mathbf{y}] \cdot \mathbf{x} = 2\mathbf{B}\mathbf{x}_1 - 2x \cdot \mathbf{y}$  has Hamming weight at least  $l$ . This would imply that  $\mathbf{B}_{\text{int}}\mathbf{x}$  cannot be annoying, which is a contradiction.

Henceforth we can assume that  $x = 0$ ; note that we now have  $\mathbf{B}_{\text{int}}\mathbf{x} = (2\mathbf{B}\mathbf{x}_1) \circ (\mathbf{B}_{\text{BCH}}\mathbf{x}_2)$ . We further claim that, if  $\mathbf{B}_{\text{int}}\mathbf{x}$  is annoying, then all of its coordinates must be even. To see that this is the case, let us consider any  $\mathbf{x}$  such that  $\mathbf{B}_{\text{int}}\mathbf{x}$  has at least one odd coordinate. From  $\mathbf{B}_{\text{int}}\mathbf{x} = (2\mathbf{B}\mathbf{x}_1) \circ (\mathbf{B}_{\text{BCH}}\mathbf{x}_2)$ ,  $\mathbf{B}_{\text{BCH}}\mathbf{x}_2$  must have at least one odd coordinate. Let us further write  $\mathbf{x}_2$  as  $\mathbf{x}_2 = \mathbf{w}_1 \circ \mathbf{w}_2$  where  $\mathbf{w}_1 \in \mathbb{Z}^m$  and  $\mathbf{w}_2 \in \mathbb{Z}^h$ . Notice that  $\mathbf{B}_{\text{BCH}}\mathbf{x}_2 = \mathbf{w}_1 \circ (Q(\mathbf{P}_{\text{BCH}}\mathbf{w}_1 - 2\mathbf{w}_2))$ . Since every coordinate of  $\mathbf{B}_{\text{BCH}}\mathbf{x}_2$  must be less than  $Q$  in magnitude, it must be the case that  $\mathbf{P}_{\text{BCH}}\mathbf{w}_1 - 2\mathbf{w}_2 = \mathbf{0}$ . In other words,  $(\mathbf{w}_1 \bmod 2)$  is a codeword of the BCH code. However, since the code has distance  $l$ , this means that, if  $\mathbf{w}_1$  has at least one odd coordinate, it must have at least  $l$  odd (non-zero) coordinates. This contradicts to our assumption that  $\mathbf{B}_{\text{int}}\mathbf{x}$  is annoying.

Therefore,  $\mathbf{B}_{\text{int}}\mathbf{x}$  must have less than  $l/2^p$  non-zero coordinates, all of which are bounded in magnitude by  $l^{10l}$ . Hence, we can bound the total number of annoying vectors by

$$(2l^{10l} + 1)^{l/2^p} \binom{n+h+g}{\lfloor \frac{l}{2^p} \rfloor} \leq (2l)^{10l^2} (n+h+g)^{l/2^p} \leq (2l)^{10l^2} ((1+l/2^p)h)^{l/2^p} \leq e(2l)^{10l^2} h^{l/2^p}$$

where the second-to-last step holds since  $g \leq \frac{1}{2} \log h \leq \frac{lh}{2^{p+1}}$  and  $n \leq \frac{h}{2^{p+1}}$ . On the other hand,

$$\frac{h^r}{h^{l/2}l} = \frac{h^{\left(\frac{1}{2} + \frac{1}{\eta} + \frac{1}{2^p}\right)l}}{h^{l/2}l} = h^{l/2^p} (h/l^\eta)^{l/\eta} \geq h^{l/2^p} (100l)^{99l^2} \geq 10^5 \left( e(2l)^{10l^2} h^{l/2^p} \right).$$

which follows from our choice of  $h$ . Combining the bounds completes the proof.  $\square$

### 6.1.3 The GAPSVP $_{p,\gamma}$ -instance

Finally, we construct  $\mathbf{B}_{\text{svp}}$  from  $\mathbf{B}_{\text{int}}$  by adding a random homogeneous constraint as in [Kho05]. For ease of notation, let  $N_g$  denote the lower bound on the number of distinct coefficient vectors guaranteed by Lemma 6.7 in the YES case. Similarly, let  $N_a$  denote the upper bound on the number of annoying vectors as given in Lemma 6.8. Combining the two Lemmas we have  $N_g \geq 10^5 N_a$ , which will be used crucially in the construction and analysis of the final lattice.

**Construction of the Final Lattice** : Let  $\rho$  be any prime number in<sup>21</sup>  $[10^{-4}N_g, 10^{-2}N_g]$ . Furthermore, let  $\mathbf{r} \stackrel{\text{unif}}{\sim} [0, \rho - 1]^{n+h+g}$  be a uniformly sampled lattice point. We construct  $\mathbf{B}_{\text{svp}}$  as

$$\mathbf{B}_{\text{svp}} = \begin{bmatrix} \mathbf{B}_{\text{int}} & 0 \\ D \cdot \mathbf{r}\mathbf{B}_{\text{int}} & D \cdot \rho \end{bmatrix} \in \mathbb{Z}^{(n+h+g+1) \times (q+h+g+2)}.$$

where  $D = l^{10l}$ . In other words, this can be thought of as adding a random linear constraint to the intermediate lattice. The choice of parameters ensures that with good probability, in the YES case, at least one of the good vectors  $\mathbf{x} \in \mathbb{Z}^{q+h+g+1}$  evaluates to 0 modulo  $\rho$  on the random constraint, and therefore we can pick  $u \in \mathbb{Z}$  such that  $\mathbf{B}_{\text{svp}}(\mathbf{x} \circ u) = (\mathbf{B}_{\text{int}}\mathbf{x}) \circ 0$  still has small  $\ell_p$  norm. On the other hand, since  $N_a \ll N_g$ , with good probability, all of annoying vectors evaluate to non-zeros, and hence will contribute a coordinate of magnitude  $D = l^{10l}$ . This intuition is formalized below.

*Proof of Lemma 6.5.* Let  $\mathbf{B}_{\text{svp}}$  be the corresponding final lattice of  $(\mathbf{B}, \mathbf{y}, t)$  as described above. Observe that given the GAPNVP $_{p,\eta}$ -instance  $(\mathbf{B}, \mathbf{y}, t)$ , we can construct  $\mathbf{B}_{\text{svp}}$  in  $t^{O(t)} \cdot \text{poly}(n, q)$ -time.

Suppose that  $(\mathbf{B}, \mathbf{y}, t)$  is a NO instance and fix a vector  $\mathbf{x} \circ u \in \mathbb{Z}^{q+h+g+2}$ . If  $\mathbf{x}$  is not an annoying vector w.r.t. the intermediate lattice  $\mathbf{B}_{\text{int}}$ , then it follows that  $\mathbf{x} \circ u$  is also not an annoying vector w.r.t. the final lattice  $\mathbf{B}_{\text{svp}}$ . Therefore, it suffices to look at vectors  $\mathbf{x} \circ u$  such that  $\mathbf{x}$  is annoying w.r.t.  $\mathbf{B}_{\text{int}}$ . By construction,  $(\mathbf{B}_{\text{svp}}(\mathbf{x} \circ u))_{(n+h+g+1)} = 0$  implies that  $\mathbf{r}\mathbf{B}_{\text{int}}\mathbf{x} = u \cdot \rho$ , which in turn is equivalent to  $\mathbf{r}\mathbf{B}_{\text{int}}\mathbf{x} \equiv 0 \pmod{\rho}$ . Since,  $\mathbf{x}_i < \rho$  for every  $i \in [q+h+g+1]$ , this event happens with probability exactly  $1/\rho$  over the random choice of  $\mathbf{r}$  (for any arbitrary  $u \in \mathbb{Z}$ ). Therefore by union bound, for every choice of annoying vector  $\mathbf{x}$  and coefficient  $u$ , we have  $\mathbf{r}\mathbf{B}_{\text{int}}\mathbf{x} \neq u \cdot \rho$ , with probability at least  $1 - N_a/\rho \geq 0.9$ . And therefore, with probability at least 0.9, for every annoying  $\mathbf{x}$ , the last coordinate of  $\mathbf{B}_{\text{svp}}(\mathbf{x} \circ u)$  has magnitude at least  $D$ .

