

# Non-Signaling Locality Lower Bounds for Dominating Set

Noah Fleming\*  
Lund & Columbia  
noah.fleming@cs.lth.se

Max Hopkins†  
IAS  
nmhopkin@ias.edu

Yuichi Yoshida‡  
National Institute of Informatics  
yyoshida@nii.ac.jp

April 6, 2026

## Abstract

Minimum dominating set is a basic local covering problem and a core task in distributed computing. Despite extensive study, in the classic LOCAL model there exist significant gaps between known algorithms and lower bounds. Chang and Li (PODC '23) prove an  $\Omega(\log n)$ -locality lower bound for a constant factor approximation, while Kuhn–Moscibroda–Wattenhofer (JACM '16) provide an algorithm beating this bound beyond  $\log \Delta$ -approximation, along with a weaker lower bound for this degree-dependent setting scaling roughly with  $\min\{\log \Delta / \log \log \Delta, \sqrt{\log n / \log \log n}\}$ . Unfortunately, this latter bound is very weak for small  $\Delta$ , and never recovers the Chang–Li bound, leaving central questions: does  $O(\log \Delta)$ -approximation require  $\Omega(\log n)$  locality, and do such bounds extend beyond LOCAL?

In this work, we take a major step toward answering these questions in the non-signaling model, which strictly subsumes the LOCAL, quantum-LOCAL, and bounded-dependence settings. We prove every  $O(\log \Delta)$ -approximate non-signaling distribution for dominating set requires locality  $\Omega(\log n / (\log \Delta \cdot \text{poly} \log \log \Delta))$ . Further, we show for some  $\beta \in (0, 1)$ , every  $O(\log^\beta \Delta)$ -approximate non-signaling distribution requires locality  $\Omega(\log n / \log \Delta)$ , which combined with the KMW bound yields a degree-independent  $\Omega(\sqrt{\log n / \log \log n})$  quantum-LOCAL lower bound for  $O(\log^\beta \Delta)$ -approximation algorithms, removing the weak degree-dependent term from KMW's classic result.

The proof is based on two new low-soundness sensitivity lower bounds for label cover, one via Impagliazzo–Kabanets–Wigderson-style parallel repetition (SICOMP '12) with degree reduction and one from a sensitivity-preserving reworking of the Dinur–Harsha framework (SICOMP '13), together with the reductions from label cover to set cover to dominating set and the sensitivity-to-locality transfer theorem of Fleming and Yoshida (SODA '26).

---

\*Supported by the Swedish Research Council under grant number 2025-06762.

†Supported by NSF Grant Number DMS-2424441

‡Supported by JSPS KAKENHI Grant Number JP24K02903.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Proof Overview . . . . .	3
1.2	Further Related Work . . . . .	10
1.3	Organization of the Paper . . . . .	11
<b>2</b>	<b>Preliminaries</b>	<b>11</b>
2.1	Label Cover . . . . .	11
2.2	Sensitivity . . . . .	12
<b>I</b>	<b>Low-Soundness Label Cover via Impagliazzo–Kabanets–Wigderson</b>	<b>13</b>
<b>3</b>	<b>Left/Right Sensitivity Regions</b>	<b>13</b>
<b>4</b>	<b>Right-Degree Reduction</b>	<b>16</b>
<b>5</b>	<b>Low-Soundness Label Cover Route via Direct Product Testing</b>	<b>19</b>
5.1	The IKW Construction and Its Label-Cover Consequence . . . . .	19
5.2	Proof of the IKW Sensitivity Theorem . . . . .	22
5.3	Analysis of the Localized Decoder . . . . .	26
<b>II</b>	<b>Low-Soundness Label Cover via Dinur–Harsha</b>	<b>28</b>
<b>6</b>	<b>Sensitivity Reduction with Source-Dependent Recovery</b>	<b>29</b>
<b>7</b>	<b>Right-Degree Reduction in the Scalar Setting</b>	<b>30</b>
<b>8</b>	<b>Label Cover under Clause Swaps and Stable Recovery</b>	<b>31</b>
8.1	Left-Predicate Label Cover under Clause Swaps . . . . .	31
8.2	Recovering Satisfying Assignments under Local Perturbations . . . . .	42
<b>9</b>	<b>Alphabet Reduction</b>	<b>47</b>
9.1	The Alphabet Reduction Gadget . . . . .	47
9.2	Reduced Label Cover . . . . .	50
<b>10</b>	<b>Decodable PCPs and Composition</b>	<b>52</b>
10.1	Decodable PCPs . . . . .	53
10.2	Composition . . . . .	54
<b>11</b>	<b>Iteration by Stages</b>	<b>60</b>
11.1	How the Constants Evolve . . . . .	61
11.2	The Fixed Alphabet/Degree Reduction on a Common Family . . . . .	65
11.3	Initialization of the Left-Predicate Label Cover under Clause Swaps . . . . .	68
11.4	One-Stage Update of the Reduction Data . . . . .	75
11.5	Proof of Theorem 1.6 . . . . .	81
<b>III</b>	<b>From Label Cover to Covering Problems</b>	<b>84</b>
<b>12</b>	<b>Set Cover</b>	<b>84</b>
12.1	Reduction from Label Cover to Set Cover . . . . .	85
12.2	IKW-Based Bound . . . . .	87
12.3	Dinur–Harsha-Based Bound . . . . .	94

<b>13 Dominating Set</b>	<b>100</b>
13.1 Reduction from Set Cover to Dominating Set . . . . .	100
13.2 IKW-Based $O(\log \Delta)$ Bound . . . . .	103
13.3 Dinur–Harsha-Based Polylogarithmic-in-Degree Bound . . . . .	106

# 1 Introduction

Minimum dominating set is a classic local covering problem asking for the smallest set of vertices in a graph such that every vertex is either selected, or has a selected neighbor. The problem is a core optimization task in distributed computing, especially in the LOCAL model, where the trade-off between approximation quality and locality has been studied extensively [10, 13, 23, 30, 40, 42, 43], yet remains far from fully understood.

State of the art bounds for dominating set in the LOCAL model come in two flavors. On the one hand, Chang and Li [10] show a strong  $\Omega(\log n)$  locality lower bound against some constant-factor approximation. On the other hand, allowing a *degree-dependent* approximation factor, in seminal work Kuhn–Moscibroda–Wattenhofer [42] prove this lower bound can be circumvented, achieving a  $(1 + \varepsilon) \log \Delta$ -approximation with locality  $O(\frac{\log n}{\log(1+\varepsilon)})$  for any  $\varepsilon > 0$ , and complement this with a corresponding degree-dependent lower bound showing any  $\log \Delta$ -approximation requires locality at least

$$\Omega\left(\min\left\{\frac{\log \Delta}{\log \log \Delta}, \sqrt{\frac{\log n}{\log \log n}}\right\}\right).$$

Unfortunately, especially for smaller  $\Delta$ , this latter bound is extremely weak, scaling just with  $\log \Delta$ , and *never* achieves the  $\Omega(\log n)$  scale of Chang and Li. This has left open two central problems:

1. Can Chang and Li’s  $\Omega(\log n)$  lower bound be extended to  $O(\log \Delta)$ -approximation to match existing algorithms?
2. Do such bounds extend beyond the LOCAL model to modern distributed settings such as quantum-LOCAL and bounded-dependence? [2, 12, 29, 34]

In particular, while the KMW bound is known to extend to quantum-LOCAL and beyond [8], Chang and Li’s is not, leaving an  $\Omega(\log n)$  bound open even for constant factor approximation.

In this work, we take a major step toward resolving these problems. We prove two new lower bounds for minimum dominating set in the non-signaling model, where an algorithm has locality  $t$  if its output distribution has the property that the marginals on any vertex set depend only on the radius- $t$  neighborhood of that set; this model strictly subsumes the LOCAL, quantum-LOCAL, and bounded-dependence models at the same locality [2, 12, 29, 34, 45]. Our first main result extends Chang and Li’s  $\Omega(\log n)$  lower bound all the way to  $O(\log \Delta)$ -approximation against non-signaling, albeit with locality that decays with larger  $\Delta$ :

**Theorem 1.1.** There exists an integer  $\Delta_0 \geq 2$  and an absolute constant  $C > 0$  such that for all sufficiently large  $n$  and every integer  $\Delta \geq \Delta_0$ , any  $C \log \Delta$ -approximate non-signaling distribution for the dominating set problem on  $n$ -vertex graphs with maximum degree at most  $\Delta$  requires locality at least

$$\Omega\left(\frac{\log n}{\log \Delta} \cdot \frac{1}{\log \log^3 \Delta}\right)$$

Since Theorem 1.1 is proved for non-signaling, it also applies to LOCAL. Thus for every fixed constant  $\Delta$  with  $\Delta \geq \Delta_0$ , we get an  $\Omega(\log n)$  lower bound matching the scale of Chang–Li while upgrading the approximation factor from some fixed constant to  $O(\log \Delta)$ , essentially matching KMW’s LOCAL algorithm.

Theorem 1.1 is weaker for large  $\Delta$ , but can be improved in combination with KMW’s bound to achieve a *degree-unconditional* lower bound of  $\tilde{\Omega}(\sqrt{\log n})$  against  $O(\log \Delta)$ -approximation in quantum-LOCAL. In fact, if one could remove the poly  $\log \log \Delta$  term from Theorem 1.1, the KMW bound becomes stronger

only when  $\Delta \geq 2^{\Theta(\sqrt{\log n \log \log n})}$ , *exactly* the point at which the classic bound's second term dominates. In other words, removing the poly  $\log \log \Delta$  term in Theorem 1.1, we could *strictly* upgrade KMW's bound to

$$\Omega\left(\sqrt{\frac{\log n}{\log \log n}}\right)$$

against  $O(\log \Delta)$ -approximation, removing the degree-based term entirely. Our second main result shows this is indeed possible if one is willing to settle for a slightly weaker approximation factor:

**Theorem 1.2.** There exists  $\beta \in (0, 1)$  and an integer  $\Delta_0 \geq 2$  such that for all sufficiently large  $n$  and every integer  $\Delta \geq \Delta_0$ , any  $O(\log^\beta \Delta)$ -approximate non-signaling distribution for the dominating set problem on  $n$ -vertex graphs with maximum degree at most  $\Delta$  requires locality at least

$$\Omega\left(\frac{\log n}{\log \Delta}\right)$$

This implies the discussed corollary for quantum-LOCAL, removing KMW's degree-dependent term:

**Corollary 1.3.** There exists  $\beta \in (0, 1)$  and an integer  $\Delta_0 \geq 2$  such that for all sufficiently large  $n$  and every integer  $\Delta \geq \Delta_0$ , any quantum-LOCAL algorithm  $O(\log^\beta \Delta)$ -approximating dominating set on  $n$ -vertex graphs with maximum degree at most  $\Delta$  requires locality at least

$$\Omega\left(\sqrt{\frac{\log n}{\log \log n}}\right)$$

Table 1 summarizes the resulting bounds.

Table 1: Trade-offs between approximation ratio and number of rounds by model. Here  $n$  is the number of vertices and  $\Delta$  is the maximum degree.

Approximation ratio	Rounds	Model	Reference
Lower bounds			
Fixed constant $c > 1$	$\Omega(\log n)$	LOCAL	[10]
$O(\log \Delta)$	$\Omega\left(\log n \cdot \frac{1}{\log \Delta \text{ poly } \log \log \Delta}\right)$	Non-signaling	Theorem 1.1
$\log^\beta \Delta$ for some $\beta > 0$	$\Omega\left(\log n \cdot \frac{1}{\log \Delta}\right)$	Non-signaling	Theorem 1.2
$\log^\beta \Delta$ for any $\beta > 0$	$\Omega\left(\min\left\{\frac{\log \Delta}{\beta \log \log \Delta}, \sqrt{\frac{\log n}{\beta \log \log n}}\right\}\right)$	quantum-LOCAL	[13, 42]
$\log^\beta \Delta$ for some $\beta > 0$	$\Omega\left(\sqrt{\frac{\log n}{\log \log n}}\right)$	quantum-LOCAL	Theorem 1.2 + [42]
Upper bounds			
$1 + \varepsilon$ for any $\varepsilon > 0$	$\text{poly}(\log n / \log(1 + \varepsilon))$	LOCAL	[30]
$(1 + \varepsilon) \log \Delta$ for any $\varepsilon > 0$	$O(\log n / \log(1 + \varepsilon))$	LOCAL	[42]

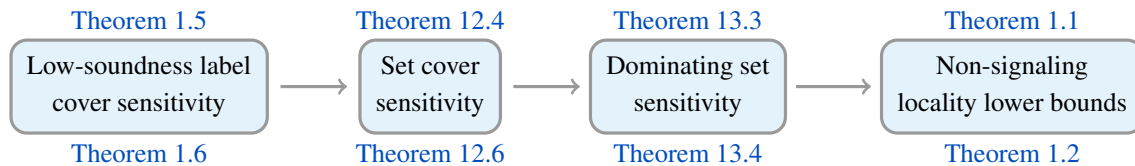
We note that the KMW bound in fact holds for even stronger approximation ratios beyond  $\log \Delta$ , in particular for any  $\beta > 0$  one gets a lower bound of  $\Omega\left(\min\left\{\frac{\log \Delta}{\beta \log \log \Delta}, \sqrt{\frac{\log n}{\beta \log \log n}}\right\}\right)$  against  $\log^\beta \Delta$ -approximation. We leave as an open problem whether our results can be extended to this regime.

## 1.1 Proof Overview

**Transfer principle.** Our lower bounds are proved via a reduction to *algorithmic sensitivity*, which measures how much an algorithm’s output distribution can change after a local perturbation of the input, e.g., a single changed constraint in a constraint satisfaction problem (CSP). In particular, we use a sensitivity-to-locality transfer principle from [28] which converts sensitivity lower bounds for graph problems into lower bounds for non-signaling algorithms. This transfer principle is simply the observation that, for any graph problem on graphs of degree at most  $\Delta$ , any non-signaling algorithm with locality  $t$ , when run on two input instances perturbed in one location, must have the same output outside a radius- $t$  ball surrounding the change. This immediately implies any such algorithm has sensitivity at most  $\Delta^t$ , so if one can show that no, say,  $\text{poly}(n)$ -sensitive algorithm exists for graphs of size  $n$  and degree  $\Delta$ , then this implies a lower bound against any locality  $t = \log_{\Delta}(n)$  non-signaling algorithm.

We defer the formal definitions of sensitivity to the preliminaries and here only overview our methods.

**Reduction chain.** Theorem 1.1 and Theorem 1.2 are proved via the following chain of reductions:



The first arrow is the classical reduction from label cover (see Section 1.1.1 for an overview of the label cover problem) to set cover, the second is the standard reduction from set cover to dominating set, and the last is the sensitivity-to-locality transfer theorem. Together they yield set-cover and dominating-set sensitivity lower bounds and, through the transfer principle of [28], the locality theorems above. For the Dinur–Harsha route, these steps are packaged as the black-box chain shown in the lower row. For the IKW route, the upper row is now the direct chain

$$\textit{Theorem 1.5} \rightarrow \textit{Theorem 12.4} \rightarrow \textit{Theorem 13.3} \rightarrow \textit{Theorem 1.1}.$$

Thus Theorem 1.5 supplies the low-soundness label-cover hardness, Theorem 12.4 turns it into a set-cover lower bound, Theorem 13.3 lifts this to bounded-degree dominating set, and the transfer principle yields Theorem 1.1.

**Key Ingredients.** Existing label cover sensitivity lower bounds [28] only handle the high-soundness regime, roughly a  $(1 - \varepsilon)$ -approximation on satisfiable instances for a fixed constant  $\varepsilon > 0$ . This is far too weak for the reductions to set cover and dominating set, which require low-soundness label cover instances of the kind used in classical PCP hardness reductions [27, 33]. Unfortunately, low soundness PCPs are famously harder to construct and analyze than their high soundness counterpart. The following two theorems, our main technical contribution, supply sensitivity lower bounds for such label cover instances, which we use to prove Theorem 1.1 and Theorem 1.2 respectively.

Before stating the results, we briefly note that both results below are stated for *left-predicate* label cover.

**Definition 1.4 (Label Cover).** Recall that a 2-CSP consists of an underlying graph  $G = (V, E)$ , an alphabet  $\{\Sigma_v\}$  of possible values for each vertex, and, for each edge  $e = (v, v')$ , a constraint  $f_e$  defining a set of satisfying assignments in  $\Sigma_v \times \Sigma_{v'}$ . The value of the instance is the maximum over assignments  $\pi$  to  $V$

of the total fraction of satisfied constraints. A *label cover* instance is a special 2-CSP consisting of a tuple  $(U, V, E, \Sigma_U, \Sigma_V, F)$  where  $G = (U, V, E)$  is a bipartite graph,  $\Sigma_U$  is a fixed alphabet for the left-hand side,  $\Sigma_V$  is a fixed alphabet for the right-hand side, and  $F = \{f_e : \Sigma_U \rightarrow \Sigma_V\}$  (the ‘edge projections’) determines the constraints by requiring  $\pi(v) = f_e(\pi(u))$  for any edge  $e = (u, v)$ .

The left-predicate label cover problem is a slight generalization of the above in which each left vertex  $u$  carries a unary predicate; if this predicate evaluates to 0, then all projection constraints incident to  $u$  are violated. Ordinary label cover is recovered by taking all left predicates to be identically 1. We use this language because, in the reductions used later, a local change in the instance before the reduction can be encoded by changing only the left-side predicates in the resulting label cover instance, while keeping the underlying graph and edge projections fixed; this is exactly the form needed for the sensitivity argument. For this overview, we suggest that the reader think of this simply as label cover.

**Theorem 1.5.** There exists  $C > 0$  such that for all sufficiently large  $N \in \mathbb{N}$  and  $k \leq o(\log N)$ , any algorithm that, on satisfiable left-predicate label cover instances on  $n = N^{\Theta(k)}$  vertices, outputs an  $\exp(-C\sqrt{k})$ -satisfying solution must have sensitivity at least  $n^{\Omega(1/k)}$ . Moreover, the hard instances can be taken to be biregular with alphabet and degree parameters at most  $2^{O(k)}$ .