Next, suppose that  $(\mathbf{B}, \mathbf{y}, t)$  is a YES instance. We condition on the event that there exists at least  $N_g$  good vectors as guaranteed by Lemma 6.7. Consider any two distinct vectors  $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{Z}^{q+h+g+1}$  provided by Lemma 6.7. Since  $\mathbf{B}_{\text{int}}\mathbf{x}_1$  and  $\mathbf{B}_{\text{int}}\mathbf{x}_2$  are distinct  $\{0, 1, 2\}$ -vectors, they are pairwise independent modulo  $\rho > 2$ . Therefore, instantiating Lemma 5.8 from [Kho05] with the lower bound on the number of good vectors  $N_g$  and our choice of  $\rho$ , it follows that with probability at least 0.9, there exists a good vector  $\mathbf{x}$  such that  $\mathbf{r}\mathbf{B}_{\text{int}}\mathbf{x} \equiv 0 \pmod{\rho}$ . Therefore, by union bound, with probability at least 0.8 (over the randomness of Lemma 6.7 and the choice of  $\mathbf{r}, \rho$ ), there exists  $\mathbf{x}$  such that, for some  $u \in \mathbb{Z}$ ,  $\|\mathbf{B}_{\text{svp}}(\mathbf{x} \circ u)\|_p^p = \|\mathbf{B}_{\text{int}}\mathbf{x}\|_p^p \leq 2^p t + r = \gamma_p^{-1} l$ , which concludes the proof.  $\square$

## 6.2 Gap Amplification for GAPSV $P_{2,\gamma}$

As in the case of GAPMDP, we can use tensor product of lattices to boost the hardness of GAPSV $P_{2,\gamma}$  to any constant factor. The tensor product of lattices is defined similarly as in the case of linear codes. Given  $\mathbf{A} \in \mathbb{Z}^{m \times m'}$ , let  $\mathcal{L}(\mathbf{A})$  be the lattice generated by  $\mathbf{A}$ . Given matrices  $\mathbf{A} \in \mathbb{Z}^{m \times m'}$  and

<sup>21</sup>Note that the density of primes in this range is at least  $1/\log N_g = 1/r \log h$ . Therefore, a random sample of size  $O(r \log h)$  in this range contains a prime with high probability. Since we can test primality for any  $\rho \in [10^{-4}N_g, 10^{-2}N_g]$  in FPT time, this gives us an FPT algorithm to sample such a prime number efficiently.



$\mathbf{B} \in \mathbb{Z}^{n \times n'}$  the tensor product of  $\mathcal{L}(\mathbf{A})$  and  $\mathcal{L}(\mathbf{B})$  is defined as

$$\mathcal{L}(\mathbf{A}) \otimes \mathcal{L}(\mathbf{B}) = \{\mathbf{A}\mathbf{X}\mathbf{B}^\top \mid \mathbf{X} \in \mathbb{Z}^{m' \times n'}\} \quad (9)$$

For brevity, let  $\mathcal{L}_1 = \mathcal{L}(\mathbf{A})$  and  $\mathcal{L}_2 = \mathcal{L}(\mathbf{B})$ . The following lemma from [HR07] summarizes the properties of tensor product instances of  $\text{GAPSVP}_{2,\gamma}$ .

**Lemma 6.9** *Let  $\mathcal{L}_1$  and  $\mathcal{L}_2$  be as above. Then we have that  $\lambda_2(\mathcal{L}_1 \otimes \mathcal{L}_2) \leq \lambda_2(\mathcal{L}_1) \cdot \lambda_2(\mathcal{L}_2)$ . Furthermore, suppose every vector  $\mathbf{y} \in \mathcal{L}_1$  satisfies at least one of the following conditions:*

- *The vector  $\mathbf{y}$  has Hamming norm at least  $l$ .*
- *All coordinates of  $\mathbf{y}$  are even and  $\mathbf{y}$  has Hamming norm at least  $l/4$ .*
- *All coordinates of  $\mathbf{y}$  are even and  $\|\mathbf{y}\|_\infty \geq l^{4l}$ .*

*Then,  $\lambda_2(\mathcal{L}_1 \otimes \mathcal{L}_2) \geq l^{1/2} \cdot \lambda_2(\mathcal{L}_2)$*

It is easy to see that given a  $\text{GAPSVP}_{2,\gamma}$  instance  $(\mathbf{A}, k)$  (as in Lemma 6.5), the tensor product gives an FPT reduction to  $\text{GAPSVP}_{2,\gamma^2}$ . Indeed, if  $(\mathbf{A}, k)$ , then  $\lambda_2(\mathcal{L}(\mathbf{A}) \otimes \mathcal{L}(\mathbf{A}))^2 \leq k^2$ . Conversely, if  $(\mathbf{A}, k)$  is a NO instance, then setting  $l = \gamma k$  in Lemma 6.9, we get  $\lambda_2(\mathcal{L}(\mathbf{A}) \otimes \mathcal{L}(\mathbf{A}))^2 \geq l^2 = \gamma^2 k^2$ . Therefore, by repeated tensoring, we have proved Theorem 6.2.

We remark here that, while there is an analogous statement as in Lemma 6.9 for  $p \neq 2$ , the third case requires  $\|\mathbf{y}\|_\infty$  to not only be large in terms of  $l$ , but also in terms of the dimension of the lattice. However, to the best of our knowledge, this can only be achieved when  $l$  is polynomial in the dimension. This is indeed the reason we fail to amplify the gap to all constants for  $p \neq 2$ .

## 7 Conclusion and Open Questions

In this work, we have shown the parameterized inapproximability of  $k$ -Minimum Distance Problem ( $k$ -MDP) and  $k$ -Shortest Vector Problem ( $k$ -SVP) in the  $\ell_p$  norm for every  $p > 1$  based on the Parameterized Inapproximability Hypothesis (PIH), which in turns follows from the Gap Exponential Time Hypothesis (Gap-ETH). While our results give an evidence of intractability of these problems, there are still many questions that remain open. First and foremost, it is still open whether the hardness of both problems can be based on more standard assumptions, such as ETH or  $W[1] \neq \text{FPT}$ . On this front, we would like to note that the only reason we need PIH (or Gap-ETH) is to arrive at the inapproximability of the non-homogeneous variants of the problems (Theorems 4.3 and 6.3), which is needed for us even if we want to only rule out exact FPT algorithms for  $k$ -MDP and  $k$ -SVP. Hence, if one could prove the hardness of approximation for these problems under weaker assumptions, then the inapproximability of  $k$ -MDP and  $k$ -SVP would still follow.