Roughly speaking, to prove lower bounds for max-degree  $\Delta$  dominating set, one should think of setting  $k = \text{poly log log } \Delta$  above, resulting in an  $n^{1/\text{poly log log } \Delta}$ -sensitivity lower bound. Traced through the reduction, the exponent becomes the additional  $\text{poly log log } \Delta$ -term in Theorem 1.1 that we’d like to remove. With this in mind, we give the following alternate truly polynomial sensitivity lower bound:

**Theorem 1.6.** There exist absolute constants  $B, \delta > 0$  such that for every function  $1 < g(n) \leq (\log n)^B$  and all sufficiently large  $n$ , any algorithm that, given a satisfiable left-predicate label cover instance with  $n = |U|$ -many left vertices, outputs an assignment that satisfies at least a  $1/g(n)$  fraction of the constraints must have sensitivity at least  $\Omega(n^\delta)$ . Moreover, the hard instances can be chosen so that  $|U| = n$ ,  $|V| = n \cdot (\log n)^{O(1)}$ , the left and right degrees are  $g(n)^{O(1)}$ ,  $|\Sigma_V| = g(n)^{O(1)}$ , and  $|\Sigma_U| \leq \exp(g(n)^{O(1)})$ .

Unfortunately, Theorem 1.6 cannot be used to prove a lower bound against  $O(\log \Delta)$ -approximation for dominating set due to the large left-alphabet size, which restricts us to  $O(\log^\beta \Delta)$ -approximation for some  $\beta > 0$ . More generally, the degree  $\Delta$  of the final dominating set instance in our reduction is at least the degree and alphabet size of the starting low soundness label cover, so it is critical we keep these parameters as small as possible. Improving the alphabet-size bound in the PCP underlying Theorem 1.6 is a central open problem in PCP theory, related to the so-called ‘sliding scale conjecture’, and doing so would have consequences far beyond just improving our dominating set lower bound [48].

Below, we give an overview of the proofs of Theorem 1.5 and Theorem 1.6. Before this, we describe the challenges that go into developing a PCP theorem that preserves sensitivity.

**The Challenge of Maintaining Sensitivity.** Classical PCP theorems for hardness of approximation track how the completeness, soundness, and structural parameters of a label cover instance change as it proceeds through the PCP. For sensitivity, this is not enough. Because sensitivity compares the behavior of an algorithm on two *nearby* source instances rather than a single instance in isolation, we need to ensure the reduction remains stable applied to instances related by any one local modification, which we call a *source swap*. In other words, instead of working with single instance, formally we will have to work with and control the behavior of an entire *swap-closed family*<sup>1</sup> of instances throughout the reduction.

<sup>1</sup>That is, once an instance is in the family, every instance obtained from it by one source swap also remains in the family; see Section 2 for the formal definition.

Recall our goal is really to transfer a sensitivity lower bound backward through the PCP reduction, which we call *sensitivity pullback*. For this, each stage of the PCP reduction must come with a recovery map that converts a labeling of the transformed instance  $I_{i+1}$  into a labeling of the source instance  $I_i$ . Once such a recovery map is available, there are three effects of a source swap that must be controlled:

- how much the transformed instance itself can change;
- how much the recovery map amplifies a perturbation of a labeling on the transformed instance; and
- how much the recovery map itself changes when the source instance changes.

This is a substantially stronger requirement than the usual completeness/soundness bookkeeping required by an ordinary PCP theorem. To make it possible, we design our reductions so that all instances in the family live on a common underlying graph, and a source swap changes only local predicates, not the graph or the edge projections.

### 1.1.1 Proof Overview of Theorem 1.5

The proof of Theorem 1.5 consists of two main steps: 1) a new “sensitivity-preserving” analysis and recovery map for (combinatorial) parallel repetition [38], which amplifies soundness (approximation factor) of a base label-cover instance at the cost of massively blowing up its size and (right)-degree, and 2) a new “sensitivity-amplifying” degree-reduction procedure that simultaneously reduces this right-degree while improving the right-sensitivity.<sup>2</sup> We overview these two steps below along with some corresponding background.

**Combinatorial Parallel Repetition.** The low-soundness combinatorial parallel repetition theorem of Impagliazzo, Kabanets, and Wigderson (IKW) [38] amplifies a base 2-CSP  $\Phi$  with alphabet  $\Sigma$  on underlying graph  $G_\Phi = (V, E)$  to a label-cover instance  $\Phi' = (U', V', E', \Sigma_{U'}, \Sigma_{V'}, F')$  defined as follows.

The underlying graph of  $\Phi'$  has left vertex set  $U' = \binom{V}{k}$  and right vertex set  $V' = \binom{V}{k'}$ . Edges in  $E'$  are generated via the following process:

1. Sample a size  $k'$  subset  $A \subset [V]$  (i.e. a random right-vertex in  $V'$ )
2. Sample a random edge for each vertex in  $A$ , and call the resulting edge-set  $E_A$
3. Sample  $k - k'$  additional random edges  $E_B \subseteq E$  without replacement.
4. Output the edge  $(\{E_A, E_B\}, A)$

In particular, every sampled pair  $\{E_A, E_B\}$  determines a left vertex  $S = \{E_A, E_B\} \in U'$ , while the right endpoint  $A$  lies in  $V' = \binom{V}{k'}$ . The left alphabet is  $\Sigma_{U'} = \Sigma^{2k}$  (one symbol for each of the  $2k$  original vertices in  $\{E_A, E_B\}$ ), and the right alphabet is  $\Sigma_{V'} = \Sigma^{k'}$  (one symbol per original vertex in  $A$ ). For an assignment  $(\pi_{U'}, \pi_{V'})$  to satisfy the edge  $(\{E_A, E_B\}, A) \in E_{IKW}$ , it must hold that<sup>3</sup>

1. The left label  $\pi_U(E_A, E_B)$  satisfies all constraints in  $E_B$  in the original CSP, and
2.  $\pi_U$ 's vertex assignments on  $A$  agree with  $\pi_V(A)$ .

We would like to show that if the base-CSP  $\Phi$  on  $N$  vertices has no (say)  $N^{1/2}$ -sensitive algorithm that outputs a  $(1 - \delta)$ -satisfying solution on satisfiable inputs, then  $\Phi'$  likewise has no  $N^{1/4} = n^{\Omega(1/k)}$ -sensitive algorithm that, on satisfiable inputs, outputs even an  $\varepsilon = \exp(-O(k'))$ -satisfying solution. This can then be instantiated with a base CSP of this form [28].

<sup>2</sup>Here, ‘right’-sensitivity corresponds to the number of changes seen on the right-hand side of the label cover due to a single local perturbation in the input instance.

<sup>3</sup>We note this is technically a left-predicate label cover instance, with the first condition below being the ‘left-predicate’. See the formal definitions in Section 2.

**Imbalanced Graphs and Sensitivity.** Unfortunately, we will not be able to show such a strong sensitivity lower bound directly due to the fact that the label-cover instance described above is extremely imbalanced. To illustrate the problem, let us first outline what the above reduction should look like.

As described at a high level above, our basic approach is to start with neighboring instances  $I$  and  $\tilde{I}$  of the original 2-CSP that differ in one constraint, and map both instances to their corresponding IKW label cover instances. To prove a lower bound against  $T$ -sensitive algorithms for the resulting label cover, we'll assume toward contradiction we in fact have such an algorithm. Our goal is to run this algorithm on the embeddings of  $I$  and  $\tilde{I}$  to produce solutions  $\pi$  and  $\tilde{\pi}$ , then give a 'recovery map' that maps these solutions back to highly satisfying solutions to the base CSP  $R(\pi)$  and  $R(\tilde{\pi})$  whose Hamming distance is  $\ll \sqrt{N}$  in expectation, violating the sensitivity lower bound for the base CSP.

What actually happens when we run this approach? Well, if  $I$  and  $\tilde{I}$  differ by one constraint, the corresponding label cover instances differ in roughly a  $\Theta(k/N)$  fraction of constraints, the probability we sample a modified constraint in steps (1-3) above. Running a  $T$ -sensitive algorithm on these embedded neighboring instances would therefore produce solutions  $\pi$  and  $\tilde{\pi}$  that differ in up to  $T|E'| \frac{k}{N}$  locations. This is a problem: due to the imbalance of the graph, even if  $T = 1$ ,  $T|E'| \frac{k}{N} \gg |V'|$ . In other words, the entire right-hand side of the two embedded label cover instances could be completely different!

We'll handle this by proving a weaker lower bound which separately controls the sensitivity on the left and right parts of the label-cover instance. In particular, we will show there is no algorithm which is  $N^{1/4}$ -sensitive on the left-hand side, and  $\frac{|E'|}{|V'|} N^{1/4}$ -sensitive on the right-hand side, normalizing for this degree imbalance. Note that this latter quantity is much less than 1, which seems like an extremely weak lower bound. Nevertheless, we will see this guarantee, which controls how much the right-hand side changes *in expectation*, is in fact still strong enough to achieve the final desired  $\text{poly}(N)$ -bound after degree reduction, which roughly amplifies this term by a corresponding  $\frac{|V'|}{|E'|}$  factor.

**The Recovery Procedure.** With this in mind, assuming the embedded instances have a  $(N^{1/4}, \frac{|E'|}{|V'|} N^{1/4})$ -sensitive algorithm, running the above approach indeed results in assignments  $\pi$  and  $\tilde{\pi}$  that agree with each other on both sides of the label cover except a  $o(1/\sqrt{N})$  fraction of the time. We therefore need to give a recovery map which, given such labelings, returns highly-satisfying  $\sqrt{N}$ -close solutions to the base CSP. We describe such a recovery procedure below based on an assignment  $\pi = (\pi_{U'}, \pi_{V'})$ :

1. Draw  $A \in \binom{V}{k'}$  uniformly and set  $a = \pi_{V'}(A) \in \Sigma^{k'}$
2. For each  $x \in V$ :
  - Draw a uniformly random edge  $e \in E$  of the original CSP with  $x \in e$
  - Draw  $O(\frac{\log N}{\epsilon})$  independent  $S_E = (E_A, E_B) \in U'$  uniformly conditioned on  $e \subseteq S_E$
  - If  $\pi_{U'}(S_E)|_A = a$  for any sampled  $S_E$ , output  $\pi_{U'}(S_E)|_x$ <sup>4</sup>
  - Else: Output a fixed symbol in  $\Sigma$ .

We first note that, essentially by design, the above procedure has low-sensitivity. Namely, run on  $\pi$  and  $\tilde{\pi}$ , the decoding only differs if we draw some  $A$  or  $S_E$  on which  $\pi \neq \tilde{\pi}$ . However, since, over the randomness of  $x$ ,  $A$  and  $S_E$  are marginally uniform, the probability we see such a disagreement scales with the normalized Hamming distance of  $\pi$  and  $\tilde{\pi}$ , and in particular is  $o(1/\sqrt{N})$ , so union bounding over these events the expected (un-normalized) distance between the final decodings remains  $o(\sqrt{N})$  as desired.

The main technical component of the argument is then to show the above recovery procedure actually outputs a, say,  $(1 - \delta/2)$ -satisfying solution to the original CSP  $\Phi$  with non-trivial probability. Once we

---

<sup>4</sup>If there are multiple such  $S_E$ , pick one at random.

have this, it is easy to produce a good decoding using an additional “low-sensitivity selection” procedure inspired by similar algorithms in replicable learning [39]:

1. Repeat the recovery procedure several times independently
2. Pick a random  $t \in [1/2, 1]$
3. Output the first solution with value at least  $1 - t\delta$

Repeating the procedure independently ensures at least one good output with high probability. Since the recovered solutions from  $\pi$  and  $\tilde{\pi}$  are very close, their values will be close, and as a result the randomized thresholding procedure in Steps 2/3 will almost certainly select close solutions to the base CSP as desired.

It remains to argue the initial recovery procedure indeed produces a high value solution to the base CSP. This is the most technical part of the argument, and is deferred to Section 5. The high level idea is to show that the decoding procedure is typically close to the *majority* decoding value of  $x$  conditioned on  $\pi_{V'}(A) = a$ . This majority-based decoding is already known to have high value due to the PCP reduction of [38], but may have poor sensitivity. The proof of closeness to majority follows via concentration-of-measure properties of the graph underlying the IKW label-cover, which control the probability that the label of a random  $x \in V$  is far from its expectation, even after conditioning on various events like  $\pi_{V'}(A) = a$ .

**Degree Reduction** The produced label-cover instance above has poor right-sensitivity and corresponding high right degree; the second key step is to give a degree-reduction procedure rectifying this. Toward this end, we use a procedure from Dinur and Harsha [20] that reduces the right degree to  $\exp(O(k'))$  (polynomial in the soundness), and, critically, give a recovery map which actually *amplifies* the corresponding right-sensitivity lower bound by a factor of  $\frac{|E'|}{|V'|}$ , the original right-degree. Plugging this into the sensitivity bound above gives the desired  $\text{poly}(N)$ -sensitivity lower bound for low-soundness label cover of Theorem 1.5.

In slightly more detail, degree-reduction proceeds by replacing each right vertex  $v$ , with degree  $D_v$ , by a cloud of  $D_v$ -many vertices. An expander is then used to wire these new vertices to the old neighborhood of  $v$ . If  $u$  was a neighbor of  $v$  and it is now a neighbor of a vertex  $v'$  in the cloud, then the edge  $(u, v')$  inherits the constraint of  $(u, v)$ . The expander ensures that the constraints of  $v$ 's edges are dispersed uniformly among the cloud, making them a good approximation to the original constraint-set. Critically this introduces some extra redundancy: the assignment for  $v$  is duplicated among the vertices in its cloud. We design a recovery procedure which takes advantage of this redundancy to increase the sensitivity lower bound by a corresponding factor in the size of the cloud (the original right-degree) as desired.

### 1.1.2 Proof Overview of Theorem 1.6

The proof of Theorem 1.6 has two main steps. First, we construct an initial label-cover family, which we call the *stage-0* construction, with very small soundness but very large alphabet, degree, and local description complexity, an additional measure needed for composition roughly capturing the complexity of computing whether a given assignment to the neighborhood of some left vertex  $u$  has a potential satisfying label for  $u$ . Second, we apply a Dinur–Harsha style iterated composition theorem [20] to move to the target soundness scale  $1/g(n)$  while reducing these structural parameters to the corresponding scale.

As we adapt the Dinur–Harsha framework, we recall only the part relevant below. Stage-0 already gives a soundness threshold much smaller than the one needed in the final theorem, but its alphabet size, degrees, and local description complexity are tuned to that overly strong regime. The iteration weakens the soundness guarantee in a controlled way and, at the same time, reshapes the local structure so that these parameters can be reduced to the new scale.

**Dinur–Harsha Composition.** A label cover instance asks, at each left vertex  $u$ , whether the labels on the neighboring right vertices form an admissible tuple, namely one that can be extended to some left label satisfying the incident projections and the left predicate. Composition replaces this admissibility check by a local proof-verification task: one supplies a short auxiliary proof together with a local decoder that probes only a few proof positions and either recovers the relevant component of an admissible tuple or rejects at the cost of increasing the soundness threshold from  $\varepsilon_i$  to the weaker scale  $\varepsilon_{i+1}$ .

After composition, the resulting instance has a simpler local certification, but its structural parameters are still not in the form needed for the next round. The cleanup steps—alphabet reduction and degree reduction—convert this composed instance back into a left-predicate label cover whose degree, alphabet, and local-description parameters are polynomial in  $1/\varepsilon_{i+1}$ . Thus, one stage has the form

$$I_i \xrightarrow{\text{compose}} J_{i+1} \xrightarrow{\text{alphabet/degree reduction}} I_{i+1}.$$

The intermediate instance  $J_{i+1}$  exposes the proof-based certification created by composition, while the cleanup step restores the interface needed for the next stage. Repeating these two steps moves the construction from the very small soundness of stage-0 to the target scale  $1/g(n)$  while keeping the structural parameters under control. The rest of the overview explains how we stabilize this template for sensitivity.

**Stage-0: Keeping Clause-Swaps Local.** The first obstacle already appears in the initialization  $I_0$  of the iteration. The hard family that we ultimately care about starts as bounded-degree satisfiable 2-CSP instances (we take the instance constructed in the lower bound proof of Fleming and Yoshida [28]), but before the Dinur–Harsha iteration begins, we must first convert that family—via a constant-size gadget—into a bounded-degree satisfiable CNF family; the instance  $I_0$  for stage-0 is the robust algebraic encoding of this CNF family into a left-predicate label cover instance. In the standard Dinur–Harsha template, the formula is represented by its clause-indicator table. It is then extended to a global low-degree object, and then checked by the usual algebraic consistency tests, which verify its degree and that it does indeed represent the clause-indicator table. This is excellent for classical soundness, because the low-degree structure gives strong global control. For sensitivity, however, it creates exactly the wrong kind of dependence on the source formula: swapping one clause for another (a *clause swap*) changes one entry of the clause table before algebraization, but after low-degree extension the encoded object typically changes at many points, so one local clause swap becomes a global change in the label-cover instance.

Our new stage-0 construction breaks this dependence into two separate tasks. The outer algebraic layer certifies only the low-degree structure of the assignment and auxiliary witness polynomials. The question “what is the value of the clause table at this particular location?” is delegated to a small Reed–Muller decode/reject gadget that either returns the requested table value or rejects. Because the CNF now enters only through this local lookup, the formula-dependent part of the construction is confined to the left predicates. Consequently, one clause swap changes only a controlled number of left predicates, while the underlying graph and edge projections remain fixed. This yields the left-predicate label cover under clause swaps that serves as the stage-0 starting family for the rest of the argument.

Finally, stage-0 must also provide a stable way to recover a satisfying assignment to the initial CNF formula for sensitivity pullback. Given an assignment to the stage-0 label cover instance, we extract a tuple of candidate low-degree polynomials which claim to encode the low-degree extension (as well as some auxiliary witnesses), and score them by the fraction of local checks they explain. We then select one of the high-scoring candidates using a random threshold. This produces a satisfying assignment with noticeable probability, in a form that remains stable under local perturbations for the later pullback argument.

**Stable Composition: Padding the Graph.** Next, we modify the composition procedure from Dinur-Harsha [20] to preserve sensitivity. After composition, the constraints on the edges incident to a left vertex  $u$  check whether the joint labeling of the right vertices that are adjacent to  $u$  are compatible with  $u$ 's label and all incident projections. The inner decoder is asked to certify exactly this local compatibility condition, rather than to reason about the full left alphabet directly.

In analyzing sensitivity, this standard approach once again hits a barrier: in a naive composed instance, the decoder's chosen proof locations determine which new edges appear in the composed graph. A local source change can then alter which locations are queried, and hence alter the graph of the composed instance itself. To prevent this, we implement the composition on a single padded graph shared by all instances in the family. Each composed left label is viewed as a fixed-size matrix whose rows correspond to the potential neighboring constraints of the underlying outer right vertex and whose columns correspond to positions in the inner decoder's proof string. The rows corresponding to actual neighbors are active, and the remaining rows are dummy padding. All coordinate-projection edges from each active row to every column are placed in advance. The decoder's actual queries are then used only inside the left predicate rather than being encoded by the edge set. With this padding, a source swap may change which rows and columns are semantically relevant, but it does not change the underlying graph.

**Stable Cleanup: Freezing the Auxiliary Choices.** After composition, one still has to repair the local parameters before the next round. This is the role of the alphabet-then-degree reduction. In a classical hardness proof these gadgets are used only as parameter-improving transformations, so one is free to choose whatever code or expander is convenient at that moment. In the sensitivity setting, that freedom is dangerous: if nearby source instances were sent through different auxiliary choices, then the reduction itself could vary globally even when the source change was local.

We therefore freeze these auxiliary choices on each swap-closed family. In particular, the code used in alphabet reduction and the expanders used in the subsequent degree-reduction step are fixed once-and-for-all on that family. This matters because the alphabet-reduction recovery samples each recovered right label from the empirical distribution induced by the incident transformed edges at that vertex, and that distribution depends not only on the transformed labeling but also on the source instance through the underlying projections. Hence "stability under perturbing a labeling on a fixed transformed instance" and "stability under changing the source instance" are different statements. Freezing the auxiliary choices keeps the reduction itself fixed, so that only the predicates and the resulting recovery maps vary across the family.

**How Iterations Fit Together.** With these modifications in place, the rest of the proof can be read as a stabilized version of Dinur-Harsha iteration. We start from the stage-0 family described above, and that stage-0 construction begins with a smaller soundness threshold than the one ultimately required in the theorem. At stage  $i$ , composition weakens the soundness guarantee from  $\varepsilon_i$  to  $\varepsilon_{i+1} = 20\varepsilon_i^{1/4}$ . The cleanup step then restores the degree, alphabet, and local-description bounds needed to run the next round at the new scale. As the starting point for the iteration, we use the high-soundness hard family of Fleming and Yoshida [28], namely a swap-closed family of satisfiable bounded-degree 2-CSP instances with constant alphabet and degree. A constant-size gadget first converts this family into the bounded-degree satisfiable CNF family used by the stage-0 construction above.

At stage  $i$  we maintain a family  $I_i$  on one common underlying graph, together with an instance map from that base family and a recovery map back to it, and we track the three stability quantities above. The stage parameter is denoted  $\varepsilon_i$ : smaller  $\varepsilon_i$  means stronger soundness, while the left and right degrees, alphabet sizes, local description complexity, and vertex blow-up at stage  $i$  are polynomial in  $1/\varepsilon_i$ . One stage has the

form

$$I_i \xrightarrow{\text{compose}} J_{i+1} \xrightarrow{\text{alphabet/degree reduction}} I_{i+1},$$

where the composition step updates the parameter from  $\varepsilon_i$  to  $\varepsilon_{i+1} = 20\varepsilon_i^{1/4}$ , and the cleanup step restores the same polynomial-in- $1/\varepsilon_{i+1}$  bounds needed for stage  $i + 1$ . Each layer loses only a polynomial factor in  $1/\varepsilon_i$  or  $1/\varepsilon_{i+1}$ , and the recursion  $\varepsilon_{i+1} = 20\varepsilon_i^{1/4}$  keeps the total product of these losses polylogarithmic. Consequently, the iteration reaches the target  $1/g(n)$  soundness threshold while still preserving a polynomial sensitivity lower bound.

## 1.2 Further Related Work

**Dominating Set, Label Cover, and the Non-Signaling Model** The non-signaling model (also known as the  $\phi$ -LOCAL or causal model) has emerged as the gold standard for distributed computing lower bounds, automatically extending to the popular LOCAL, quantum-LOCAL, and bounded-dependence models [2, 4, 12, 29]. Fleming and Yoshida [28] first connected non-signaling locality to sensitivity, obtaining non-signaling lower bounds for problems such as minimum vertex cover and maximum cut, but their methods (based on high-soundness label cover) do not extend directly to other classic distributed problems like dominating set. For dominating set,  $\Omega(\log n)$  lower bounds were only known in the LOCAL model and only for constant-factor approximation. In particular, Chang and Li [10] prove an  $\Omega(\log n)$  lower bound for some constant approximation, while Kuhn, Moscibroda, and Wattenhofer [42] give a  $O(\frac{\log n}{\log(1+\varepsilon)})$ -round LOCAL algorithm for  $(1 + \varepsilon) \log \Delta$ -approximate dominating set along with a weaker degree-dependent  $\Omega(\min\{\log \Delta / \log \log \Delta, \sqrt{\log n / \log \log n}\})$ -round lower bound against  $\text{polylog}(\Delta)$ -approximation. The latter bound was also known to hold for quantum-LOCAL [8, 13, 42], but has remained open for the bounded-dependence and non-signaling settings.

Reductions from low soundness label cover to set cover and dominating set are a classical approach in hardness of approximation. Lund and Yannakakis [47] introduced the set-system framework underlying hardness for minimization problems, Feige [27] established the threshold hardness of set cover via reduction from label cover, and Chlebík and Chlebíková [11] proved bounded-degree dominating-set hardness. These are the reductions that underlie Sections 12 and 13; our use differs in that we reconstruct them in a form compatible with the later sensitivity pullback, rather than using them only to transfer approximation gaps.

**Sensitivity** The notion of sensitivity was introduced by Murai and Yoshida [50] to study the stability of network centrality measures. Subsequent work focused largely on problem-specific sensitivity analyses and low-sensitivity algorithms for broader graph problems [53, 55, 56]. The notion is also related to average sensitivity [26, 31, 44, 54], where the average is taken over deleted elements, as well as to differential privacy [24] and stability in learning [39]. We draw on some techniques from this literature.

On the lower-bound side, Fleming and Yoshida [28] established the first lower bounds on the sensitivity of approximation algorithms for label cover, but only in the high-soundness regime. Since our non-signaling lower bounds for dominating set proceed through the aforementioned classic reductions from label cover to set cover and dominating set, the missing ingredient was analogous lower bounds in the delicate low-soundness regime that are strong enough to survive those reductions; this paper supplies them.

**Low Soundness PCPs** Low-soundness label cover is a central problem in PCP-based hardness of approximation [33]. The relevant constructions in this regime include the low-soundness two-query PCP of Moshkovitz and Raz [49], the closely related PCP composition framework of Dinur and Harsha [20], parallel repetition and direct product testing [38, 52], and new derandomized direct product theorems [5, 6, 7,

15, 16, 17, 22, 51] based on high dimensional expanders. It is worth noting that latter give a promising approach towards a “best of both worlds” bound (i.e. combining the approximation factor of Theorem 1.1 and locality bound of Theorem 1.2), in part due to a new local recovery algorithm for related instances [19], but major barriers to this approach remain including 1) a ‘routing/embedding’ step in such reductions that ruins sensitivity, and 2) the poor alphabet-soundness trade-off achieved by current HDX-based constructions.

We emphasize that the contribution of this work is not a new PCP theorem; it is a method for proving sensitivity lower bounds that can exploit these often highly involved low-soundness label cover constructions while retaining the parameters needed for downstream applications to set cover and dominating set.

### 1.3 Organization of the Paper

Sections 3 to 5 develop the Impagliazzo–Kabanets–Wigderson route to low-soundness label cover and prove Theorem 1.5. Sections 6 to 11 develop the Dinur–Harsha route and prove Theorem 1.6. Sections 12 and 13 then organize the covering reductions by problem and prove Theorems 1.1 and 1.2.

## 2 Preliminaries

For a positive integer  $n$ , let  $[n]$  denote the set  $\{1, 2, \dots, n\}$ . We use bold symbols such as  $\pi$  to denote random assignments, random labelings, and other random output objects produced by algorithms or recovery maps; deterministic assignments and labelings are written without boldface, and symbols such as  $\Pi$  are reserved for probability distributions and couplings.

### 2.1 Label Cover

We work with a slightly generalized label cover formalism. A label cover instance is a bipartite projection game, and here we allow a unary left predicate on each left vertex. Ordinary label cover is recovered by taking all of these predicates to be identically 1.

**Definition 2.1** (Left-predicate label cover instance). A *left-predicate label cover instance* is a tuple

$$I = (U, V, E, \Sigma_U, \Sigma_V, \mathcal{P}, F), \quad \mathcal{P} = \{P_u : \Sigma_U \rightarrow \{0, 1\}\}_{u \in U}, \quad F = \{f_e : \Sigma_U \rightarrow \Sigma_V\}_{e \in E},$$

where  $G = (U, V, E)$  is a bipartite graph,  $\Sigma_U$  and  $\Sigma_V$  are finite alphabets, each  $P_u$  is a unary predicate on left labels, and each  $f_e$  is a projection. An assignment is a pair  $\pi = (\pi_U, \pi_V)$  with  $\pi_U : U \rightarrow \Sigma_U$  and  $\pi_V : V \rightarrow \Sigma_V$ . An edge  $e = (u, v)$  is satisfied by  $\pi$  if

$$P_u(\pi_U(u)) = 1 \quad \text{and} \quad f_{(u,v)}(\pi_U(u)) = \pi_V(v).$$

Thus a left label that fails its predicate satisfies none of its incident edges. For an assignment  $\pi$ , let

$$\text{val}_I(\pi) := \frac{1}{|E|} \cdot \#\{(u, v) \in E : P_u(\pi_U(u)) = 1 \text{ and } f_{(u,v)}(\pi_U(u)) = \pi_V(v)\}$$

denote the fraction of satisfied edges, and write  $\text{val}(I) := \max_{\pi} \text{val}_I(\pi)$ . When every predicate  $P_u$  is identically 1, we recover the usual label cover / projection-game notion. We will refer to this special case simply as *ordinary label cover*. When the predicates are clear from context, we may abbreviate the tuple by omitting  $\mathcal{P}$ . For  $\delta \in (0, 1)$ ,  $\text{LabelCover}_{\delta}$  is the problem that, given a satisfiable instance, asks for an assignment of value at least  $\delta$ .

For two assignments  $\pi, \tilde{\pi} : X \rightarrow \Sigma$  over the same domain  $X$ , let  $d_H(\pi, \tilde{\pi}) = \#\{x \in X : \pi(x) \neq \tilde{\pi}(x)\}$  denote their Hamming distance. For two assignments  $\pi = (\pi_U, \pi_V)$  and  $\tilde{\pi} = (\tilde{\pi}_U, \tilde{\pi}_V)$  for a left-predicate label cover instance  $I = (U, V, E, \Sigma_U, \Sigma_V, \mathcal{P}, F)$ , we define  $d_H(\pi, \tilde{\pi}) := d_H(\pi_U, \tilde{\pi}_U) + d_H(\pi_V, \tilde{\pi}_V)$ .

**Definition 2.2** (Earth mover's distance). Let  $(\mathcal{X}, d)$  be a metric space and let  $P, Q$  be probability distributions on  $\mathcal{X}$ . The *earth mover's distance* (a.k.a. 1-Wasserstein distance) between  $P$  and  $Q$  is  $\text{EMD}(P, Q) := \min_{\Gamma \in \Pi(P, Q)} \mathbf{E}_{(X, Y) \sim \Gamma} [d(X, Y)]$ , where  $\Pi(P, Q)$  denotes the set of couplings of  $P$  and  $Q$ . Throughout the paper,  $\mathcal{X}$  will be the set of assignments and  $d$  will be the Hamming distance  $d_H$ .

For randomized algorithms for minimization problems, we use the following convention throughout: an algorithm is an  $\alpha$ -approximation on feasible instances if it outputs a feasible solution almost surely and its expected objective value is at most  $\alpha$  times optimum. Equivalently, if the objective function of an instance  $I$  is denoted by  $\text{cost}_I$ , then

$$\mathbf{E}[\text{cost}_I(A(I))] \leq \alpha \text{opt}(I)$$

for every feasible  $I$ . When the problem is clear from context, we simply say that  $A$  satisfies the  $\alpha$ -approximation guarantee.

## 2.2 Sensitivity

We define a variant of sensitivity that is more convenient when studying the label cover problem, called swap sensitivity. Let  $I = (U, V, E, \Sigma_U, \Sigma_V, \mathcal{P}, F)$  be a left-predicate label cover instance. For  $e \in E$  and  $f : \Sigma_U \rightarrow \Sigma_V$ , let  $I^{e \leftarrow f}$  denote the instance obtained from  $I$  by replacing the projection  $f_e$  with  $f$ . For  $u \in U$  and a predicate  $Q : \Sigma_U \rightarrow \{0, 1\}$ , let  $I^{u \leftarrow Q}$  denote the instance obtained from  $I$  by replacing  $P_u$  with  $Q$ . Then, the *swap sensitivity* of a randomized algorithm  $A$  on  $I$  is defined as

$$\text{SwapSens}(A, I) := \max \left\{ \begin{array}{l} \max_{e \in E} \max_{f: \Sigma_U \rightarrow \Sigma_V} \text{EMD}(A(I), A(I^{e \leftarrow f})), \\ \max_{u \in U} \max_{Q: \Sigma_U \rightarrow \{0, 1\}} \text{EMD}(A(I), A(I^{u \leftarrow Q})) \end{array} \right\}.$$

For a family of instances  $\mathcal{I}$ , set  $\text{SwapSens}(A, \mathcal{I}) := \max_{I \in \mathcal{I}} \text{SwapSens}(A, I)$ . For  $s \in (0, 1)$  and a family of instances  $\mathcal{I}$ , let  $\mathcal{I}_{\text{sat}} \subseteq \mathcal{I}$  denote the satisfiable instances. We define  $\text{SwapSens}_s(\mathcal{I})$  to be the minimum of  $\text{SwapSens}(A, \mathcal{I})$  over all algorithms  $A$  such that for every  $I \in \mathcal{I}_{\text{sat}}$ , if  $\pi \sim A(I)$  then  $\mathbf{E}[\text{val}_I(\pi)] \geq s$ .

We say that a family of left-predicate label cover instances  $\mathcal{I}$  is *swap-closed* if for any instance  $I = (U, V, E, \Sigma_U, \Sigma_V, \mathcal{P}, F) \in \mathcal{I}$ ,  $e \in E$ , and  $u \in U$ , every projection swap  $I^{e \leftarrow f}$  and every predicate swap  $I^{u \leftarrow Q}$  also belongs to  $\mathcal{I}$ . For instances  $I$  and  $\tilde{I}$  on the same underlying graph and the same alphabets, the *swap distance*  $\text{SwapDist}(I, \tilde{I})$  is the minimum number of projection substitutions and left-predicate substitutions needed to transform  $I$  into  $\tilde{I}$ . Note that for any algorithm  $A$  and  $I, \tilde{I} \in \mathcal{I}$  for a swap-closed family of instances  $\mathcal{I}$ , we have  $\text{EMD}(A(I), A(\tilde{I})) \leq \text{SwapSens}(A, \mathcal{I}) \cdot \text{SwapDist}(I, \tilde{I})$ . For a family of instances  $\mathcal{I}$ , we define  $\text{SwapClo}(\mathcal{I})$  as the *swap-closure* of  $\mathcal{I}$ , i.e., the family of instances obtained by repeatedly swapping projections and/or left predicates in instances from  $\mathcal{I}$ .

**Lemma 2.3** (Neighboring witness extraction). Let  $\mathcal{I}$  be a family of instances equipped with  $\text{SwapDist}$ . Let  $A$  be an algorithm on  $\mathcal{I}$ . If  $\frac{\text{EMD}(A(I), A(\tilde{I}))}{\text{SwapDist}(I, \tilde{I})} \geq \lambda$  for some  $I, \tilde{I} \in \mathcal{I}$  with  $\text{SwapDist}(I, \tilde{I}) = k > 0$ , then along every shortest path  $I = I^{(0)}, I^{(1)}, \dots, I^{(k)} = \tilde{I}$  satisfying  $\text{SwapDist}(I^{(t)}, I^{(t+1)}) = 1$  for every  $t \in \{0, \dots, k-1\}$ , there exists  $t$  with  $\text{EMD}(A(I^{(t)}), A(I^{(t+1)})) \geq \lambda$ . In particular, any sensitivity lower bound can be witnessed by a neighboring pair.