Another obvious question from our work is whether  $k$ -SVP in the  $\ell_1$  norm is in FPT. Khot's reduction unfortunately does not work for  $\ell_1$ ; indeed, in the work of Haviv and Regev [HR07], they arrive at the hardness of approximating SVP in the  $\ell_1$  norm by embedding SVP instances in  $\ell_2$  to instances in  $\ell_1$  using an earlier result of Regev and Rosen [RR06]. The Regev-Rosen embedding inherently does not work in the FPT regime either, as it produces non-integral lattices. Similar issue applies to an earlier hardness result for SVP on  $\ell_1$  of [Mic00], whose reduction produces irrational bases.

An additional question regarding  $k$ -SVP is whether we can prove hardness of approximation for every constant factor for  $p \neq 2$ . As described earlier, the gap amplification techniques of [Kho05, HR07] require the distance  $k$  to be dependent on the input size  $nm$ , and hence are not applicable for

us. To the best of our knowledge, it is unknown whether this dependency is necessary. If they are indeed required, it would also be interesting to see whether other different techniques that work for our settings can be utilized for gap amplification instead of those from [Kho05, HR07].

Furthermore, the Minimum Distance Problem can be defined for linear codes in  $\mathbb{F}_p$  for any larger field of size  $p > 2$  as well. It turns out that our result does not rule out FPT algorithms for  $k$ -MDP over  $\mathbb{F}_p$  with  $p > 2$ . The issue here is that, in our proof of existence of Sparse Covering Codes (Lemma 5.4), we need the co-dimension of the code to be small compared to its distance. In particular, the co-dimension  $h - m$  has to be at most  $(d/2 + O(1)) \log_p h$  where  $d$  is the distance. While the BCH code over binary alphabet satisfies this property, we are not aware of any linear codes that satisfy this for larger fields. It is an intriguing open question to determine whether such codes exist, or whether the reduction can be made to work without existence of such codes.

Since the current reductions for both  $k$ -MDP and  $k$ -SVP are randomized, it is still an intriguing open question whether we can find deterministic reductions from PIH (or Gap-ETH) to these problems. As stated in the introduction, even in the non-parameterized setting, NP-hardness of SVP through deterministic reductions is not known. On the other hand, MDP is known to be NP-hard even to approximate under deterministic reductions; in fact, even the Dumer et al.'s reduction [DMS03] that we employ can be derandomized, as long as one has a deterministic construction for Locally Dense Codes [CW12, Mic14]. In our settings, if one can deterministically construct Sparse Covering Codes (i.e. derandomize Lemma 5.4), then we would also get a deterministic reduction for  $k$ -MDP.

Finally, another interesting research direction is to attempt to prove more concrete running time lower bounds for  $k$ -MDP and  $k$ -SVP. For instance, it is easy to see that  $k$ -MDP can be solved (exactly) in  $N^{O(k)}$  time, where  $N = nm$  is the input size. On the other hand, while not stated explicitly above, it is simple to check that our proof implies that  $k$ -MDP cannot be solved (or even approximated) in time  $N^{o(k)}$  for some small constant  $c > 0$ , assuming Gap-ETH. Would it be possible to improve this running time lower bound to the tight  $N^{o(k)}$ ? Similar questions also apply to  $k$ -SVP.

## Acknowledgment

We are grateful to Ishay Haviv for providing insights on how the gap amplification for  $p \neq 2$  from [HR07] works. Pasin would like to thank Danupon Nanongkai for introducing him to the  $k$ -Even Set problem and for subsequent useful discussions.

## 8 References

- [ABSS97] Sanjeev Arora, László Babai, Jacques Stern, and Z. Sweedyk. The hardness of approximate optima in lattices, codes, and systems of linear equations. *J. Comput. Syst. Sci.*, 54(2):317–331, 1997.
- [AD97] Miklós Ajtai and Cynthia Dwork. A public-key cryptosystem with worst-case/average-case equivalence. In *STOC*, pages 284–293, 1997.
- [Ajt96] Miklós Ajtai. Generating hard instances of lattice problems (extended abstract). In *STOC*, pages 99–108, 1996.
- [Ajt98] Miklós Ajtai. The shortest vector problem in  $\ell_2$  is NP-hard for randomized reductions (extended abstract). In *STOC*, pages 10–19, 1998.

- [AK14] Per Austrin and Subhash Khot. A simple deterministic reduction for the gap minimum distance of code problem. *IEEE Trans. Information Theory*, 60(10):6636–6645, 2014.
- [App17] Benny Applebaum. Exponentially-hard gap-CSP and local PRG via local hardcore functions. In *FOCS*, pages 836–847, 2017.
- [AS18] Divesh Aggarwal and Noah Stephens-Davidowitz. (gap/s)eth hardness of SVP. In *STOC*, 2018. To appear.
- [BGGS16] Arnab Bhattacharyya, Ameet Gadekar, Suprovat Ghoshal, and Rishi Saket. On the hardness of learning sparse parities. In *ESA*, pages 11:1–11:17, 2016.
- [BGS17] Huck Bennett, Alexander Golovnev, and Noah Stephens-Davidowitz. On the quantitative hardness of CVP. In *FOCS*, pages 13–24, 2017.
- [BMvT78] Elwyn R. Berlekamp, Robert J. McEliece, and Henk C. A. van Tilborg. On the inherent intractability of certain coding problems (corresp.). *IEEE Trans. Information Theory*, 24(3):384–386, 1978.
- [BR60] R. C. Bose and Dwijendra K. Ray-Chaudhuri. On a class of error correcting binary group codes. *Information and Control*, 3(1):68–79, 1960.
- [BS08] Eli Ben-Sasson and Madhu Sudan. Short PCPs with polylog query complexity. *SIAM J. Comput.*, 38(2):551–607, 2008.
- [CCK<sup>+</sup>17] Parinya Chalermsook, Marek Cygan, Guy Kortsarz, Bundit Laekhanukit, Pasin Manurangsi, Danupon Nanongkai, and Luca Trevisan. From Gap-ETH to FPT-inapproximability: Clique, dominating set, and more. In *FOCS*, pages 743–754, 2017.
- [CFHW17] Marek Cygan, Fedor V. Fomin, Danny Hermelin, and Magnus Wahlström. Randomization in parameterized complexity (dagstuhl seminar 17041). *Dagstuhl Reports*, 7(1):103–128, 2017.
- [CFJ<sup>+</sup>14] Marek Cygan, Fedor Fomin, Bart MP Jansen, Lukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, and Saket Saurabh. Open problems for fpt school 2014. 2014.
- [CFK<sup>+</sup>15] Marek Cygan, Fedor V. Fomin, Lukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michal Pilipczuk, and Saket Saurabh. *Parameterized Algorithms*. Springer, 2015.
- [CFM17] Rajesh Chitnis, Andreas Emil Feldmann, and Pasin Manurangsi. Parameterized approximation algorithms for directed steiner network problems. *CoRR*, abs/1707.06499, 2017.
- [CN99] Jin-yi Cai and Ajay Nerurkar. Approximating the SVP to within a factor  $(1 + 1/\dim^\xi)$  is NP-hard under randomized reductions. *J. Comput. Syst. Sci.*, 59(2):221–239, 1999.
- [CW12] Qi Cheng and Daqing Wan. A deterministic reduction for the gap minimum distance problem. *IEEE Trans. Information Theory*, 58(11):6935–6941, 2012.
- [DF99] Rodney G. Downey and Michael R. Fellows. *Parameterized Complexity*. Monographs in Computer Science. Springer, 1999.