*Proof.* Let  $I = I^{(0)}, I^{(1)}, \dots, I^{(k)} = \tilde{I}$  be a shortest path as above, where  $k = \text{SwapDist}(I, \tilde{I})$ . By repeated use of the triangle inequality,

$$\text{EMD}(A(I), A(\tilde{I})) \leq \sum_{t=0}^{k-1} \text{EMD}(A(I^{(t)}), A(I^{(t+1)})).$$

Hence, the average of the  $k$  summands is at least  $\frac{\text{EMD}(A(I), A(\tilde{I}))}{k} \geq \lambda$ , so at least one summand is at least  $\lambda$ .  $\square$

For comparison, if one restricts attention to projection swaps while keeping the left-predicates fixed, define the *edge-update sensitivity at  $I$*  by  $\text{EdgeUpdateSens}(A, I) := \max_{\tilde{I}} \text{EMD}(A(I), A(\tilde{I}))$ , where  $\tilde{I}$  ranges over instances obtained from  $I$  by one edge deletion or one edge insertion (with an arbitrary projection on the inserted edge). If  $I^{e\setminus}$  denotes deleting an edge  $e$  from  $I$ , then a swap  $I \rightarrow I^{e\leftarrow f}$  can be realized as the two-edge-update path  $I \rightarrow I^{e\setminus} \rightarrow I^{e\leftarrow f}$ ; hence by the triangle inequality for EMD,

$$\max_{e \in E} \max_{f: \Sigma_U \rightarrow \Sigma_V} \text{EMD}(A(I), A(I^{e\leftarrow f})) \leq 2 \cdot \max_{J \in \{I\} \cup \{I^{e\setminus}: e \in E\}} \text{EdgeUpdateSens}(A, J).$$

In particular, every swap-sensitivity lower bound that is proved using only projection swaps implies the same lower bound in the edge-deletion/update model up to a factor 2.

## Part I

# Low-Soundness Label Cover via Impagliazzo–Kabanets–Wigderson

In this chapter we prove Theorem 1.5, a sensitivity lower bound for low-soundness label cover based on the combinatorial parallel repetition theorem of Impagliazzo, Kabanets, and Wigderson (IKW) [38]. The section is split into three components. In Section 3, we introduce left/right sensitivity regions and the reduction lemma that uses them to transfer sensitivity lower bounds through reductions. In Section 4, we analyze a right-degree reduction procedure that reduces the right-degree of label cover while simultaneously amplifying its right-sensitivity by a corresponding factor. Finally, in Section 5, we introduce the IKW construction and a new corresponding low-sensitivity recovery procedure that combined with degree-reduction completes the proof.

## 3 Left/Right Sensitivity Regions

In this section, we set up an abstract framework for transferring sensitivity lower bounds through reductions. Concretely, we consider an instance transformation  $\mathcal{T}$  together with a recovery map  $\mathcal{R}$  that converts a labeling of the transformed instance back to a labeling of the source instance, and we identify conditions under which such a pair lets us pull a sensitivity lower bound back to the source family.

To make this framework work in label cover, we need to track perturbations on the left and right separately, since a reduction may amplify them by different factors. If a label cover algorithm has sensitivity  $k$ , then swapping a single constraint can alter at most  $k_L$  labels on the left and  $k_R$  labels on the right, with

$k_L + k_R \leq k$ . It is therefore more informative to reason about the pair  $(k_L, k_R)$  than only their sum. For our reduction arguments, we package this information as an upward-closed region of feasible left/right budget vectors, which is the form that composes cleanly when couplings are chained along a swap path.

To formalize this idea, we record which expected left/right perturbation budgets are attainable under couplings.

For points  $x, \tilde{x} \in \mathbb{R}^2$ , we write  $x \preceq \tilde{x}$  if  $x_1 \leq \tilde{x}_1$  and  $x_2 \leq \tilde{x}_2$  (the coordinate-wise order). Let  $S, S' \subseteq \mathbb{R}^2$  be sets of points. We write  $S \preceq_H S'$  when for any  $x \in S$ , there exists  $x' \in S'$  such that  $x \preceq x'$  (*Hoare order*). We also write

$$\uparrow S := \{x \in \mathbb{R}^2 : \exists y \in S \text{ such that } y \preceq x\}$$

for the upward closure of  $S$ . For a scalar  $c \in \mathbb{R}_{\geq 0}$  and a set  $S \subseteq \mathbb{R}^2$ , we write  $c \cdot S := \{cx : x \in S\}$ . For a matrix  $M \in \mathbb{R}_{\geq 0}^{2 \times 2}$  and a set  $S \subseteq \mathbb{R}^2$ , we write  $M \cdot S := \{Mx : x \in S\}$ .

For two assignments  $\pi = (\pi_U, \pi_V), \tilde{\pi} = (\tilde{\pi}_U, \tilde{\pi}_V)$  to a label cover instance, we define their *distance vector*  $\vec{d}(\pi, \tilde{\pi})$  as  $(d_H(\pi_U, \tilde{\pi}_U), d_H(\pi_V, \tilde{\pi}_V))$ . Let  $\Pi, \tilde{\Pi}$  be distributions over assignments of a label cover instance. We define the *attainable expected distance set* as:

$$\text{Att}(\Pi, \tilde{\Pi}) := \left\{ \mathbf{E}_{(\pi, \tilde{\pi}) \sim \mathcal{D}} \vec{d}(\pi, \tilde{\pi}) : \mathcal{D} \text{ ranges over all couplings of } \Pi \text{ and } \tilde{\Pi} \right\},$$

Its upward closure  $\uparrow \text{Att}(\Pi, \tilde{\Pi})$  is the set of all budget vectors  $x$  for which there exists a coupling of  $\Pi$  and  $\tilde{\Pi}$  whose expected distance vector is at most  $x$ .

We now aggregate these attainable budgets into a single region that must work uniformly over every source swap. For an ordinary label cover instance  $I = (U, V, E, \Sigma_U, \Sigma_V, F)$ , define the *feasible swap-sensitivity region* as

$$\text{SSR}(A, I) := \bigcap_{e \in E, f: \Sigma_U \rightarrow \Sigma_V} \uparrow \text{Att}(A(I), A(I^{e \leftarrow f})).$$

Thus,  $x \in \text{SSR}(A, I)$  if and only if for every single-edge swap of  $I$ , there exists a coupling of the two output distributions whose expected distance vector is at most  $x$ . If  $I = (U, V, E, \Sigma_U, \Sigma_V, P, F)$  is a left-predicate label cover instance, we extend this definition by also intersecting over swaps of a single left predicate:

$$\text{SSR}(A, I) := \left( \bigcap_{e \in E, f: \Sigma_U \rightarrow \Sigma_V} \uparrow \text{Att}(A(I), A(I^{e \leftarrow f})) \right) \cap \left( \bigcap_{u \in U, Q: \Sigma_U \rightarrow \{0,1\}} \uparrow \text{Att}(A(I), A(I^{u \leftarrow Q})) \right).$$

For a family of instances  $\mathcal{I}$ , define

$$\text{SSR}(A, \mathcal{I}) := \bigcap_{I \in \mathcal{I}} \text{SSR}(A, I).$$

For  $s \in (0, 1)$  and an arbitrary family of label cover instances  $\mathcal{I}$ , let  $\mathcal{I}_{\text{sat}} := \{I \in \mathcal{I} : I \text{ is satisfiable}\}$ . We define the optimal feasible swap-sensitivity region by

$$\text{SSR}_s(\mathcal{I}) := \bigcup_{A \text{ achieves soundness } s \text{ on } \mathcal{I}_{\text{sat}}} \text{SSR}(A, \mathcal{I}),$$

and note that, by construction, every set  $\text{SSR}(\cdot)$  and  $\text{SSR}_s(\cdot)$  is upward closed. The following reduction lemma is the mechanism that transfers sensitivity lower bounds through such reductions: feasible budgets compose directly when couplings are glued along a swap path.

We now record how these feasible regions behave under reductions.

**Definition 3.1** (Sensitivity-Region-Preserving Reduction). Let  $\mathcal{I}$  be a family of ordinary label cover instances or left-predicate label cover instances on the same underlying graph and domain. Let  $\mathcal{T}$  be a procedure that transforms an instance  $I \in \mathcal{I}$  into another instance  $I'$ , and let  $\mathcal{R}$  be a (possibly randomized) procedure that transforms an assignment  $\pi' = (\pi'_U, \pi'_V)$  for  $I'$  to an assignment  $\pi = (\pi_U, \pi_V)$  for  $I$  ( $\mathcal{R}$  might depend on  $I$ ). For a distribution  $\Pi'$  over assignments for  $I'$ , we write  $\mathcal{R}(\Pi')$  for the distribution obtained by sampling  $\pi' \sim \Pi'$  and then sampling  $\pi \sim \mathcal{R}(\pi')$ . For  $s, s' \in [0, 1]$ ,  $C_{\mathcal{T}} > 0$ ,  $C_{\mathcal{R}} \in \mathbb{R}_{\geq 0}^{2 \times 2}$ , we say that the pair  $(\mathcal{T}, \mathcal{R})$  is a  $(s, s', C_{\mathcal{T}}, C_{\mathcal{R}})$ -sensitivity-region-preserving reduction for  $\mathcal{I}$  if the following hold:

1. If  $I$  is satisfiable, then  $I' = \mathcal{T}(I)$  is also satisfiable.
2. If a (possibly random) assignment  $\pi'$  for  $I'$  satisfies  $\mathbf{E}[\text{val}_{I'}(\pi')] \geq s'$ , then the assignment  $\pi = \mathcal{R}(\pi')$  satisfies  $\mathbf{E}[\text{val}_I(\pi)] \geq s$ .
3. Let  $I, \tilde{I} \in \mathcal{I}$  be two instances. Then, we have  $\text{SwapDist}(\mathcal{T}(I), \mathcal{T}(\tilde{I})) \leq C_{\mathcal{T}} \cdot \text{SwapDist}(I, \tilde{I})$ . In particular, this implies that  $\mathcal{T}$  outputs instances on a common underlying graph and domain across all  $I \in \mathcal{I}$ , so that  $\text{SwapDist}(\cdot, \cdot)$  is well-defined on  $\mathcal{T}(\mathcal{I})$ .
4. Let  $I, \tilde{I} \in \mathcal{I}$  be two instances and let  $\Pi', \tilde{\Pi}'$  be distributions over assignments for  $\mathcal{T}(I)$  and  $\mathcal{T}(\tilde{I})$ , respectively. Then, we have

$$C_{\mathcal{R}} \cdot \uparrow \text{Att}(\Pi', \tilde{\Pi}') \subseteq \uparrow \text{Att}(\mathcal{R}(\Pi'), \mathcal{R}(\tilde{\Pi}')).$$

**Lemma 3.2.** Let  $\mathcal{I}$  be a swap-closed family of instances on the same underlying graph and domain. Suppose there exists an  $(s, s', C_{\mathcal{T}}, C_{\mathcal{R}})$ -sensitivity-region-preserving reduction  $(\mathcal{T}, \mathcal{R})$  for  $\mathcal{I}$ . Write  $t := \lceil C_{\mathcal{T}} \rceil$ . Then we have

$$C_{\mathcal{R}} \cdot (t \cdot \text{SSR}_{s'}(\text{SwapClo}(\mathcal{I}))) \subseteq \text{SSR}_s(\mathcal{I}),$$

where  $\mathcal{I}' = \mathcal{T}(\mathcal{I}) := \{\mathcal{T}(I) : I \in \mathcal{I}\}$ .

*Proof.* Let  $A'$  be any algorithm on  $\text{SwapClo}(\mathcal{I}')$  that achieves soundness  $s'$  on its satisfiable members. We build an algorithm  $A$  for  $\mathcal{I}$  by simulating  $A'$  on the image of each instance under  $\mathcal{T}$ . Given  $I \in \mathcal{I}$ , construct  $I' := \mathcal{T}(I)$  and run  $A'$  on  $I'$  to obtain a (possibly random) assignment  $\pi'$  for  $I'$ . Finally, return the assignment  $\pi \sim \mathcal{R}(\pi')$  for  $I$ .

We first verify the approximation guarantee of  $A$ . If  $I$  is satisfiable, then Item 1 of Definition 3.1 ensures that  $I'$  is satisfiable as well. Since  $I' \in \mathcal{I}' \subseteq \text{SwapClo}(\mathcal{I}')$ , the assumption on  $A'$  yields  $\mathbf{E}[\text{val}_{I'}(\pi')] \geq s'$ , and Item 2 implies  $\mathbf{E}[\text{val}_I(\pi)] \geq s$ .

We next bound the feasible sensitivity region of  $A$ . Fix any  $x \in \text{SSR}(A', \text{SwapClo}(\mathcal{I}'))$ . We claim that  $C_{\mathcal{R}}(tx) \in \text{SSR}(A, \mathcal{I})$ .

Let  $I, \tilde{I} \in \mathcal{I}$  satisfy  $\text{SwapDist}(I, \tilde{I}) = 1$ , and write  $I' = \mathcal{T}(I)$  and  $\tilde{I}' = \mathcal{T}(\tilde{I})$ . By Item 3, we have  $\text{SwapDist}(I', \tilde{I}') \leq C_{\mathcal{T}}$ . Hence there exists a swap path  $I' = I'_0, I'_1, \dots, I'_m = \tilde{I}'$  with  $m \leq \lceil C_{\mathcal{T}} \rceil = t$ , where each consecutive pair differs by one swap and each  $I'_j \in \text{SwapClo}(\mathcal{I}')$ .

Let  $\Pi_j := A'(I'_j)$  denote the output distribution of  $A'$  on  $I'_j$ . Because  $x \in \text{SSR}(A', \text{SwapClo}(\mathcal{I}'))$ , for each  $j \in [m]$  there exists a coupling  $\Gamma_j$  of  $\Pi_{j-1}$  and  $\Pi_j$  such that

$$\mathbf{E}_{(\pi_{j-1}, \pi_j) \sim \Gamma_j} \vec{d}(\pi_{j-1}, \pi_j) \preceq x.$$

By repeated gluing of these couplings, we obtain a joint distribution on  $(\pi_0, \dots, \pi_m)$  whose  $(j-1, j)$  marginal is  $\Gamma_j$  for every  $j$ . Under this joint distribution, the triangle inequality for Hamming distance on

each side gives

$$\vec{d}(\pi_0, \pi_m) \preceq \sum_{j=1}^m \vec{d}(\pi_{j-1}, \pi_j).$$

Taking expectations yields

$$\mathbf{E}[\vec{d}(\pi_0, \pi_m)] \preceq \sum_{j=1}^m \mathbf{E}_{(\pi_{j-1}, \pi_j) \sim \Gamma_j} \vec{d}(\pi_{j-1}, \pi_j) \preceq mx \preceq tx.$$

Hence  $tx \in \uparrow \text{Att}(\Pi_0, \Pi_m)$ . Applying Item 4 gives

$$C_{\mathcal{R}}(tx) \in \uparrow \text{Att}(\mathcal{R}(\Pi_0), \mathcal{R}(\Pi_m)) = \uparrow \text{Att}(A(I), A(\tilde{I})).$$

Since the source one-swap pair  $(I, \tilde{I})$  was arbitrary, this proves  $C_{\mathcal{R}}(tx) \in \text{SSR}(A, \mathcal{I})$ , and hence the claim.

Finally, the claim holds for every admissible  $A'$  (namely, every  $A'$  that achieves soundness  $s'$  on the satisfiable members of  $\text{SwapClo}(\mathcal{I}')$ ). Taking the union over all such  $A'$  yields

$$C_{\mathcal{R}} \cdot (t \cdot \text{SSR}_{s'}(\text{SwapClo}(\mathcal{I}'))) \subseteq \text{SSR}_s(\mathcal{I}),$$

as required. □

## 4 Right-Degree Reduction

In this section we describe a procedure which converts a label cover instance to one in which every right-vertex has the same low degree, while preserving any sensitivity lower bound that we have on the original instance.

**Definition 4.1** (Expanders). A  $d$ -regular graph  $G = (V, E)$  on  $n$  vertices is said to be an  $[n, d, \lambda]$ -expander if every eigenvalue of its normalized adjacency matrix other than the top one is at most  $\lambda$  in absolute value (equivalently, the second-largest absolute eigenvalue is at most  $\lambda$ ).

**Proposition 4.2** (Explicit Quasi-Ramanujan graphs, [3]). There exists a uniform algorithm that, for every  $d > 3$  and every sufficiently large  $n \in \mathbb{N}$  with  $nd$  even, constructs an  $[n, d, \lambda]$ -expander with  $\lambda = O(d^{-1/2})$  in time polynomial in  $n$  and  $d$ .

Let  $d \geq 4$  be right-degree that we would like our label cover instance to have. Degree reduction proceeds by replacing each right-vertex  $v$ , with degree  $D_v$ , by a “cloud” of  $D_v$ -many new vertices. We replace the edges to  $v$  by edges from the  $D_v$ -many vertices to the neighborhood of  $v$  by using a  $d$ -regular expander graph as a template; this disperses the information stored at  $v$  among the  $D_v$  vertices. We describe this procedure, which can be seen pictorially in Figure 1, formally next.

Because Proposition 4.2 only guarantees expanders for pairs  $(n, d)$  satisfying its hypotheses, we formulate the degree-reduction map relative to a fixed choice of expanders. Fix an even integer  $d \geq 4$  and a parameter  $\lambda \in [0, 1)$ . Let  $I = (U, V, E, \Sigma_U, \Sigma_V, P, F)$  be a left-predicate label cover instance whose right degrees satisfy  $\deg(v) \geq d$  for every  $v \in V$ . For each right vertex  $v \in V$ , let  $D_v$  denote its degree and set  $\delta_V := \min_{v \in V} D_v$ . Since  $d$  is even, the parity condition  $D_v d$  even from Proposition 4.2 is automatic for every  $v \in V$ . A  $(d, \lambda)$ -expander package for  $I$  is a choice, for each  $v \in V$ , of a  $[D_v, d, \lambda]$ -expander  $H_v = ([D_v], E'')$ . Once such a package is fixed, we define the transformation  $\mathcal{T}_{\text{DR}}(I; d)$  as follows. We fix

an arbitrary ordering of the neighbors of each right vertex  $v \in V$ , and denote the  $i$ -th neighbor of  $v$  in the bipartite graph  $G = (U, V, E)$  by  $\Gamma_G(v, i)$ . Let  $\Gamma_{H_v}(i, k)$  denote the  $k$ -th neighbor of  $i$  in  $H_v$ .

---

**Algorithm 1:** Right-Degree Reduction  $\mathcal{T}_{\text{DR}}$

---

**Input:** A left-predicate label cover instance  $I = (U, V, E, \Sigma_U, \Sigma_V, P, F)$ , an integer  $d \geq 4$ , a  $(d, \lambda)$ -expander package, and an ordering on the neighbors of each vertex  
**Output:** An instance  $I' = (U, V', E', \Sigma_U, \Sigma_V, P, F')$  with right-degree  $d$ .

```

1 for  $v \in V$  do
2   Let  $D_v$  be the degree of  $v$ ;
3   Let  $H_v$  be the  $[D_v, d, \lambda]$ -expander for  $v$  in the expander package;
4   Add to  $V'$  a vertex  $(v, i)$  for each  $i \in [D_v]$ ;
5   for  $i \in [D_v]$  do
6     Define the edges in  $E'$  that are incident to  $(v, i)$  as follows;
7     The  $k$ -th neighbor of  $(v, i)$  is the  $j$ -th neighbor of  $v$  in  $G$ ;
8 for  $e' = (u, (v, i)) \in E'$  do
9   The constraint  $f_{e'} : \Sigma_U \rightarrow \Sigma_V \in F'$  for  $e'$  is  $f_{(u,v)}$ ;
10 return  $I'$ 

```

---

We observe some properties of the instance output by degree-reduction. First,  $|V'| = |E|$ . For  $G' = (V', E')$ , each neighborhood satisfies  $\Gamma_{G'}((v, i), k) := \Gamma_G(v, \Gamma_{H_v}(i, k))$ . Finally, observe that the degree of each vertex  $u \in U$  in  $G'$  becomes  $d \cdot \deg_G(u)$ .

The recovery map  $\mathcal{R}_{\text{DR}}$  converts an assignment  $\pi' = (\pi'_U, \pi'_V)$  for  $I' = \mathcal{T}_{\text{DR}}(I; d)$  into a random assignment  $\pi$  for  $I = (U, V, E, \Sigma_U, \Sigma_V, P, F)$ . It copies the left labels, i.e., sets  $\pi_U(u) := \pi'_U(u)$  for each  $u \in U$ . For each right vertex  $v \in V$ , it samples an index  $i_v \in [D_v]$  uniformly and independently, and sets  $\pi_V(v) := \pi'_V(v, i_v)$ .

**Lemma 4.3.** Let  $d \geq 4$  be even, and let  $\lambda \in [0, 1)$ . Let  $\mathcal{I}$  be a swap-closed family of left-predicate label cover instances on a common underlying graph and domain, whose minimum right degree is at least  $\delta_V \geq d$ . For each right vertex  $v$  of this common underlying graph, let  $D_v$  denote its (common) degree across the family, and fix a  $[D_v, d, \lambda]$ -expander  $H_v$ . Let  $\mathcal{I}' = \text{SwapClo}(\mathcal{T}_{\text{DR}}(\mathcal{I}; d))$ . Then  $\mathcal{I}'$  consists of right-regular left-predicate label cover instances with right degree  $d$ , and for any  $\varepsilon > 0$ , we have

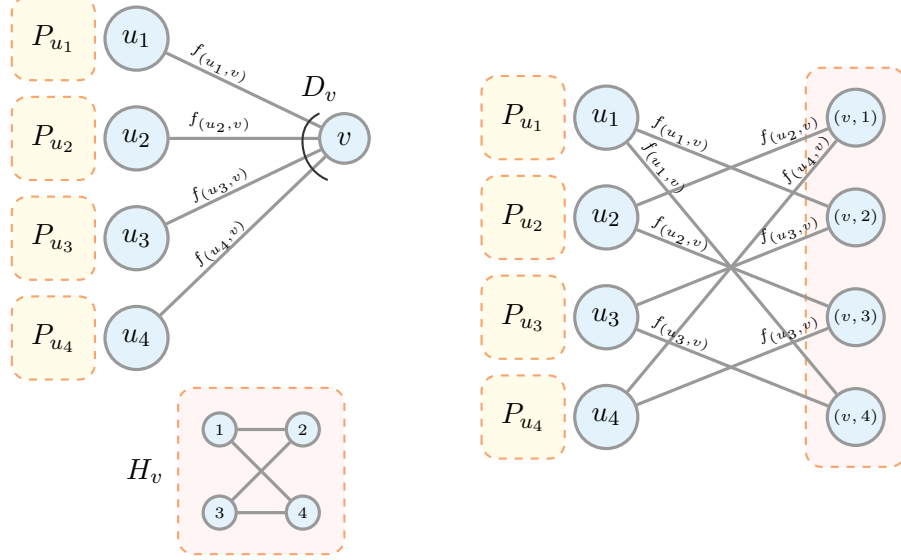
$$C_{\mathcal{R}} \cdot (d \cdot \text{SSR}_{\varepsilon + O(\lambda)}(\mathcal{I}')) \subseteq \text{SSR}_{\varepsilon}(\mathcal{I}),$$

where  $C_{\mathcal{R}} = \begin{pmatrix} 1 & 0 \\ 0 & 1/\delta_V \end{pmatrix}$ .

*Proof.* We show that the pair  $(\mathcal{T}_{\text{DR}}, \mathcal{R}_{\text{DR}})$  is a  $(\varepsilon, \varepsilon + O(\lambda), C_{\mathcal{T}}, C_{\mathcal{R}})$ -sensitivity-region-preserving reduction with  $C_{\mathcal{T}} = d$  and  $C_{\mathcal{R}} = \begin{pmatrix} 1 & 0 \\ 0 & 1/\delta_V \end{pmatrix}$ . Hence, the claim follows by Lemma 3.2.

Item 1 of Definition 3.1 is immediate. If  $I$  is satisfiable by  $(\pi_U, \pi_V)$ , then setting  $\pi'_U := \pi_U$  and  $\pi'_V(v, i) := \pi_V(v)$  for every  $v \in V$  and  $i \in [D_v]$  satisfies every copied edge projection in  $I'$ , while the left predicates are unchanged.

For Item 2, we use the approximation guarantee proved in [20, Theorem 5.4] for the degree-reduction gadget. Their argument depends only on the copied edge projections inside each cloud, while our transformation leaves the left predicates unchanged. Therefore the same recovery map  $\mathcal{R}_{\text{DR}}$  yields the stated  $\varepsilon + O(\lambda)$  versus  $\varepsilon$  guarantee in the left-predicate setting as well.



(a) The neighborhood of  $v$  before degree-reduction and auxiliary expander graph  $H_v$ . (b) The portion of the label-cover instance corresponding to  $v$  after degree-reduction.

Figure 1: Before and after applying degree-reduction on a right vertex  $v$  with graph  $H_v$ .

For Item 3, consider a single swap in the source instance. If it is an edge-projection swap on an original edge  $(u, v)$ , then exactly the  $d$  copied edges incident with the cloud of  $v$  change, so the target instances differ by at most  $d$  edge-projection swaps. If it is a predicate swap at a left vertex  $u$ , then only the single left predicate at  $u$  changes, because  $\mathcal{T}_{\text{DR}}$  leaves the family  $P$  unchanged. Thus  $C_{\mathcal{T}} = d$ .

It remains to verify  $C_{\mathcal{R}}$ . Fix any two deterministic assignments  $\pi' = (\pi'_U, \pi'_V)$  and  $\tilde{\pi}' = (\tilde{\pi}'_U, \tilde{\pi}'_V)$  for  $I'$ . Consider the natural coupling of  $\mathcal{R}_{\text{DR}}(\pi')$  and  $\mathcal{R}_{\text{DR}}(\tilde{\pi}')$  that uses the *same* random index  $i_v \sim [D_v]$  (uniformly and independently over  $v \in V$ ) in both recovered assignments. Under this coupling, the left labels are copied deterministically, so  $d_{\text{H}}(\pi_U, \tilde{\pi}_U) = d_{\text{H}}(\pi'_U, \tilde{\pi}'_U)$ . On the right side, for each  $v \in V$  we have

$$\Pr [\pi_V(v) \neq \tilde{\pi}_V(v)] = \frac{|\{i \in [D_v] : \pi'_V(v, i) \neq \tilde{\pi}'_V(v, i)\}|}{D_v} \leq \frac{1}{\delta_V} \cdot |\{i \in [D_v] : \pi'_V(v, i) \neq \tilde{\pi}'_V(v, i)\}|.$$

Summing over all  $v \in V$  and taking expectation shows

$$\mathbf{E} [d_{\text{H}}(\pi_V, \tilde{\pi}_V)] \leq \frac{1}{\delta_V} \cdot d_{\text{H}}(\pi'_V, \tilde{\pi}'_V).$$

Hence this coupling yields

$$\mathbf{E} [\vec{d}(\pi, \tilde{\pi})] \preceq C_{\mathcal{R}} \cdot \vec{d}(\pi', \tilde{\pi}').$$

Let  $\Pi', \tilde{\Pi}'$  be arbitrary distributions over assignments for  $I'$  and fix any

$$x \in \uparrow \text{Att}(\Pi', \tilde{\Pi}').$$

By definition of upward closure, there exists a coupling  $\mathcal{D}$  of  $\Pi'$  and  $\tilde{\Pi}'$  such that

$$\mathbf{E}_{(\pi', \tilde{\pi}') \sim \mathcal{D}} \vec{d}(\pi', \tilde{\pi}') \preceq x.$$

Sample  $(\pi', \tilde{\pi}') \sim \mathcal{D}$ , and conditioned on this pair, apply the shared-index coupling of the recovery randomness described above. Averaging the deterministic bound gives

$$\mathbf{E} [\vec{d}(\pi, \tilde{\pi})] \preceq C_{\mathcal{R}} \cdot \mathbf{E}_{(\pi', \tilde{\pi}') \sim \mathcal{D}} \vec{d}(\pi', \tilde{\pi}') \preceq C_{\mathcal{R}} \cdot x.$$

Hence  $C_{\mathcal{R}} \cdot x \in \uparrow \text{Att}(\mathcal{R}_{\text{DR}}(\Pi'), \mathcal{R}_{\text{DR}}(\tilde{\Pi}'))$ . Since  $x$  was arbitrary, this verifies Item 4.  $\square$

**Corollary 4.4** (Forbidden-point consequence). In the setting of Lemma 4.3, let  $y = (y_L, y_R) \in \mathbb{R}_{\geq 0}^2$  satisfy

$$y \notin \text{SSR}_{\varepsilon}(\mathcal{I}).$$

Then

$$\text{SSR}_{\varepsilon+O(\lambda)}(\mathcal{I}') \cap \{x \in \mathbb{R}_{\geq 0}^2 : C_{\mathcal{R}}(dx) \preceq y\} = \emptyset.$$

Equivalently,

$$\text{SSR}_{\varepsilon+O(\lambda)}(\mathcal{I}') \cap ([0, y_L/d] \times [0, \delta_V y_R/d]) = \emptyset.$$

*Proof.* Suppose for contradiction that  $x \in \text{SSR}_{\varepsilon+O(\lambda)}(\mathcal{I}')$  and  $C_{\mathcal{R}}(dx) \preceq y$ . Then Lemma 4.3 gives

$$C_{\mathcal{R}}(dx) \in \text{SSR}_{\varepsilon}(\mathcal{I}).$$

Since  $\text{SSR}_{\varepsilon}(\mathcal{I})$  is upward closed and  $C_{\mathcal{R}}(dx) \preceq y$ , it follows that  $y \in \text{SSR}_{\varepsilon}(\mathcal{I})$ , contradicting the hypothesis. The displayed rectangle is exactly the set  $\{x \in \mathbb{R}_{\geq 0}^2 : C_{\mathcal{R}}(dx) \preceq y\}$  because  $C_{\mathcal{R}} = \begin{pmatrix} 1 & 0 \\ 0 & 1/\delta_V \end{pmatrix}$ .  $\square$

## 5 Low-Soundness Label Cover Route via Direct Product Testing

### 5.1 The IKW Construction and Its Label-Cover Consequence

We are now ready to prove the main sensitivity lower bound of this chapter, Theorem 1.5. We re-state the result for convenience.

**Theorem 5.1.** There exists  $C > 0$  such that for all sufficiently large  $N_{\text{IKW}} \in \mathbb{N}$  and  $k \leq o(\log N_{\text{IKW}})$ , any algorithm that, on satisfiable left-predicate label cover instances on  $n = N_{\text{IKW}}^{\Theta(k)}$  vertices, outputs an  $\exp(-C\sqrt{k})$ -satisfying solution must have sensitivity at least  $n^{\Omega(1/k)}$ . Moreover, the hard instances can be taken to be biregular with alphabet and degree parameters at most  $2^{O(k)}$ .

Theorem 5.1 is proved using a modification of the combinatorial parallel repetition (direct product) theorem of Impagliazzo, Kabanets, and Wigderson (IKW) [38]. We note above and throughout this section, we write  $N_{\text{IKW}}$  for the size of the base 2CSP used by the direct-product construction, and  $n$  in the statement for the final label cover size. We start by overviewing the construction and relevant parameters below.

**Parameter guide.** Let  $N_{\text{IKW}}$  denotes the number of variables in the base 2CSP. This separates the base-instance size used by the IKW construction from the final label cover left-side size  $n_{\text{LC}}$  used in later sections.

We first recall the label cover instance analyzed by IKW [38] in the perfect completeness regime. The construction starts with a ‘slightly hard’ (constant soundness) base 2-CSP  $\Phi$  on a  $d$ -regular graph  $G = ([N_{\text{IKW}}], E)$  with alphabet  $\Sigma$ . We first describe the constraint graph of IKW’s amplified instance, which we

denote by  $\text{IKW}_k^\Phi$ . Fix  $k \in \mathbb{N}$  and some  $k' = \Theta(\sqrt{k})$  following [38].<sup>5</sup> The underlying graph of  $\text{IKW}_k^\Phi$  has left vertex set  $U = \binom{E}{k}$  and right vertex set  $V = \binom{[N_{\text{IKW}}]}{k'}$ , and we view it as a left-predicate label cover instance  $(U, V, E_{\text{IKW}}, \Sigma_U, \Sigma_V, P, F)$ . Edges in  $E_{\text{IKW}}$  are generated via the following process:

1. Sample a size  $k'$  subset  $A \subset [N_{\text{IKW}}]$  (i.e. a random vertex in  $V$ )
2. Sample a random edge for each vertex in  $A$ , and call the resulting edge-set  $A_{E,2}$ <sup>6</sup> where  $A_{E,2} \subseteq E$  and  $|A_{E,2}| = k'$
3. Sample  $k - k'$  additional random edges  $B_{E,2} \subseteq E$  without replacement<sup>7</sup> so that  $A_{E,2} \cap B_{E,2} = \emptyset$ .
4. Output the edge  $(\{A_{E,2}, B_{E,2}\}, A)$

In particular, every sampled pair  $(A_{E,2}, B_{E,2})$  determines a left vertex  $S = (A_{E,2}, B_{E,2}) \in U$ , while the right endpoint  $A$  lies in  $V = \binom{[N_{\text{IKW}}]}{k'}$ . The left alphabet is  $\Sigma_U = \Sigma^{2k}$  (one symbol for each of the  $2k$  original vertices in  $(A_{E,2}, B_{E,2})$ ), and the right alphabet is  $\Sigma_V = \Sigma^{k'}$  (one symbol per original vertex in  $A$ ). For  $(\pi_U, \pi_V)$  to satisfy the edge  $(\{A_{E,2}, B_{E,2}\}, A) \in E_{\text{IKW}}$ , it must hold that

1. The left label  $\pi_U(A_{E,2}, B_{E,2})$  satisfies all constraints in  $B_{E,2}$  in the original CSP, and
2.  $\pi_U$ 's vertex assignments on  $A$  agree with  $\pi_V(A)$ .

The left-predicates  $P_u$  enforce the first of these constraints, and the associated projection  $f_{\{A_{E,2}, B_{E,2}\}, A} \in F$  enforces the second. To align with our global notation, we denote the left labeling by  $\pi_U$  and write  $\tilde{\pi}_U$  for the corresponding labeling produced on a neighboring instance. Formally,  $\pi_U : U \rightarrow \Sigma_U$  and  $\tilde{\pi}_U : U \rightarrow \Sigma_U$ , and similarly  $\pi_V : V \rightarrow \Sigma_V$  and  $\tilde{\pi}_V : V \rightarrow \Sigma_V$ .

IKW [38] prove the above CSP has soundness  $\exp(-\Omega(\sqrt{k}))$ , which is quasi-polynomially related to the alphabet-size and left-degree  $2^{O(k)}$ . This will us to push to lower soundness without ruining the sensitivity. In particular, the main result of this section is the following sensitivity lower bound for  $\text{IKW}_k^\Phi$ :

**Theorem 5.2** (IKW Sensitivity). For every constant  $c > 0$  there are constants  $C_1, C_2 > 0$  such that for any  $\delta, \tau > 0$ , any large enough  $N_{\text{IKW}} \in \mathbb{N}$ , any  $k \leq o(\log N_{\text{IKW}})$ , and any regular 2-CSP  $\Phi$  on  $N_{\text{IKW}}$  variables such that  $\text{Max-}\Phi(1, 1 - \delta)$  has no  $N_{\text{IKW}}^c$ -sensitive algorithm,  $\text{Max-}\text{IKW}_k^\Phi(1, e^{-C_1 \delta \sqrt{k}})$  has no  $(N_{\text{IKW}}^{c-\tau}, \frac{|V|}{|E_{\text{IKW}}|} N_{\text{IKW}}^{c-\tau})$ -sensitive algorithm that succeeds with probability more than  $1 - C_2 \delta$ .