- [DF13] Rodney G. Downey and Michael R. Fellows. *Fundamentals of Parameterized Complexity*. Texts in Computer Science. Springer, 2013.
- [DFVW99] Rodney G. Downey, Michael R. Fellows, Alexander Vardy, and Geoff Whittle. The parametrized complexity of some fundamental problems in coding theory. *SIAM J. Comput.*, 29(2):545–570, 1999.
- [DGMS07] Erik D. Demaine, Gregory Gutin, Dániel Marx, and Ulrike Stege. 07281 open problems – structure theory and FPT algorithmics for graphs, digraphs and hypergraphs. In *Structure Theory and FPT Algorithmics for Graphs, Digraphs and Hypergraphs*, 08.07. - 13.07.2007, 2007.
- [Din02] Irit Dinur. Approximating  $SVP_\infty$  to within almost-polynomial factors is NP-hard. *Theor. Comput. Sci.*, 285(1):55–71, 2002.
- [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. *ECCC*, 23:128, 2016.
- [DKRS03] Irit Dinur, Guy Kindler, Ran Raz, and Shmuel Safra. Approximating CVP to within almost-polynomial factors is NP-hard. *Combinatorica*, 23(2):205–243, 2003.
- [DM18] Irit Dinur and Pasin Manurangsi. ETH-hardness of approximating 2-CSPs and directed steiner network. In *ITCS*, pages 36:1–36:20, 2018.
- [DMS03] Ilya Dumer, Daniele Micciancio, and Madhu Sudan. Hardness of approximating the minimum distance of a linear code. *IEEE Trans. Information Theory*, 49(1):22–37, 2003.
- [FGMS12] Michael R. Fellows, Jiong Guo, Dániel Marx, and Saket Saurabh. Data reduction and problem kernels (dagstuhl seminar 12241). *Dagstuhl Reports*, 2(6):26–50, 2012.
- [GKS12] Petr A. Golovach, Jan Kratochvíl, and Ondrej Suchý. Parameterized complexity of generalized domination problems. *Discrete Applied Mathematics*, 160(6):780–792, 2012.
- [GMSS99] Oded Goldreich, Daniele Micciancio, Shmuel Safra, and Jean-Pierre Seifert. Approximating shortest lattice vectors is not harder than approximating closest lattice vectors. *Inf. Process. Lett.*, 71(2):55–61, 1999.
- [Gol06] Oded Goldreich. On promise problems: A survey. In *Theoretical Computer Science, Essays in Memory of Shimon Even*, pages 254–290, 2006.
- [Hoc59] Alexis Hocquenghem. Codes correcteurs d’erreurs. *Chiffres*, 2:147–156, September 1959.
- [HR07] Ishay Haviv and Oded Regev. Tensor-based hardness of the shortest vector problem to within almost polynomial factors. In *STOC*, pages 469–477, 2007.
- [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.

- [Joh90] D. S. Johnson. Handbook of theoretical computer science. volume A (Algorithms and Complexity), chapter 2, A catalog of complexity classes, pages 67–161. Elsevier, 1990.
- [Kho05] Subhash Khot. Hardness of approximating the shortest vector problem in lattices. *J. ACM*, 52(5):789–808, 2005.
- [KLM18] Karthik C. S., Bundit Laekhanukit, and Pasin Manurangsi. On the parameterized complexity of approximating dominating set. In *STOC*, 2018. To appear.
- [Len83] Hendrik Willem Lenstra. Integer programming with a fixed number of variables. *Math. Oper. Res.*, 8(4):538–548, 1983.
- [Lin15] Bingkai Lin. The parameterized complexity of  $k$ -Biclique. In *SODA*, pages 605–615, 2015.
- [LLL82] Arjen Klaas Lenstra, Hendrik Willem Lenstra, and László Lovász. Factoring polynomials with rational coefficients. *Mathematische Annalen*, 261(4):515–534, 1982.
- [LRSZ17] Daniel Lokshtanov, M. S. Ramanujan, Saket Saurabh, and Meirav Zehavi. Parameterized complexity and approximability of directed odd cycle transversal. *CoRR*, abs/1704.04249, 2017.
- [Maj17] Ruhollah Majdoddin. Parameterized complexity of CSP for infinite constraint languages. *CoRR*, abs/1706.10153, 2017.
- [Man17] Pasin Manurangsi. Almost-polynomial ratio ETH-hardness of approximating densest  $k$ -subgraph. In *STOC*, pages 954–961, 2017.
- [Mic00] Daniele Micciancio. The shortest vector in a lattice is hard to approximate to within some constant. *SIAM J. Comput.*, 30(6):2008–2035, 2000.
- [Mic01] Daniele Micciancio. The hardness of the closest vector problem with preprocessing. *IEEE Trans. Information Theory*, 47(3):1212–1215, 2001.
- [Mic12] Daniele Micciancio. Inapproximability of the shortest vector problem: Toward a deterministic reduction. *Theory of Computing*, 8(1):487–512, 2012.
- [Mic14] Daniele Micciancio. Locally dense codes. In *CCC*, pages 90–97, 2014.
- [MR09] Daniele Micciancio and Oded Regev. Lattice-based cryptography. In *Post-quantum cryptography*, pages 147–191. Springer, 2009.
- [MR16] Pasin Manurangsi and Prasad Raghavendra. A birthday repetition theorem and complexity of approximating dense CSPs. *CoRR*, abs/1607.02986, 2016.
- [NV10] Phong Q. Nguyen and Brigitte Vallée, editors. *The LLL Algorithm - Survey and Applications*. Information Security and Cryptography. Springer, 2010.
- [Reg03] Oded Regev. New lattice based cryptographic constructions. In *STOC*, pages 407–416, 2003.
- [Reg05] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In *STOC*, pages 84–93, 2005.

- [Reg06] Oded Regev. Lattice-based cryptography. In *CRYPTO*, pages 131–141, 2006.
- [Reg10] Oded Regev. The learning with errors problem (invited survey). In *CCC*, pages 191–204, 2010.
- [RR06] Oded Regev and Ricky Rosen. Lattice problems and norm embeddings. In *STOC*, pages 447–456, 2006.
- [Ste93] Jacques Stern. Approximating the number of error locations within a constant ratio is NP-complete. In *AAECC*, pages 325–331, 1993.
- [Var97a] Alexander Vardy. Algorithmic complexity in coding theory and the minimum distance problem. In *STOC*, pages 92–109, 1997.
- [Var97b] Alexander Vardy. The intractability of computing the minimum distance of a code. *IEEE Trans. Information Theory*, 43(6):1757–1766, 1997.
- [vEB81] Peter van Emde-Boas. *Another NP-complete partition problem and the complexity of computing short vectors in a lattice*. Report. Department of Mathematics. University of Amsterdam. Department, Univ., 1981.