The key point for the later degree-reduction step is that Theorem 5.2 already gives exactly the kind of *single forbidden budget point* required by Corollary 4.4; one does not need to characterize the entire source feasible region. Concretely, if

$$y_{N_{\text{IKW}}, k} := \left( N_{\text{IKW}}^{c-\tau}, \frac{|V|}{|E_{\text{IKW}}|} N_{\text{IKW}}^{c-\tau} \right),$$

then Theorem 5.2 says that this point is forbidden for the IKW family. Since the IKW graph is right-regular, its common right degree is

$$\delta_V = \frac{|E_{\text{IKW}}|}{|V|}.$$

<sup>5</sup>We remark that [18, 21] study a related test for  $k' = O(k)$ , which leads to improved soundness. However, the analysis, and hence implicitly the decoding, in these works requires a ‘smoothing’ operation over the original assignment that may ruin the sensitivity of the decoding on the original instance.

<sup>6</sup>We remark this notation is due to a different CSP analyzed earlier in IKW’s work. We follow their notation regardless to make references to their proof clear.

<sup>7</sup>In fact we will assume in the analysis that the vertices in this set are entirely disjoint. This does not change the argument in the regime where  $k \leq o(\log(N_{\text{IKW}}))$  due to the low probability of re-sampling. The distinction of sampling with or without replacement is not made in [38] for this reason.

Thus, after right-degree reduction, Corollary 4.4 multiplies the weak second coordinate by exactly the compensating factor  $\delta_V$ , turning the forbidden point above into a balanced forbidden square.

*Proof of Theorem 5.1.* Let  $\mathcal{F}_{N_{\text{IKW}}}$  be the constant-alphabet, bounded-degree, regular base 2-CSP family from Fleming–Yoshida [28]. Thus there exist constants  $c, \delta_0 > 0$  such that  $\text{Max-}\Phi(1, 1 - \delta_0)$  has no  $N_{\text{IKW}}^c$ -sensitive algorithm for  $\Phi \in \mathcal{F}_{N_{\text{IKW}}}$ . For each  $k$ , let

$$\mathcal{I}_{N_{\text{IKW}},k}^{\text{IKW}} := \text{SwapClo}(\{\text{IKW}_k^\Phi : \Phi \in \mathcal{F}_{N_{\text{IKW}}}\})$$

be the swap-closed family generated by the corresponding IKW instances.

Applying Theorem 5.2 with  $\delta = \delta_0$  and any fixed  $\tau \in (0, c/2)$ , we obtain a forbidden budget point

$$y_{N_{\text{IKW}},k} := \left( N_{\text{IKW}}^{c-\tau}, \frac{|V|}{|E_{\text{IKW}}|} N_{\text{IKW}}^{c-\tau} \right)$$

for the family  $\mathcal{I}_{N_{\text{IKW}},k}^{\text{IKW}}$  at soundness

$$\varepsilon_k := \exp\left(-C_1 \delta_0 \sqrt{k}\right).$$

In the notation of Section 3, this simply means that the point  $y_{N_{\text{IKW}},k}$  is not contained in the feasible sensitivity region of any algorithm achieving soundness  $\varepsilon_k$  on this family.

Next choose an even degree parameter

$$d_\star = \Theta\left(\exp\left(C_\star \sqrt{k}\right)\right)$$

for a sufficiently large absolute constant  $C_\star$ , and let  $\lambda_\star = O(d_\star^{-1/2})$  be the corresponding expander parameter from Proposition 4.2. By increasing  $C_\star$  if necessary, we may ensure that

$$O(\lambda_\star) \leq \frac{1}{2} \varepsilon_k.$$

Since the original IKW right degree is  $N_{\text{IKW}}^{\Theta(k)}$ , it is in particular at least  $d_\star$  for all sufficiently large  $N_{\text{IKW}}$ , so the right-degree reduction from Section 4 applies to the family  $\mathcal{I}_{N_{\text{IKW}},k}^{\text{IKW}}$ .

Let

$$\mathcal{I}'_{N_{\text{IKW}},k} := \text{SwapClo}(\mathcal{T}_{\text{DR}}(\mathcal{I}_{N_{\text{IKW}},k}^{\text{IKW}}; d_\star))$$

be the resulting degree-reduced family. Because the IKW graph is right-regular with  $\delta_V = |E_{\text{IKW}}|/|V|$ , Corollary 4.4 gives

$$\text{SSR}_{\varepsilon_k + O(\lambda_\star)}(\mathcal{I}'_{N_{\text{IKW}},k}) \cap \left[0, \frac{N_{\text{IKW}}^{c-\tau}}{d_\star}\right]^2 = \emptyset.$$

Equivalently, no algorithm on  $\mathcal{I}'_{N_{\text{IKW}},k}$  that achieves soundness  $\varepsilon_k + O(\lambda_\star)$  can have scalar sensitivity at most  $N_{\text{IKW}}^{c-\tau}/d_\star$ , since scalar sensitivity  $t$  implies pair sensitivity at most  $(t, t)$ . By the choice of  $d_\star$ , the target soundness remains

$$\varepsilon_k + O(\lambda_\star) = \exp\left(-\Omega(\sqrt{k})\right).$$

The alphabet sizes are unchanged by degree reduction. Since the base alphabet is constant, the IKW alphabets are  $|\Sigma|^{2k}$  on the left and  $|\Sigma|^{k'}$  on the right, hence both are at most  $2^{O(k)}$ . The transformed right degree is exactly  $d_\star = 2^{O(\sqrt{k})}$ . The transformed left degree is the original IKW left degree multiplied by  $d_\star$ , and is therefore at most  $2^{O(k)}$  as well.

Finally, the left side of the IKW graph has size

$$|U| = \binom{|E(\Phi)|}{k} = N_{\text{IKW}}^{\Theta(k)},$$

because the base CSP has bounded degree and hence  $|E(\Phi)| = \Theta(N_{\text{IKW}})$ . After right-degree reduction, the right side has size

$$|V'| = |E_{\text{IKW}}| = N_{\text{IKW}}^{\Theta(k)},$$

since the IKW left degree is at most  $2^{O(k)}$ . Thus the final total number of vertices satisfies

$$n = N_{\text{IKW}}^{\Theta(k)}.$$

Moreover,

$$\frac{N_{\text{IKW}}^{c-\tau}}{d_\star} = N_{\text{IKW}}^{c-\tau-o(1)} = N_{\text{IKW}}^{\Omega(1)} = n^{\Omega(1/k)},$$

because  $d_\star = \exp(O(\sqrt{k}))$  and  $k \leq o(\log N_{\text{IKW}})$ . This proves Theorem 5.1.  $\square$

## 5.2 Proof of the IKW Sensitivity Theorem

We briefly outline the proof of Theorem 5.2. We start with two neighboring instances  $I_\Phi$  and  $\tilde{I}_\Phi$  of the original CSP. We map both instances through the IKW encoding and show that the outputs remain reasonably close. Assuming a low-sensitivity algorithm for the IKW CSP, we run it on the encodings of  $I_\Phi$  and  $\tilde{I}_\Phi$  to obtain two solutions to the IKW CSP that remain reasonably close. We then give a simple local decoding algorithm for the IKW CSP that preserves sensitivity while outputting a highly satisfying solution to the base CSP with non-trivial probability. Finally, we run this decoder sufficiently many times to ensure that we obtain a highly satisfying solution with high probability, and apply a low-sensitivity selection rule to output a good, low-sensitivity solution to the original CSP, contradicting the base CSP's lower bound.

*Proof of Theorem 5.2.* We start with the first claim that the IKW encoding has relatively low (normalized) sensitivity. In particular, given an instance  $I_\Phi$ , write  $\text{IKW}_k(I_\Phi)$  to denote its encoded form. We claim:

**Claim 5.3** (Encoding Blowup). Let  $I_\Phi, \tilde{I}_\Phi$  be one-swap neighboring instances of a  $d$ -regular 2-CSP on  $N_{\text{IKW}}$  variables. The fraction of constraints differing between  $\text{IKW}_k(I_\Phi)$  and  $\text{IKW}_k(\tilde{I}_\Phi)$  is at most  $O(\frac{k}{dN_{\text{IKW}}})$ .

*Proof.* The proof is immediate from the above description of the construction. Namely, the constraints in the new CSP only differ if the changed constraint is sampled in  $B_{E,2}$ . Since  $B_{E,2}$  is a random subset of  $k - k'$  edges, the probability we sample a set including the changed edge is  $1 - \frac{\binom{|E|-1}{k-k'}}{\binom{|E|}{k-k'}} \leq O(\frac{k}{dN_{\text{IKW}}})$ .  $\square$

Now, fix  $\varepsilon := e^{-C_1\delta\sqrt{k}}$  for some sufficiently small constant  $C_1 > 0$ , and assume for the sake of contradiction there exists a  $(N_{\text{IKW}}^{c-\tau}, \frac{|V|}{|E_{\text{IKW}}|} N_{\text{IKW}}^{c-\tau})$ -sensitive algorithm which, given a satisfiable instance of  $\text{IKW}_k^\Phi$ , outputs a (possibly randomized) solution which is  $\varepsilon$ -satisfying with probability at least  $1 - C_2\delta$  for some sufficiently small  $C_2 > 0$ . We will show this implies a  $o(N_{\text{IKW}}^c)$ -sensitive algorithm for the original CSP  $\Phi$  that is  $(1 - O(\delta))$ -satisfying in expectation, violating our original lower bound.<sup>8</sup>

<sup>8</sup>Technically we should achieve  $1 - \delta$  expected accuracy as written in the statement, but this simply amounts to shifting the constant factors  $C_1$  and  $C_2$  appropriately.

Let  $I_\Phi$  be a satisfiable instance to the original CSP and  $\tilde{I}_\Phi$  a one-swap neighboring instance. It is clear from construction that the encoding  $\text{IKW}_k(I_\Phi)$  is satisfiable, and by Claim 5.3 that  $\text{IKW}_k(I_\Phi)$  and  $\text{IKW}_k(\tilde{I}_\Phi)$  differ in at most  $O(|E_{\text{IKW}}| \frac{k}{dN_{\text{IKW}}})$  constraints. Thus, running our  $(N_{\text{IKW}}^{c-\tau}, \frac{|V|}{|E_{\text{IKW}}|} N_{\text{IKW}}^{c-\tau})$ -sensitive algorithm  $\mathcal{A}_{\text{IKW}}$  for  $\text{IKW}_k^\Phi$  on these instances, we are guaranteed output solutions  $\pi = (\pi_U, \pi_V)$  and  $\tilde{\pi} = (\tilde{\pi}_U, \tilde{\pi}_V)$  satisfying

1. **U-Close:** The expected normalized Hamming distance between  $\pi_U$  and  $\tilde{\pi}_U$  is at most

$$O\left(\frac{k}{dN_{\text{IKW}}^{1-c+\tau}} \cdot \frac{|E_{\text{IKW}}|}{|U|}\right) \leq \frac{2^{O(k)}}{dN_{\text{IKW}}^{1-c+\tau}}$$

2. **V-Close:** The expected normalized Hamming distance between  $\pi_V$  and  $\tilde{\pi}_V$  is at most

$$\frac{1}{|V|} \cdot |E_{\text{IKW}}| \frac{k}{dN_{\text{IKW}}} \cdot \frac{|V|}{|E_{\text{IKW}}|} N_{\text{IKW}}^{c-\tau} \leq O\left(\frac{k}{N_{\text{IKW}}^{1-c+\tau}}\right)$$

3. **High value:**  $\pi$  is  $\varepsilon$ -satisfying with probability at least  $1 - C_2\delta$ .

We now present a low-sensitivity decoding algorithm for the above solutions that outputs close, highly satisfying solutions to  $\Phi$  in expectation. We begin by introducing some notation.

Let us focus for now on a fixed realization  $\pi = (\pi_U, \pi_V)$  of  $\pi = (\pi_U, \pi_V)$  (the analysis for a realization  $\tilde{\pi} = (\tilde{\pi}_U, \tilde{\pi}_V)$  of  $\tilde{\pi} = (\tilde{\pi}_U, \tilde{\pi}_V)$  is identical). Recall that the constraints of our CSP are defined by first selecting a random subset  $A \subset \binom{[N_{\text{IKW}}]}{k'}$  of  $k'$  vertices in the original CSP. For a given assignment  $a \in \Sigma^{k'}$  to these vertices, define the set of ‘consistent hyperedges’ in  $U = \binom{E}{k}$  as

$$\text{Cons}_{A,a} := \{(A_{E,2}, B_{E,2}) \in U : A_{E,2} \supset A \wedge \pi_U(A_{E,2}, B_{E,2})|_A = a\}.$$

and define the set of consistent hyperedges (vertices in  $U$ ) that contain a particular  $x \in [N_{\text{IKW}}] \setminus A$  as

$$\text{Cons}_{A,a}^x := \{(A_{E,2}, B_{E,2}) \in U : A_{E,2} \supset A \wedge \pi_U(A_{E,2}, B_{E,2})|_A = a \wedge x \in (A_{E,2}, B_{E,2})\}.$$

IKW define a decoding for every such pair  $(A, a)$  by taking the plurality response conditioned on being in Cons. Namely, they consider the following randomized decoding family  $\{g_A : [N_{\text{IKW}}] \rightarrow \Sigma\}$ :

1. Sample a right vertex  $A \in \binom{[N_{\text{IKW}}]}{k'}$  and set  $a = \pi_V(A) \in \Sigma^{k'}$
2. Define  $g_A(x)$  as the plurality value over  $S_E = (A_{E,2}, B_{E,2})$  (left vertices) in Cons containing  $x$ :<sup>9</sup>

$$g_A(x) = \text{maj}_{S_E \in \text{Cons}_{A,a}^x}(\pi_U(S_E)|_x).$$

We define  $\tilde{g}_A$  similarly using  $\tilde{\pi}_V$ .

A direct inspection of the proof of [38, Theorem 6.4] gives the following useful facts about the above randomized decoding:<sup>10</sup>

**Theorem 5.4.** If  $\pi = (\pi_U, \pi_V)$  has value at least  $\varepsilon$ , then an  $\Omega(\varepsilon)$  fraction of  $A \in V$  satisfy:

<sup>9</sup>Ties may be broken arbitrarily. If  $\text{Cons}_{A,a}^x$  is empty, we fix  $g_{A,a}(x)$  to be some arbitrary symbol in  $\Sigma$ , say 0.

<sup>10</sup>In particular, see the explanation at the bottom of Pg. 43 of [38] along with the discussion of soundness for [38, Theorem 6.1] in the second paragraph of Pg. 42.

1.  $g_A$  is a  $(1 - \delta)$ -satisfying solution to the base CSP  $I_\Phi$
2.  $\text{Cons}_{A,a}$  has measure at least  $\Omega(\varepsilon)$
3. Almost all consistent sets nearly compute  $g_A(x)$ :

$$\Pr_{S_E=(A_E, B_{E,2}) \in U} \left[ \pi_U(S_E)|_{B_{E,2}} \stackrel{O(\delta)}{\neq} g_A(B_{E,2}) \mid S_E \in \text{Cons}_{A,a} \right] \leq O(\varepsilon^3),$$

where  $\stackrel{O(\delta)}{\neq}$  denotes the two functions disagreeing on more than an  $O(\delta)$  fraction of the domain.

Now, given the above, note that a random choice of  $g_A$  gives a successful decoding with probability  $\Omega(\varepsilon)$ . The natural strategy would be to sample roughly  $O(\frac{\log \frac{1}{\delta}}{\varepsilon})$  independent  $(A, a)$  pairs to ensure that one satisfies the above properties with high probability, and apply a low-sensitivity procedure to pick one of the outputs  $g_{A,a}$  with high value (indeed we will perform a variant of this approach shortly).

However, there is a problem with this approach: as defined above, the value of  $g_{A,a}(x)$  depends on a large and dependent subset of  $\pi_U$ . A priori, this means the (expected)  $\frac{2^{O(k)}}{dN_{\text{IKW}}^{1-c+\tau}}$  fraction of changes between  $\pi_U$  and  $\tilde{\pi}_U$  might actually result in our decodings  $g_A$  and  $\tilde{g}_A$  being far, ruining the sensitivity.

We avoid this issue by introducing a simpler randomized decoding for  $\text{IKW}_k^\Phi$  that is highly local and uncorrelated. We will argue that these properties allow us to preserve sensitivity, while items (2) and (3) of Theorem 5.4 ensure that the decoding remains  $1 - O(\delta)$  satisfying with non-trivial probability. The resulting randomized decoding family is given by the following process:

1. Sample  $(A, a)$  as described before, i.e. draw  $A \in \binom{[N_{\text{IKW}}]}{k'}$  uniformly and set  $a = \pi_V(A) \in \Sigma^{k'}$
2. For each  $x \in [N_{\text{IKW}}]$ :
  - Draw a uniformly random edge  $e \in E$  of the original CSP with  $x \in e$
  - Draw  $O(\frac{\log N_{\text{IKW}}}{\varepsilon})$  independent samples  $S_E = (A_{E,2}, B_{E,2}) \in U$  uniformly conditioned on  $e \subseteq S_E$
  - If any of the  $S_E \in \text{Cons}_{A,a}$ , pick a random such  $S_E$  and output  $\pi_U(S_E)|_x$
  - Else: Output 0.

We remark while it would be more natural to skip the first step and simply sample  $O(\frac{\log N_{\text{IKW}}}{\varepsilon})$  random  $(A_{E,2}, B_{E,2})$  containing  $x$ , sampling the common edge  $e$  turns out to simplify the correctness analysis.