## A A Simple Proof of PIH from Gap-ETH

In this section, we provide a simple proof of PIH assuming Gap-ETH. The proof is folklore and is known in the community, although we are not aware of it being fully written down anywhere. The main technical ingredient is the following subexponential time reduction from 3SAT to 2CSP:

**Lemma A.1** *For any  $\varepsilon > 0$  and  $\Delta, k \in \mathbb{N}$  such that  $k \geq 2$ , there exists a randomized reduction that, given any 3SAT instance  $\Phi$  on  $n$  variables where each variable appears in at most  $\Delta$  clauses, produces a 2CSP instance  $\Gamma = (G = (V, E), \Sigma, \{C_{uv}\}_{(u,v) \in E})$  such that*

- (YES) If  $\text{SAT}(\Phi) = 1$ , then  $\text{val}(\Gamma) = 1$ .
- (NO) If  $\text{SAT}(\Phi) < 1 - \varepsilon$ , then  $\text{val}(\Gamma) < 1 - \frac{\varepsilon}{3000\Delta^4}$  with high probability.
- (Running Time) The reduction runs in time  $2^{O_{\Delta, \varepsilon}(n/k)}$ , and,
- (Parameter) The number of variables of  $\Gamma$  is equal to  $k$ , i.e.,  $|V| = k$ .

Note that, for Gap-ETH, it can be assumed without loss of generality that each variable of the 3CNF formula appears in at most  $O(1)$  number of clauses (see e.g. [MR16, page 21]). Observe that, since the reduction runs in time  $2^{O(n/k)}$ ;  $\Gamma$  is of size at most  $2^{O(n/k)}$ . This yields the following corollary, which in fact provides a running time lower bound which is even stronger than PIH:

**Corollary A.2** *Assuming Gap-ETH, there exists  $\varepsilon > 0$  such that, for any computable function  $T$ , there is no  $T(k) \cdot |\Sigma|^{o(k)}$  algorithm for  $\text{GAP2CSP}_\varepsilon$ .*

We note here that the running time lower bound in the above lemma is asymptotically optimal, because we can solve 2CSP exactly in  $\text{poly}(k) \cdot |\Sigma|^k$  time by just enumerating all possible assignments and outputting the one that satisfies the maximum number of constraints.

We now turn our attention to the proof of Lemma A.1.

*Proof of Lemma A.1.* The reduction is simple and is in fact a special case of the reduction from [DM18]. Let  $X$  and  $\mathcal{C}$  denote the set of variables and the set of clauses of  $\Phi$  respectively. We

assume w.l.o.g. that each variable appears in at least one clause. For each  $C \in \mathcal{C}$ , we use  $\text{var}(C)$  to denote the variables appearing in  $C$ . Furthermore, for each  $S \subseteq \mathcal{C}$ , let  $\text{var}(S) = \cup_{C \in S} \text{var}(C)$ . For convenience, we will also assume that  $m$  is divisible by  $k$  where  $m = |\mathcal{C}|$  is the number of clauses.

Randomly partition  $\mathcal{C}$  into  $k$  parts  $S_1, \dots, S_k$  each of size  $m/k$ . We can then define  $\Gamma$  as follows:

- The vertex set  $V$  is  $[k]$  and the constraint graph  $G$  is the complete graph over  $V$ .
- Let  $\Sigma = 2^{3(m/k)}$ . For each  $i \in [k]$ , we associate elements of  $\Sigma$  with the partial assignments  $f_i : \text{var}(S_i) \rightarrow \{0, 1\}$  that satisfies all clauses in  $S_i$ . In other words, we view  $\Sigma$  as  $\Sigma_i = \{f_i : \text{var}(S_i) \rightarrow \{0, 1\} \mid f_i \text{ satisfies all } C \in S_i\}$ .
- For each edge  $\{i, j\} \subseteq [k]$ , we define the constraint  $C_{ij}$  to contain all  $f_i \in \Sigma_i$  and  $f_j \in \Sigma_j$  that agree with each other, i.e.,  $C_{ij} = \{(f_i, f_j) \in \Sigma_i \times \Sigma_j \mid \forall x \in \text{var}(S_i) \cap \text{var}(S_j), f_i(x) = f_j(x)\}$ .

It is clear that the reduction runs in  $2^{O(m/k)} = 2^{O(\Delta n / (\epsilon k))}$  time and  $|V| = k$ . Further, in the YES case where there exists an assignment  $f : X \rightarrow \{0, 1\}$  that satisfies all the clauses of  $\Phi$ , we can simply pick  $f_i$  to be  $f|_{\text{var}(S_i)}$ , the restriction of  $f$  on  $\text{var}(S_i)$ . Clearly,  $\psi$  is a valid assignment for  $\Gamma$  and it satisfies all the constraints. Hence,  $\text{val}(\Gamma) = 1$ .

To show the NO case of the reduction, we will need an additional property of a random partition  $S_1, \dots, S_k$ , which is formally stated below and proved later:

**Proposition A.3** *With high probability, for every  $i \neq j$ , we have  $\text{var}(S_i) \cap \text{var}(S_j) < \frac{1000n\Delta^3}{k^2}$ .*

Now, conditioned on  $\text{var}(S_i) \cap \text{var}(S_j) < \frac{1000n\Delta^3}{k^2}$  for every  $i \neq j \in [k]$ , we will prove the NO case by showing the contrapositive of the statement. Suppose that there exists an assignment  $\psi : V \rightarrow \Sigma$  such that  $\text{val}(\psi) \geq 1 - \epsilon'$  where  $\epsilon' = \frac{\epsilon}{3000\Delta^4}$ . For notational convenience, let  $f_i = \psi(i)$ . Observe that  $\text{val}(\psi) \geq 1 - \epsilon'$  is simply equivalent to  $\Pr_{\{i,j\} \subseteq [k]} [f_i \approx f_j] \leq \epsilon'$  where  $f_i \approx f_j$  denote  $f_i(x) \neq f_j(x)$  for some  $x \in \text{var}(S_i) \cap \text{var}(S_j)$ .

We call a variable  $x \in X$  *globally consistent* if  $f_i(x)$  is equal for every  $i \in [k]$  such that  $\text{var}(S_i)$  contains  $x$ . We claim that the number of globally consistent variables is at least  $(1 - 1000\epsilon'\Delta^3)n$ . To see this, let us count the number of  $(x, \{i, j\}) \in X \times \binom{[k]}{2}$  such that  $x \in \text{var}(S_i) \cap \text{var}(S_j)$  and  $f_i(x) \neq f_j(x)$ . Notice that such tuple exists only for  $\{i, j\}$  such that  $f_i \approx f_j$ , and there are only  $\epsilon' \binom{k}{2}$  such pairs  $\{i, j\}$ 's. Recall also that  $\text{var}(S_i) \cap \text{var}(S_j) \leq \frac{1000n\Delta^3}{k^2}$  for any  $i \neq j \in [k]$ . As a result, the number of such  $(x, \{i, j\})$ 's is at most  $\epsilon' \binom{k}{2} \cdot \frac{1000n\Delta^3}{k^2} \leq 1000\epsilon'\Delta^3n$ . On the other hand, each variable  $x \in X$  that is not globally consistent must contribute to at least one such pair; hence, at most  $1000\epsilon'\Delta^3n$  variables are not globally consistent.