We will denote the above randomized decoding as  $\mathbf{g}_A^{\text{rand}}$ , where rand stands for the internal randomness over the choices of  $e$  and  $\{S_E\}$ . We first argue that  $\mathbf{g}_A^{\text{rand}}$  and  $\tilde{\mathbf{g}}_A^{\text{rand}}$  are very close in expectation over a random choice of  $(A, a)$  and the internal randomness of the above decoding:

**Claim 5.5** (Sensitivity of  $\mathbf{g}_A^{\text{rand}}$ ). Over the randomness of  $\mathcal{A}_{\text{IKW}}$ ,  $A \in V$ , and choices of  $e, S_E$ :

$$\mathbf{E} \left[ \text{dist}(\mathbf{g}_A^{\text{rand}}, \tilde{\mathbf{g}}_A^{\text{rand}}) \right] \leq O \left( \frac{\log N_{\text{IKW}}}{\varepsilon} \cdot \frac{2^{O(k)}}{N_{\text{IKW}}^{1-c+\tau}} \right)$$

where  $\text{dist}(\cdot, \cdot)$  is the normalized hamming distance.

*Proof.* Observe that by construction,  $g_A(x)$  and  $\tilde{g}_A$  only disagree at  $x$  if either

1.  $\pi_V(A) \neq \tilde{\pi}_V(A)$
2.  $\exists S_E : \pi_U(S_E) \neq \tilde{\pi}_U(S_E)$

over the  $S_E \ni x$  sampled in defining  $g_A^{\text{rand}}(x)$ . Furthermore, since  $\text{dist}(g_A^{\text{rand}}, \tilde{g}_A^{\text{rand}}) = \mathbb{E}_{x, \{S_E\}}[\mathbf{1}[g_A^{\text{rand}}(x) \neq \tilde{g}_A^{\text{rand}}(x)]]$ , we may then upper bound the expected disagreement as

$$\mathbf{E} \left[ \text{dist}(g_A^{\text{rand}}, \tilde{g}_A^{\text{rand}}) \right] \leq \Pr_{\mathcal{A}_{\text{IKW}}, A, x, \{S_E\}} \left[ \pi_V(A) \neq \tilde{\pi}_V(A) \vee \exists S_E : \pi_U(S_E) \neq \tilde{\pi}_U(S_E) \right].$$

To bound the righthand side, note that even conditioned on the randomness of  $\mathcal{A}_{\text{IKW}}$ , the distribution of  $A$  and each individual  $S_E$  are marginally uniform over  $U$  and  $V$  respectively (where the marginalization is over  $x$ , a uniformly random vertex of the base 2-CSP). Thus, by a union bound, we finally have

$$\begin{aligned} \mathbf{E} \left[ \text{dist}(g_A^{\text{rand}}, \tilde{g}_A^{\text{rand}}) \right] &\leq \mathbf{E}_{\mathcal{A}_{\text{IKW}}} [\text{dist}(\pi_V, \tilde{\pi}_V)] + O\left(\frac{\log N_{\text{IKW}}}{\varepsilon}\right) \mathbf{E}_{\mathcal{A}_{\text{IKW}}} [\text{dist}(\pi_U, \tilde{\pi}_U)] \\ &\leq \frac{\log(N_{\text{IKW}})2^{O(k)}}{\varepsilon N_{\text{IKW}}^{1-c+\tau}} \end{aligned}$$

as desired.  $\square$

Our second claim is that a random  $g_A^{\text{rand}}$  still gives a highly satisfying assignment to the original CSP with non-trivial probability. Namely, we argue that for any  $(A, a)$  satisfying the properties of Theorem 5.4, a random  $g_A^{\text{rand}}$  is likely to be  $1 - O(\delta)$  satisfying:

**Claim 5.6** (Correctness of  $g_A^{\text{rand}}$ ). If  $\pi = (\pi_U, \pi_V)$  and  $A$  satisfy the properties in Theorem 5.4, then:

$$\Pr \left[ g_A^{\text{rand}} \text{ is } 1 - O(\delta) \text{ satisfying} \right] \geq 0.99$$

where the probability is over  $g_A^{\text{rand}}$ 's internal randomness.

We defer the proof of the claim, which is technical and relies on sampling properties of  $\text{IKW}_k^\Phi$ .

We now complete the proof assuming the above. As discussed, we need a low-sensitivity method for generating a highly satisfying solution. We use the following randomized thresholding trick inspired by similar methods in replicable machine learning (see e.g. [1, 9, 14, 25, 35, 36, 37, 39, 41, 46]). Let  $c'$  be the promised constant such that  $g_A^{\text{rand}}$  is  $1 - c'\delta$  satisfying with 99% probability in Claim 5.6, then:

1. Pick  $m = O\left(\frac{\log \frac{1}{\delta}}{\varepsilon}\right)$  random  $\{g_{A,i}^{\text{rand}}\}_{i \in [m]}$
2. Pick a random threshold  $t \in [c', 2c']$
3. Output the first  $g_{A,i}^{\text{rand}}$  which is at least  $1 - t\delta$  satisfying (if one exists, else output all 0).

We now analyze the correctness and sensitivity of the above algorithm individually:

**Correctness** By Claim 5.6, as long as  $\mathcal{A}_{\text{IKW}}$  indeed produced a  $\varepsilon$ -satisfying solution (which occurs except with probability  $O(\delta)$ ), a random choice of  $g_A^{\text{rand}}$  is  $1 - t\delta$  correct with probability at least  $\Omega(\varepsilon)$ . Thus, conditioned on the above, the probability one of the sampled  $g_A^{\text{rand}}$  is at least  $1 - t\delta$  correct is at least  $1 - O(\delta)$ . The expected value of the output assignment is therefore also at least  $1 - O(\delta)$  as well (since  $t$  is at most some constant  $2c'$ ).

**Sensitivity** Denote the  $m = O(\frac{\log \frac{1}{\delta}}{\varepsilon})$  sampled decoder outputs on the starting instances  $I_\Phi$  and  $\tilde{I}_\Phi$  as  $\{\mathbf{g}_{A,i}^{\text{rand}}\}_{i \in [m]}$  and  $\{\tilde{\mathbf{g}}_{A,i}^{\text{rand}}\}_{i \in [m]}$  respectively. Let  $T_{\text{bad}}$  denote the event that the random threshold  $t$  falls *between* the value of any pair  $(\mathbf{g}_{A,i}^{\text{rand}}, \tilde{\mathbf{g}}_{A,i}^{\text{rand}})$ . We claim the expected hamming distance between our final solutions  $\mathbf{g}_{\text{out}}$  and  $\tilde{\mathbf{g}}_{\text{out}}$  is at most

$$\mathbf{E}[\text{dist}(\mathbf{g}_{\text{out}}, \tilde{\mathbf{g}}_{\text{out}})] \leq \mathbf{E} \left[ T_{\text{bad}} + \sum_{i \in [m]} \text{dist}(\mathbf{g}_{A,i}^{\text{rand}}, \tilde{\mathbf{g}}_{A,i}^{\text{rand}}) \right]$$

where the expectation is over  $\mathcal{A}_{\text{IKW}}$ , the choice of  $(A, a)$ , the choice of randomness for  $\mathbf{g}^{\text{rand}}, \tilde{\mathbf{g}}^{\text{rand}}$ , and the choice of  $t$ . This follows since as long as the event  $T_{\text{bad}}$  does not occur, randomized thresholding always picks  $\mathbf{g}_{\text{out}}$  and  $\tilde{\mathbf{g}}_{\text{out}}$  to be some pair  $(\mathbf{g}_{A,i}^{\text{rand}}, \tilde{\mathbf{g}}_{A,i}^{\text{rand}})$ , in which case we pay the distance between these two functions.

Moreover, since  $t$  is a random independent threshold in  $[c', 2c']$ , the probability the threshold  $1 - t\delta$  lands between any two of our functions is at most the sum of the pairwise distances divided by  $\delta c'$ , so we get

$$\begin{aligned} \mathbf{E}[\text{dist}(\mathbf{g}_{\text{out}}, \tilde{\mathbf{g}}_{\text{out}})] &\leq \frac{2}{\delta c'} \mathbf{E} \left[ \sum_{i \in [m]} \text{dist}(\mathbf{g}_{A,a,i}^{\text{rand}}, \tilde{\mathbf{g}}_{A,a,i}^{\text{rand}}) \right] \\ &\leq \frac{2}{\delta c'} \sum_{i \in [m]} \mathbf{E} \left[ \text{dist}(\mathbf{g}_{A,a,i}^{\text{rand}}, \tilde{\mathbf{g}}_{A,a,i}^{\text{rand}}) \right] \\ &\leq \frac{2^{O(k)} \log \frac{1}{\delta} \log N_{\text{IKW}}}{\delta \varepsilon^2} \cdot \frac{1}{N_{\text{IKW}}^{1-c+\tau}}. \end{aligned}$$

Noting that we have assumed  $k \leq o(\log(N_{\text{IKW}}))$ , the leading term is sub-polynomial in  $n$ , and in particular the entire expression is  $o(\frac{1}{N_{\text{IKW}}^{1-c}})$ . Thus the resulting un-normalized expected hamming distance is  $o(N_{\text{IKW}}^c)$  and we get the desired contradiction.  $\square$

### 5.3 Analysis of the Localized Decoder

We now prove the deferred soundness claim for the localized decoder. This is the sampler-based concentration step invoked in the proof of Theorem 5.2.

*Proof of Claim 5.6.* It is enough to upper bound the *expected* distance of  $\mathbf{g}_A^{\text{rand}}$  from IKW's decoding  $g_A$  as

$$\Pr_{r,x}[\mathbf{g}_A^{\text{rand}}(x) \neq g_A(x)] \leq O(\delta) \tag{1}$$

where  $x$  is distributed uniformly. Markov's inequality then implies that with (say) 99% probability, our random decoding is  $1 - O(\delta)$  close to  $g_A$ , and therefore  $1 - O(\delta)$  satisfying.

Recall our algorithm operates on  $x$  by first sampling  $e \ni x$ , then  $O(\frac{\log N_{\text{IKW}}}{\varepsilon})$  random  $S_E \ni e$ , and outputting the value of  $\pi_U(S_E)$  on a random such  $S_E \in \text{Cons}_{A,a}^x$  (if one exists), and otherwise outputting 0. We can therefore upper bound Equation (1) by the probability of 'losing' the following game:

1. Sample a uniformly random edge  $e$
2. Sample  $\{S_E\} \ni e$  as described

3. If  $\{S_E\}$  has no face in  $\text{Cons}_{A,a}^x$ , lose. Otherwise
  - Pick a random  $T_E \subset \{S_E\}$  that's in  $\text{Cons}_{A,a}^x$ 
    - If  $\pi_U(T_E)|_e \neq g_A|_e$ , lose
    - Else, win.

To analyze this game, we consider its underlying graph structure. Write  $E \setminus A$  to denote the edges in the induced subgraph of  $G$  on  $[N_{\text{IKW}}] \setminus A$ .<sup>11</sup> We are interested in the bipartite graph  $G' = (L, R, E)$  with

$$L = E \setminus A \quad \text{and} \quad R = \bigsqcup_{A_E} \binom{E \setminus A}{k - k'}$$

with edges given by inclusion. In other words, the right-hand side of this graph consists of disjoint copies of  $\binom{E \setminus A}{k - k'}$  corresponding to the choice of  $B_E$  for every choice of  $A_E$  (itself a choice of edges incident to  $A$ ), and thus are in one-to-one correspondence with choices of  $S_E = (A_E, B_E)$ , i.e. the neighborhood of  $A$  in the IKW graph.<sup>12</sup>

Now, call an edge in this graph corresponding to  $(e, S_E)$  ‘good’ if  $\pi_U(S_E)|_e = g_A(e)$  (and ‘bad’ otherwise), and write  $\mu = \frac{|\text{Cons}_{A,a}|}{|R|}$  to be the relative size of  $\text{Cons}$  (recall we are promised  $\mu \geq \Omega(\varepsilon)$ ). We will prove the following two facts which together bound the probability of losing this game by  $O(\delta)$ :

1. At most an  $O(\frac{1}{\sqrt{k}})$ -fraction of  $e \in L$  have fewer than  $\frac{\mu}{2}$  or more than  $2\mu$  of their edges into  $\text{Cons}_{A,a}$
2. At most an  $O(\delta)$  fraction of edges  $(e, S_E)$  in the induced subgraph on  $\text{Cons}_{A,a}$  are bad

We first explain why these two facts imply the bound. Assume that in the first step of the game we sample  $e$  that avoids the bad event in Fact 1 above. Then  $e$  has at least an  $\varepsilon/2$ -fraction of its neighbors in  $\text{Cons}_{A,a}$ , and the probability that we completely miss  $\text{Cons}_{A,a}$  when sampling the  $\{S_E\}$  is vanishingly small (at most  $1/\text{poly}(N_{\text{IKW}})$ ), and may therefore be ignored. In particular, this means that the output of this process is  $1/\text{poly}(N_{\text{IKW}})$ -close in total variation distance to drawing a random ‘good’  $e$ , then a random  $S_E$  conditioned on landing in  $\text{Cons}_{A,a}$ .

Finally, combining Facts 1 and 2, we know the total fraction of bad edges in the induced subgraph between  $\text{Cons}_{A,a}$  and the ‘good’  $e$  is at most  $O(\delta)$ . Moreover, since all the degrees of good  $e$  in the induced subgraph are within a constant factor, the distribution over edges in our game generated by picking a uniformly random good  $e$ , then a random edge in the induced subgraph is close to uniform in the sense that no edge has more than a constant factor times its measure in the uniform distribution (since the degree of all good  $e$  are within a constant factor of each other). As a result, the probability we hit a bad edge (of which there are at most  $O(\delta)$  by Fact 2) and lose is at most  $O(\delta)$ . Finally, adding back the probability we lose from sampling a ‘bad’  $e$  in the first step gives a total failure probability of  $O(\delta) + O(\frac{1}{\sqrt{k}}) \leq O(\delta)$  as desired

It is left to prove the two claimed facts. The second fact is immediate from Property 3 of Theorem 5.4, which states that almost all  $S_E$  in  $\text{Cons}$  have at most an  $O(\delta)$  fraction of bad edges. The first follows from the fact that our described graph  $G'$  is an excellent *sampler*. A regular bipartite graph  $G = (L, R, E)$  is called a (multiplicative)  $(\mu, \beta, \delta)$ -sampler if for every subset  $T \subset L$  of density at least  $\mu$ :

$$\Pr_{v \in R} \left[ \left| \mathbf{E}_{w \in L} [1_T(w) \mid w \sim v] - \mathbf{E}_{w \in L} [1_T] \right| > \delta \mathbf{E}_{w \in L} [1_T] \right] \leq \beta.$$

<sup>11</sup>We can ignore the case where  $x \in A$ , as by definition we always match  $g_A$  on these vertices, so it is enough to work on  $E \setminus A$ .

<sup>12</sup>Technically, we should restrict in each component of  $R$  to choices in  $E \setminus A$  that do not intersect  $A_E$ . However the fraction of such nodes is  $O(k/N_{\text{IKW}})$ , so this may be safely ignored in the analysis.

In other words  $G$  is a sampler the neighborhood of a random  $v \in R$  almost always sees the set  $T$  in roughly the correct proportion.

We first claim it is easy to see  $G'$  is an  $(\mu, e^{-\Omega(\delta^2 \mu k)}, \delta)$ -sampler for  $\delta \in (0, 1)$ . This is essentially immediate from the fact that  $G'$  is isomorphic to (several stacked copies of) the inclusion graph of  $\left([m], \binom{[m]}{k-k'}\right)$  for the appropriate choice of  $m$ . Sampling on this latter object is exactly Chernoff without replacement (we have a ‘marked’ subset  $T \subset [m]$ , sample  $t$  random elements of  $[m]$  without replacement, and ask what fraction of these  $t$  elements lie in  $T$ ). Thus standard Chernoff-Hoeffding bounds imply the stated bounds. Finally, notice that ‘stacking’ these graphs in the sense that we have many copies of the right-hand side with edges to the same  $[m]$  maintains the sampling guarantee.

Now, the above is not quite what we want. In particular, it promises that if we have any set of size roughly  $\mu$  on  $L$  is ‘seen’ in the correct fraction by  $R$  with extremely high probability  $e^{-\Omega(\mu k)}$ . However, the guarantee we need to show is the reverse of this, that all but an  $O(\frac{1}{\sqrt{k}})$  fraction of  $L$  sees Cons in roughly the right proportion (i.e. setting  $\delta = 1/2$  is sufficient), where Cons is a set of measure at least  $\exp(-O(\sqrt{k}))$ . Thankfully, it is a standard fact that whenever a graph  $G' = (L, R, E)$  is an  $(\delta, \beta, \mu)$ -sampler, the flipped graph  $G'_{\text{flip}} = (R, L, E)$  is an  $(O(\beta), O(\mu), O(\delta))$  sampler (see, e.g., [38]). In particular, appropriately setting  $\mu = O(\frac{1}{\sqrt{k}})$  in the original bound, we get that the flip graph is a  $(\mu(\text{Cons}_A), O(1/\sqrt{k}), 1/2)$ -sampler which immediately implies the first fact and completes the proof.  $\square$

## Part II

# Low-Soundness Label Cover via Dinur–Harsha

In this chapter we prove our sensitivity lower bound for low-soundness label cover based on the composition theorem of Dinur and Harsha [20], which we restate next.

**Theorem 1.6.** There exist absolute constants  $B, \delta > 0$  such that for every function  $1 < g(n) \leq (\log n)^B$  and all sufficiently large  $n$ , any algorithm that, given a satisfiable left-predicate label cover instance with  $n = |U|$ -many left vertices, outputs an assignment that satisfies at least a  $1/g(n)$  fraction of the constraints must have sensitivity at least  $\Omega(n^\delta)$ . Moreover, the hard instances can be chosen so that  $|U| = n$ ,  $|V| = n \cdot (\log n)^{O(1)}$ , the left and right degrees are  $g(n)^{O(1)}$ ,  $|\Sigma_V| = g(n)^{O(1)}$ , and  $|\Sigma_U| \leq \exp(g(n)^{O(1)})$ .