Let us define a global assignment  $f : X \rightarrow \{0, 1\}$  as follows: for each  $x \in X$ , pick an arbitrary partition  $S_i$  such that  $x \in \text{var}(S_i)$  and let  $f(x) = f_i(x)$ . Observe that, for every clause  $C$  such that all its variables are globally consistent,  $C$  is satisfied by  $f$ ; this is simply because, if  $C$  is in partition  $S_j$ , then  $f(x) = f_j(x)$  for all  $x \in \text{var}(C)$  due to global consistency of  $x$  and  $f_j$  must satisfy  $C$  from our definition of  $\Gamma$ . As a result,  $f$  satisfies all but at most  $1000\epsilon'\Delta^4n \leq 3000\epsilon'\Delta^4m$  clauses, which means that  $\text{val}(\Phi) \geq 1 - 3000\epsilon'\Delta^4 = 1 - \epsilon$  as desired.  $\square$

*Proof of Proposition A.3.* For convenience, let  $\ell = \frac{1000n\Delta^3}{k^2}$ . Let us fix a pair  $i \neq j$ . We will show that  $\text{var}(S_i) \cap \text{var}(S_j) < \ell$  with high probability; taking union bound over all  $\{i, j\}$ 's yields the desired

statement.

To bound  $\Pr[\text{var}(S_i) \cap \text{var}(S_j) < \ell]$ , let us define one more notation: we call a set  $T \subseteq X$  of variables *independent* if no two variables in  $T$  appears in the same clause. We claim that, for any subset  $T'$  of at least  $\ell$  variables, there exists an independent subset  $T \subseteq T'$  of at least  $\lceil \ell / (2\Delta + 1) \rceil$  variables. This is true because we can use the following algorithm to find  $T$ : as long as  $T'$  is not empty, pick an arbitrary variable  $x \in T'$ , put  $x$  into  $T$ , and remove  $x$  and all variables that share at least one clause with  $x$  from  $T'$ . Since each time we add an element to  $T$ , we remove at most  $2\Delta + 1$  elements from  $T'$ ; we can conclude that  $|T| \geq \lceil \ell / (2\Delta + 1) \rceil$  as desired.

Hence, to show that  $\text{var}(S_i) \cap \text{var}(S_j) < \ell$  w.h.p., it suffices to show that  $\text{var}(S_i) \cap \text{var}(S_j)$  does not contain an independent set of variable of size  $\lceil \ell / (2\Delta + 1) \rceil$  w.h.p.

Let  $r = \lceil \ell / (2\Delta + 1) \rceil$ ; notice that  $r \geq \frac{100n\Delta^2}{k^2}$ . Consider any independent subset of variables  $T \subseteq X$  of size  $r$ . We will upper bound the probability that  $T$  is contained in  $\text{var}(S_i) \cap \text{var}(S_j)$ . To do so, first observe that the events  $T \subseteq \text{var}(S_i)$  and  $T \subseteq \text{var}(S_j)$  are negatively correlated, i.e.,

$$\Pr[T \subseteq \text{var}(S_i) \cap \text{var}(S_j)] \leq \Pr[T \subseteq \text{var}(S_i)] \Pr[T \subseteq \text{var}(S_j)] = (\Pr[T \subseteq \text{var}(S_i)])^2. \quad (10)$$

Observe that since two distinct  $x, x' \in T$  do not appear in the same clause, the events  $x \in \text{var}(S_i)$  and  $x' \in \text{var}(S_i)$  are negatively correlated. That is, if  $T = \{x_{p_1}, \dots, x_{p_r}\}$ , then we can bound  $\Pr[T \subseteq \text{var}(S_i)]$  as follows:

$$\Pr[T \subseteq \text{var}(S_i)] = \prod_{j \in [r]} \Pr[x_{p_j} \in S_i \mid x_{p_1}, \dots, x_{p_{j-1}} \in S_i] \leq \prod_{j \in [r]} \Pr[x_{p_j} \in S_i]. \quad (11)$$

Now, consider any  $x \in T$ . Since  $x$  is contained in at most  $\Delta$  clauses, we have

$$\Pr[x \in S_i] \leq 1 - \frac{\binom{m-\Delta}{m/k}}{\binom{m}{m/k}} \leq 1 - \left(1 - \frac{\Delta}{m - m/k}\right)^{m/k} \leq \frac{\Delta(m/k)}{m - m/k} = \frac{\Delta}{k-1}. \quad (12)$$

where the last inequality follows from Bernoulli's inequality.

Combining (10), (11) and (12), we have

$$\Pr[T \subseteq \text{var}(S_i) \cap \text{var}(S_j)] \leq \left(\frac{\Delta}{k-1}\right)^{2r}.$$

Finally, by taking union bound over all  $r$ -size independent sets of variables, the probability that  $\text{var}(S_i) \cap \text{var}(S_j)$  contains an  $r$ -size independent sets of variables is at most

$$\binom{n}{r} \left(\frac{\Delta}{k-1}\right)^{2r} \leq \left(\frac{en}{r}\right)^r \left(\frac{\Delta}{k-1}\right)^{2r} = \left(\frac{en\Delta^2}{r(k-1)^2}\right)^r \leq (1/2)^r = o(1),$$

which concludes our proof.  $\square$



## B FPT Inapproximability of GAPNVP

In this section, we will prove the inapproximability of GAPNVP (Theorem 6.3). The reduction goes through the following intermediate problem GAPLVS (analogous to the GAPMLD):

$\eta$ -Gap Lattice Vector Sum (GAPLVS $_{\eta}$ )

**Input:** A matrix  $\mathbf{A} \in \mathbb{Z}^{m \times n}$ , a vector  $\mathbf{y} \in \{0, 1\}^m$  and a parameter  $k \in \mathbb{N}$ .

**Parameter:**  $k$

**Question:** Distinguish between the following two cases:

- (YES) There exists  $\mathbf{x} \in \{0, 1\}^n$  such that  $\|\mathbf{x}\|_0 \leq k$  and  $\mathbf{A}\mathbf{x} = \mathbf{y}$ .
- (NO) For all choices of  $\mathbf{x} \in \mathbb{Z}^n$  such that  $\|\mathbf{x}\|_0 \leq \eta \cdot k$  and  $r \in \mathbb{Z} \setminus \{0\}$ , we have  $\mathbf{A}\mathbf{x} \neq r \cdot \mathbf{y}$ .

The above problem is well defined for any  $\eta \geq 1$ . The reduction for Theorem 6.3 goes through the following steps. First, we use Lemma B.1 to give an FPT reduction from GAP2CSP $_{\varepsilon}$  to GAPLVS $_{1+\varepsilon/3}$ . Next, we use the gap amplification operation (Lemma B.2) to give a FPT reduction from GAPLVS $_{1+\varepsilon/3}$  to all GAPLVS $_{\eta}$  for all  $\eta \geq 1$ . Finally, we use Lemma B.3 to reduce GAPLVS $_{\eta}$  to GAPNVP $_{p,\eta}$  (for any  $p \geq 1$ ).

**Lemma B.1** *For all  $\varepsilon > 0$ , there is an FPT reduction from GAP2CSP $_{\varepsilon}$  to GAPLVS $_{1+\varepsilon/3}$ .*

*Proof.* Let  $\Gamma = (G := (V, E), \Sigma, \{C_{uv}\}_{(u,v) \in E})$  be the GAP2CSP $_{\varepsilon}$  instance. For clarity, we assume that  $G$  is directed. From  $\Gamma$ , we construct  $\mathbf{A} \in \{-1, 0, 1\}^{m \times n}$ ,  $\mathbf{y} \in \{0, 1\}^m$  where  $m = |V| + |E| + 2|E||\Sigma|$  and  $n = |V||\Sigma| + \sum_{(u,v) \in E} |C_{uv}|$  exactly the same as in the proof of Lemma 4.2 with only one difference: we set  $\mathbf{A}_{(e,\sigma_0,0),(e,\sigma_0,\sigma_1)}$  and  $\mathbf{A}_{(e,\sigma_1,1),(e,\sigma_0,\sigma_1)}$  to be -1 instead of 1 as in Lemma 4.2.

We set the sparsity parameter to be  $k = |V| + |E|$ . Note that the matrix  $\mathbf{A}$  and the vector  $\mathbf{y}$  here are  $\{-1, 0, 1\}$ -objects over  $\mathbb{Z}$  instead of  $\mathbb{F}_2$ . Additionally, by the construction of  $\mathbf{A}$ , for every  $\mathbf{x} \in \mathbb{Z}^n$ , the vector  $\mathbf{A}\mathbf{x}$  must satisfy equations (2) and (3) from Section 4 whereas, due to our change, equation (4) now becomes

$$(\mathbf{A}\mathbf{x})_{(e,\sigma,b)} = \mathbf{x}_{(u_b,\sigma)} - \sum_{\substack{(\sigma_0,\sigma_1) \in C_e \\ \sigma_b = \sigma}} \mathbf{x}_{(e,\sigma_0,\sigma_1)}.$$

We now argue completeness and soundness of the reduction.

**Completeness.** If  $\Gamma$  is a YES instance, let  $\psi : V \mapsto \Sigma$  be the satisfying labeling. We construct  $\mathbf{x} \in \{0, 1\}^n$  as follows: we set  $\mathbf{x}_{(u,\psi(u))} = 1$  for every vertex  $u \in V$ . Additionally, for every  $e = (u, v) \in E$ , we set  $\mathbf{x}_{((u,v),\psi(u),\psi(v))} = 1$ , and all the remaining coordinates of  $\mathbf{x}$  are set to zero. It is easy to see that  $\mathbf{x}$  as constructed is  $(|V| + |E|)$ -sparse, and satisfies  $\mathbf{A}\mathbf{x} = \mathbf{y}$ .

**Soundness.** Let  $\mathbf{x} \in \mathbb{Z}^n$  be such that  $\|\mathbf{x}\|_0 \leq (1 + \varepsilon/3)k$  and  $\mathbf{A}\mathbf{x} = r \cdot \mathbf{y}$  for some  $r \neq 0$ . As before, for every vertex  $u \in V$ , we construct the set  $S_u := \{\sigma | \mathbf{x}_{(u,\sigma)} \neq 0\}$ . Similarly, for every edge  $e \in E$ , we construct the set  $T_e := \{(\sigma_0, \sigma_1) | \mathbf{x}_{(e,\sigma_0,\sigma_1)} \neq 0\}$ . By construction it follows that for every  $u \in V$ ,  $(\mathbf{A}\mathbf{x})_u = y_u = r \neq 0$  and therefore  $|S_u| \geq 1$ . Similarly, for every edge  $e \in E$  we have  $|T_e| \geq 1$ .

Furthermore, we define  $E_{\text{unique}} := \{e \in E \mid |T_e| = 1\}$ . Since the arguments used in Equations (5)-(6) (from proof of Lemma 4.2) depend only on the fact that  $\|\mathbf{x}\|_0 \leq (1 + \varepsilon/3)k$ , again we have  $|E_{\text{unique}}| \geq (1 - \varepsilon)|E|$ .

Now, consider any labeling  $\psi : V \mapsto \Sigma$  such that  $\psi(u) \in S_u$  for every  $u \in V$ . We shall show that any such labeling must satisfy every edge  $e \in E_{\text{unique}}$ . Towards that end, we fix an edge  $e = (u, v) \in E_{\text{unique}}$  and let  $T_e = \{(\sigma_0^*, \sigma_1^*)\}$ . Since  $|T_e| = 1$ , using equation (3) we have

$$r = \mathbf{y}_e = \sum_{(\sigma_0, \sigma_1) \in T_e} \mathbf{x}_{(e, \sigma_0, \sigma_1)} = \mathbf{x}_{(e, \sigma_0^*, \sigma_1^*)}$$

Observe that  $S_u = \{\sigma_0^*\}$ ; this follows immediately from substituting (the new) equation (4) with  $\sigma = \sigma_0^*, b = 0$  and from  $\mathbf{y}_{(e, \sigma_0^*, 0)} = 0$ . Therefore,  $\psi(u) = \sigma_0^*$  and, similarly,  $\psi(v) = \sigma_1^*$ . Furthermore, since  $(\sigma_0^*, \sigma_1^*) \in C_{(u, v)}$ ,  $\psi$  satisfies the edge  $e$ . Since, this happens for every  $e \in E_{\text{unique}}$ , the labeling  $\psi$  satisfies every constraint in  $E_{\text{unique}}$ . This completes the proof.  $\square$

We now give the lemma for the gap amplification step.

**Lemma B.2** *For any  $c \in \mathbb{N}$  and  $\eta \geq 1$ , there exists an FPT reduction from  $\text{GAPLVS}_\eta$  to  $\text{GAPLVS}_{\eta^c}$ .*

*Proof.* As in Section 4.2, we define a composition operation  $\oplus_{\mathcal{L}}$  on  $\text{GAPLVS}$  instances. Given  $\text{GAPLVS}_{\eta_1}$  instance  $(\mathbf{A}, \mathbf{y}_1, k_1)$  (where  $\mathbf{A} \in \mathbb{Z}^{u \times v}, \mathbf{y}_1 \in \mathbb{Z}^u$ ) and  $\text{GAPLVS}_{\eta_2}$  instance  $(\mathbf{B}, \mathbf{y}_2, k_2)$  (where  $\mathbf{B} \in \mathbb{Z}^{u' \times v'}, \mathbf{y}_2 \in \mathbb{Z}^{u'}$ ), we denote the composed instance as  $(\mathbf{C}, \mathbf{y}, k') = (\mathbf{A}, \mathbf{y}_1, k_1) \oplus_{\mathcal{L}} (\mathbf{B}, \mathbf{y}_2, k_2)$ , with  $k' = k_2 + k_1 k_2$ . The matrix  $\mathbf{C} \in \mathbb{Z}^{(u'+uv') \times (v'+vv')}$  and the vector  $\mathbf{y} \in \mathbb{Z}^{u'+uv'}$  is constructed exactly as in Section 4.2. We argue the completeness and soundness properties of the operator.

**Completeness** : Let  $(\mathbf{A}, \mathbf{y}_1, k_1)$  and  $(\mathbf{B}, \mathbf{y}_2, k_2)$  be YES instances. Then there exists vector  $\mathbf{x}_1 \in \{0, 1\}^v$  such that  $\|\mathbf{x}_1\|_0 \leq k_1$  and  $\mathbf{A}\mathbf{x}_1 = \mathbf{y}_1$ . Similarly there exists vector  $\mathbf{x}_2 \in \{0, 1\}^{v'}$  such that  $\|\mathbf{x}_2\|_0 \leq k_2$  and  $\mathbf{B}\mathbf{x}_2 = \mathbf{y}_2$ . Borrowing notation from Section 4.2, for any  $\mathbf{x} \in \mathbb{Z}^n$ , we define the sub-vectors  $\mathbf{x}^0, \mathbf{x}^1, \dots, \mathbf{x}^{v'}$ . We construct the vector  $\mathbf{x} \in \{0, 1\}^{v'+vv'}$  as follows. We set  $\mathbf{x}^0 = \mathbf{x}_2$  and for every  $i \in [v']$  we set  $\mathbf{x}^i = x_i^0 \cdot \mathbf{x}_1$ . It is easy to check that  $\mathbf{x}$  as constructed satisfies  $\|\mathbf{x}\|_0 \leq k_2 + k_1 k_2$  and  $\mathbf{C}\mathbf{x} = \mathbf{y}$ .

**Soundness** : Let  $(\mathbf{A}, \mathbf{y}_1, k_1)$  and  $(\mathbf{B}, \mathbf{y}_2, k_2)$  be NO instances. As in the the proof of Lemma 4.5, we consider the row-blocks  $S_0, S_1, \dots, S_{v'}$  which form an identical partition of  $[u' + uv']$ . Let  $\mathbf{x} \in \mathbb{Z}^{v'+vv'}$  be any vector such that  $\mathbf{C}\mathbf{x} = r \cdot \mathbf{y}$  for some  $r \in \mathbb{Z} \setminus \{0\}$ . Since  $\mathbf{x}^0$  satisfies the constraints along  $S_0$ , we have  $\mathbf{B}\mathbf{x}^0 = r \cdot \mathbf{y}_2$ , and therefore  $\|\mathbf{x}^0\|_0 > \eta_2 \cdot k_2$ . Fix any  $i \in [v']$  such that  $x_i^0 \neq 0$ . Then the sub-vector  $\mathbf{x}^i$  must satisfy the constraints along row-block  $S_i$  i.e.,  $\mathbf{A}\mathbf{x}^i = r x_i^0 \cdot \mathbf{y}_1$  and therefore, we have  $\|\mathbf{x}^i\|_0 > \eta_1 k_1$ . Since, this happens for every such choice of  $i \in [v']$  such that  $x_i^0 \neq 0$ , we have  $\|\mathbf{x}\|_0 > \eta_2 k_2 + \eta_1 \eta_2 k_1 k_2$ .

Therefore, given  $\text{GAPLVS}_\eta$  instance  $(\mathbf{A}, \mathbf{y}, k)$ , we can construct the  $i$ -wise composed instance  $(\mathbf{A}^{(i)}, \mathbf{y}^{(i)}, k(k+1)^{i-1}) = (\mathbf{A}, \mathbf{y}, k) \oplus_{\mathcal{L}} (\mathbf{A}^{(i-1)}, \mathbf{y}^{(i-1)}, k(k+1)^{i-2})$ . Using the above arguments, it is easy to check that  $(\mathbf{A}^{(i)}, \mathbf{y}^{(i)}, k(k+1)^{i-1})$  is a  $\text{GAPLVS}_{\eta^{3i/2}}$  instance. Setting  $i \geq \lceil 3c/2 \rceil$  gives us the required gap.  $\square$

Finally, the following lemma completes the reduction to GAPNVP.

**Lemma B.3** *For all choices of  $\eta \geq 1$  and  $p \geq 1$ , there is an FPT reduction from  $k$ -GAPLVS $_\eta$  to  $k$ -GAPNVP $_{p,\eta}$ . Furthermore, the resulting GAPNVP $_{p,\eta}$  instance  $(\mathbf{A}', \mathbf{y}', k)$  (where  $\mathbf{A}' \in \mathbb{Z}^{n' \times m}$  and  $\mathbf{y}' \in \mathbb{Z}^{n'}$ ) satisfies the following properties:*

- For the YES case, there exists  $\mathbf{x} \in \mathbb{Z}^m$  such that  $(\mathbf{A}'\mathbf{x} - \mathbf{y}')$  is a  $\{0, 1\}$ -vector and  $\|\mathbf{A}'\mathbf{x} - r \cdot \mathbf{y}'\|_0 \leq k$ .
- For the NO case, for all  $\mathbf{x} \in \mathbb{Z}^m$  and  $r \in \mathbb{Z} \setminus \{0\}$ , we have  $\|\mathbf{A}'\mathbf{x} - r \cdot \mathbf{y}'\|_0 > \eta \cdot k$

*Proof.* Let  $(\mathbf{A}, \mathbf{y}, k)$  be a GAPLVS $_\eta$  instance where  $\mathbf{A} \in \mathbb{Z}^{n \times m}$  and  $\mathbf{y} \in \mathbb{Z}^n$ . We construct the matrix  $\mathbf{A}' \in \mathbb{Z}^{(\lceil \eta k + 1 \rceil n + m) \times m}$  and  $\mathbf{y}' \in \mathbb{Z}^{\lceil \eta k + 1 \rceil n + m}$  identical to the proof of Lemma 5.2, and claim that  $(\mathbf{A}', \mathbf{y}', k)$  is a GAPNVP $_\eta$  instance with the additional properties as guaranteed by the lemma.

Indeed, if  $(\mathbf{A}, \mathbf{y}, k)$  is a YES instance, then there exists  $\mathbf{x} \in \{0, 1\}^m$  such that  $\|\mathbf{x}\|_0 \leq k$  such that  $\mathbf{A}\mathbf{x} = \mathbf{y}$ . Consequently,  $\mathbf{A}'\mathbf{x} - \mathbf{y}'$  is 0 on the first  $\lceil \eta k + 1 \rceil n$  coordinates, and it is just  $\mathbf{x}$  on the last  $m$  coordinates. Conversely, if  $(\mathbf{A}, \mathbf{y}, k)$  is a NO instance, then as before we consider two cases. If  $\mathbf{x} \in \mathbb{Z}^m$  such that  $\|\mathbf{x}\|_0 > \eta k$ , then it is easy to see that  $\mathbf{A}'\mathbf{x} - \mathbf{y}'$  has Hamming weight greater than  $\eta k$  on the last  $m$ -coordinates. On the other hand if  $\|\mathbf{x}\|_0 \leq \eta k$ , then we know that  $\mathbf{A}\mathbf{x} \neq r \cdot \mathbf{y}$  for any  $r \in \mathbb{Z} \setminus \{0\}$ . Therefore  $\|\mathbf{A}\mathbf{x} - r \cdot \mathbf{y}\|_0 \geq 1$  which implies that  $\|\mathbf{A}'\mathbf{x} - r \cdot \mathbf{y}'\|_0 \geq \eta k$ .

Finally, since in the YES case the witness is a  $\{0, 1\}$ -vector and in the NO case the guarantees are in the Hamming norm, the completeness and soundness properties are satisfied by  $(\mathbf{A}', \mathbf{y}', k)$  for any value of  $p \geq 1$ .  $\square$

**Putting Things Together:** Given a GAP2CSP $_\varepsilon$ -instance  $\Gamma$ , using Lemma B.1 we construct a GAPLVS $_{1+\varepsilon/3}$ -instance. Using Lemma B.2, we reduce GAPLVS $_{1+\varepsilon/3}$  to GAPLVS $_\eta$  for any choice of  $\eta \geq 1 + \varepsilon/3$ . Finally, using Lemma B.3, we reduce the GAPLVS $_\eta$  instance to GAPNVP $_{p,\eta}$  instance. Since all the steps in the reduction are FPT, this completes the proof.