This section is organized as follows. First, in Section 6 we introduce a notion of sensitivity-preserving reduction similar to the one used in Section 5.2; the key differences are that we now only track sensitivity, rather than sensitive regions, and we allow the recovery map to depend on the source instance. In Section 8, from the input family of CNF formulas, we construct the label cover instance that will be the basis for the ‘stage-0’ instance of our iterative construction of the final label cover instance that is done in Section 11. Each stage of this iterative construction proceeds by the composition procedure adapted from Dinur–Harsha [20], which we describe in Section 10, followed by a right-degree reduction which we design in Section 7, and an right-alphabet reduction procedure given in Section 9. The stage-0 instance is constructed, these procedures are put together, and the final iterative construction is done in Section 11.

## 6 Sensitivity Reduction with Source-Dependent Recovery

In Part I we introduced a left/right budget-valued framework for transferring sensitivity lower bounds through reductions. For the Dinur–Harsha route we only need its scalar analogue: we collapse the pair  $(k_L, k_R)$  to total Hamming sensitivity, and we additionally allow the recovery map to depend on the source instance. This contributes an additive term  $D$  when the source instance changes but the transformed assignment is held fixed. The resulting scalar sensitivity-reduction interface is parameterized by three constants:

- $C_{\mathcal{T}}$ : A transformation constant which bounds how far the transformed instance can move under one source swap,
- $C_{\mathcal{R}}$ : A fixed-source Lipschitz constant for the recovery map, and
- $D$ : An additive source-swap term, measuring how much the recovery map itself can change when the source instance changes but the transformed assignment is held fixed.

We record their effects in the following definition.

**Definition 6.1** (Sensitivity-Preserving Reduction). Let  $\mathcal{I}$  be a family of ordinary label cover instances or left-predicate label cover instances on the same underlying graph and domain. Let  $\mathcal{T}$  be a procedure that transforms an instance  $I \in \mathcal{I}$  into another instance  $I'$ , and for each  $I \in \mathcal{I}$  let  $\mathcal{R}_I$  be a (possibly randomized) procedure that transforms an assignment  $\pi'$  for  $I'$  into an assignment  $\pi$  for  $I$ . For  $s, s' \in [0, 1]$  and  $C_{\mathcal{T}}, C_{\mathcal{R}}, D \geq 0$ , we say that  $(\mathcal{T}, \{\mathcal{R}_I\}_{I \in \mathcal{I}})$  is an  $(s, s', C_{\mathcal{T}}, C_{\mathcal{R}}, D)$ -sensitivity-preserving reduction for  $\mathcal{I}$  if the following hold:

1. If  $I$  is satisfiable, then  $I' = \mathcal{T}(I)$  is also satisfiable.
2. If a (possibly random) assignment  $\pi'$  for  $I'$  satisfies  $\mathbf{E}[\text{val}_{I'}(\pi')] \geq s'$ , then the assignment  $\pi = \mathcal{R}_I(\pi')$  satisfies  $\mathbf{E}[\text{val}_I(\pi)] \geq s$ .
3. The transformed instances  $\mathcal{T}(I)$ , for  $I \in \mathcal{I}$ , all live on one common underlying graph and common alphabets, and for every source pair  $I, \tilde{I} \in \mathcal{I}$  with  $\text{SwapDist}(I, \tilde{I}) = 1$  we have

$$\text{SwapDist}(\mathcal{T}(I), \mathcal{T}(\tilde{I})) \leq C_{\mathcal{T}}.$$

4. For every fixed  $I \in \mathcal{I}$  and every two assignments  $\pi', \tilde{\pi}'$  for  $\mathcal{T}(I)$ ,

$$\text{EMD}(\mathcal{R}_I(\pi'), \mathcal{R}_I(\tilde{\pi}')) \leq C_{\mathcal{R}} d_{\text{H}}(\pi', \tilde{\pi}').$$

5. For every source pair  $I, \tilde{I} \in \mathcal{I}$  with  $\text{SwapDist}(I, \tilde{I}) = 1$  and every assignment  $\pi'$  on the common domain of  $\mathcal{T}(I)$  and  $\mathcal{T}(\tilde{I})$ ,

$$\text{EMD}(\mathcal{R}_I(\pi'), \mathcal{R}_{\tilde{I}}(\pi')) \leq D.$$

**Lemma 6.2.** Let  $\mathcal{I}$  be a swap-closed family of instances on the same underlying graph and domain. Suppose there exists an  $(s, s', C_{\mathcal{T}}, C_{\mathcal{R}}, D)$ -sensitivity-preserving reduction  $(\mathcal{T}, \{\mathcal{R}_I\}_{I \in \mathcal{I}})$  for  $\mathcal{I}$ . Write  $\mathcal{I}' = \mathcal{T}(\mathcal{I}) := \{\mathcal{T}(I) : I \in \mathcal{I}\}$ . Then every algorithm  $A'$  on  $\text{SwapClo}(\mathcal{I}')$  that achieves soundness  $s'$  on its satisfiable members induces an algorithm  $A$  on  $\mathcal{I}$  that achieves soundness  $s$  and satisfies

$$\text{SwapSens}(A, \mathcal{I}) \leq C_{\mathcal{T}} C_{\mathcal{R}} \text{SwapSens}(A', \text{SwapClo}(\mathcal{I}')) + D.$$

Consequently,

$$\text{SwapSens}_{s'}(\text{SwapClo}(\mathcal{I}')) \geq \frac{\text{SwapSens}_s(\mathcal{I}) - D}{C_{\mathcal{T}}C_{\mathcal{R}}}.$$

*Proof.* Let  $A'$  be any algorithm on  $\text{SwapClo}(\mathcal{I}')$  that achieves soundness  $s'$  on its satisfiable members. Define an algorithm  $A$  on  $\mathcal{I}$  as follows: on input  $I \in \mathcal{I}$ , let  $I' := \mathcal{T}(I)$ , run  $A'$  on  $I'$  to obtain a (possibly random) assignment  $\pi'$  for  $I'$ , and output  $\pi \sim \mathcal{R}_I(\pi')$ .

If  $I$  is satisfiable, then Item 1 of Definition 6.1 ensures that  $I'$  is satisfiable as well. Since  $I' \in \mathcal{I}' \subseteq \text{SwapClo}(\mathcal{I}')$ , the assumption on  $A'$  yields  $\mathbf{E}[\text{val}_{I'}(\pi')] \geq s'$ , and Item 2 implies  $\mathbf{E}[\text{val}_I(\pi)] \geq s$ .

Let  $S := \text{SwapSens}(A', \text{SwapClo}(\mathcal{I}'))$ . Fix any source pair  $I, \tilde{I} \in \mathcal{I}$  with  $\text{SwapDist}(I, \tilde{I}) = 1$ . Write  $I' := \mathcal{T}(I)$  and  $\tilde{I}' := \mathcal{T}(\tilde{I})$ . By the same swap-path argument used in the proof of Lemma 3.2, Item 3 implies

$$\text{EMD}(A'(I'), A'(\tilde{I}')) \leq C_{\mathcal{T}}S.$$

We decompose the perturbation as

$$\text{EMD}(A(I), A(\tilde{I})) \leq \text{EMD}(\mathcal{R}_I(A'(I')), \mathcal{R}_I(A'(\tilde{I}'))) + \text{EMD}(\mathcal{R}_I(A'(\tilde{I}')), \mathcal{R}_{\tilde{I}}(A'(\tilde{I}'))).$$

Couple  $A'(I')$  and  $A'(\tilde{I}')$  optimally and apply Item 4 pointwise under that coupling to bound the first term by

$$C_{\mathcal{R}} \text{EMD}(A'(I'), A'(\tilde{I}')) \leq C_{\mathcal{T}}C_{\mathcal{R}}S.$$

For the second term, sample  $\pi' \sim A'(\tilde{I}')$  and apply Item 5 pointwise to obtain the bound  $D$ . Thus,

$$\text{EMD}(A(I), A(\tilde{I})) \leq C_{\mathcal{T}}C_{\mathcal{R}}S + D.$$

Since the source one-swap pair  $(I, \tilde{I})$  was arbitrary, this proves

$$\text{SwapSens}(A, \mathcal{I}) \leq C_{\mathcal{T}}C_{\mathcal{R}} \text{SwapSens}(A', \text{SwapClo}(\mathcal{I}')) + D.$$

Minimizing the left-hand side over all algorithms  $A$  of this form and rearranging gives the final inequality.  $\square$

## 7 Right-Degree Reduction in the Scalar Setting

We use the same right-degree reduction gadget as in Section 4. Concretely, we use the same transformation  $\mathcal{T}_{\text{DR}}$ , detailed in Algorithm 1, the same recovery map  $\mathcal{R}_{\text{DR}}$ , and the same fixed auxiliary choices of expanders and neighbor orderings across a common family. Here we record only the scalar consequence of that gadget for the sensitivity-reduction framework above.

**Lemma 7.1.** Let  $d \geq 4$  be even and  $\lambda \in [0, 1)$ . Let  $\mathcal{I}$  be a swap-closed family of left-predicate label cover instances on a common underlying graph and common alphabets, whose minimum right degree is at least  $\delta_V \geq d$ . For each right vertex  $v$  of this common underlying graph, let  $D_v$  denote its common degree across the family, and fix a  $[D_v, d, \lambda]$ -expander  $H_v$ . Let  $\mathcal{I}' = \text{SwapClo}(\mathcal{T}_{\text{DR}}(\mathcal{I}; d))$ . Then  $\mathcal{I}'$  consists of right-regular left-predicate label cover instances with right degree  $d$ , and for any  $\varepsilon > 0$ , we have

$$\text{SwapSens}_{\varepsilon+O(\lambda)}(\mathcal{I}') \geq \frac{1}{d} \text{SwapSens}_{\varepsilon}(\mathcal{I}).$$

*Proof.* The structural statement that  $\mathcal{I}'$  consists of right-regular left-predicate label cover instances with right degree  $d$  follows directly from the construction in Section 4.

We verify the hypotheses of Lemma 6.2 with

$$C_{\mathcal{T}} = d, \quad C_{\mathcal{R}} = 1, \quad D = 0.$$

We use exactly the same transformation  $\mathcal{T}_{\text{DR}}$  and recovery map  $\mathcal{R}_{\text{DR}}$  as in Lemma 4.3.

Completeness, soundness, and the bound  $C_{\mathcal{T}} = d$  for how the transformed instance changes under one source swap are verified exactly as in the proof of Lemma 4.3. In particular, the soundness loss in passing from  $\mathcal{I}$  to  $\mathcal{I}'$  is the same  $O(\lambda) = O(d^{-1/2})$  loss used there.

For Item 4, use the same shared-index coupling as in the proof of Lemma 4.3. That proof shows, coordinate-wise, that the recovered left labels differ in expectation by at most  $d_{\text{H}}(\pi'_{\mathcal{U}}, \tilde{\pi}'_{\mathcal{U}})$  and the recovered right labels differ in expectation by at most  $\delta_V^{-1} d_{\text{H}}(\pi'_{\mathcal{V}}, \tilde{\pi}'_{\mathcal{V}})$ . Summing the two coordinates and using  $\delta_V \geq 1$  gives

$$\text{EMD}(\mathcal{R}_{\text{DR}}(\pi'), \mathcal{R}_{\text{DR}}(\tilde{\pi}')) \leq d_{\text{H}}(\pi'_{\mathcal{U}}, \tilde{\pi}'_{\mathcal{U}}) + \frac{1}{\delta_V} d_{\text{H}}(\pi'_{\mathcal{V}}, \tilde{\pi}'_{\mathcal{V}}) \leq d_{\text{H}}(\pi', \tilde{\pi}').$$

Thus  $C_{\mathcal{R}} = 1$ .

For Item 5, fix a source one-swap pair  $I, \tilde{I} \in \mathcal{I}$  and an assignment  $\pi'$  on the common target domain. Because the expanders  $H_v$  and the neighbor orderings are fixed across the family, the recovery map  $\mathcal{R}_{\text{DR}}$  depends only on  $\pi'$  and these auxiliary choices, not on the source predicates or source edge projections. Thus

$$\mathcal{R}_{\text{DR}, I}(\pi') = \mathcal{R}_{\text{DR}, \tilde{I}}(\pi')$$

as distributions, and so the drift term is  $D = 0$ .

Applying Lemma 6.2 now yields

$$\text{SwapSens}_{\varepsilon+O(\lambda)}(\mathcal{I}') \geq \frac{\text{SwapSens}_{\varepsilon}(\mathcal{I})}{d},$$

as claimed. □

## 8 Label Cover under Clause Swaps and Stable Recovery

For each CNF formula  $\Phi$ , we construct a left-predicate label cover instance  $I_{\Phi}$  with low soundness and explicit control over how  $I_{\Phi}$  changes when clauses of the underlying CNF are swapped; this is the construction used later to initialize the stage-0 family in Section 11. For the satisfiable CNF family used in that stage-0 initialization, we also prove a recovery theorem with perturbation guarantees, showing how to recover a satisfying assignment of  $\Phi$  from a labeling of  $I_{\Phi}$  with noticeable value. The first subsection constructs the instances  $I_{\Phi}$ , and the second proves the recovery statement needed for stage 0.

### 8.1 Left-Predicate Label Cover under Clause Swaps

We begin with the left-predicate label cover, focusing on how the resulting label cover instance changes under clause swaps in the original CNF. The obstacle is that the standard algebraic robust-PCP template hardwires the clause predicate into one global low-degree polynomial. That is exactly what makes the classical soundness argument work, but it is incompatible with locality: a single clause swap is highly local in the source CNF's table of clause indicators, but typically becomes global after low-degree extension, ruining sensitivity.

Our construction separates the algebraic soundness task from the formula lookup task. The outer algebraic layer certifies only the low-degree structure of the assignment and witness polynomials, while a separate local Reed–Muller evaluation gadget supplies the relevant value of the source CNF’s clause-indicator table at the distinguished point. This ensures that the formula-dependent part of the verifier stays inside the left predicates, so local changes in the source formula appear as local predicate swaps while the underlying graph and edge projections remain fixed. The result is a left-predicate label cover instance that serves as the stage-0 family for the later construction.

The main result of the section is the following left-predicate label cover under clause swaps for sufficiently large  $n$ .

**Theorem 8.1** (Left-predicate label cover under clause swaps). Fix any constant  $c > 0$ . There exist  $n_0 = n_0(c)$  and an absolute constant  $c_{lc} \geq 1$  such that the following holds for every integer  $n \geq n_0$ . Let  $\Phi$  be any CNF formula on  $n$  variables, and let  $a := 4c + 7$ ,  $m := \frac{\log n}{\log \log n}$ ,  $q := \Theta((\log n)^a)$ , and  $\varepsilon \geq \log^{-c} n$ . Then there is an explicit reduction  $\mathcal{T}_{\text{Robust}}$  that maps  $\Phi$  to a left-predicate label cover instance  $I_\Phi = (U, V, E, \Sigma_U, \Sigma_V, \{P_u^\Phi\}_{u \in U}, F = \{f_e\}_{e \in E})$  with the following properties.

1. **Completeness.** If  $\Phi$  is satisfiable, then  $\text{val}(I_\Phi) = 1$ .
2. **Soundness.** If  $\Phi$  is unsatisfiable, then  $\text{val}(I_\Phi) \leq \varepsilon$ .
3. **Change under clause swaps.** If  $\Phi$  and  $\Phi'$  differ by one clause swap, then  $I_\Phi$  and  $I_{\Phi'}$  have the same underlying graph, the same left and right alphabets, and the same edge projections, and only  $2q^{8m+c_{lc}}$  left predicates change.
4. **Parameters.** Every instance  $I_\Phi$  satisfies

$$|V| \leq q^{3m+c_{lc}}, \quad q^{11m-c_{lc}} \leq |U| \leq q^{11m+c_{lc}}, \quad |\Sigma_V| = q^{3m+5}, \quad |\Sigma_U| \leq q^{q^{c_{lc}}},$$

$$d_u \leq q^{c_{lc}} \text{ for every } u \in U, \quad q^{8m+6} \leq d_v \leq q^{8m+c_{lc}} \text{ for every } v \in V,$$

We now begin the proof of this theorem; the construction of the label-cover instance proceeds in several stages.

**Preprocessing.** Fix a CNF formula  $\Phi$ . Choose a field  $\mathbb{F}$  of size  $q := |\mathbb{F}| = \Theta((\log n)^a)$  and a subset  $H \subseteq \mathbb{F}$  with  $\{0, 1\} \subseteq H$  and  $|H| = \Theta(\log n)$ . After the usual padding, we identify the variable set with  $H^m$ , where  $m = \log n / \log \log n$ .

An assignment  $a \in \{0, 1\}^n$  is now viewed as a function  $a : H^m \rightarrow \{0, 1\}$ , and we let  $A : \mathbb{F}^m \rightarrow \mathbb{F}$  be its low-degree extension. Its degree is at most  $D_A := m(|H| - 1)$ . Define the *raw clause table*  $B_\Phi : H^{3m+3} \rightarrow \{0, 1\}$  by

$$B_\Phi(u, v, w, b_1, b_2, b_3) = 1 \iff x_u^{b_1} \vee x_v^{b_2} \vee x_w^{b_3} \text{ is a clause of } \Phi.$$

for  $b_1, b_2, b_3 \in \{0, 1\}$ , and 0 otherwise. In other words,  $B_\Phi$  is supported on clause-encoding points in  $H^{3m} \times \{0, 1\}^3$ , and one clause swap changes  $B_\Phi$  at exactly two points. Let  $\tilde{B}_\Phi : \mathbb{F}^{3m+3} \rightarrow \mathbb{F}$  be the low-degree extension of  $B_\Phi$ ; its degree is at most  $D_B := (3m+3)(|H| - 1)$ . We use the cyclic permutation  $\rho(i, j, k, b_1, b_2, b_3) := (j, k, i, b_1, b_2, b_3)$ . Thus for a clause point  $x = (u, v, w, b_1, b_2, b_3)$ , the first  $m$  coordinates of  $x, \rho(x), \rho^2(x)$  are  $u, v, w$  respectively, matching the literal order in the clause-check polynomial below. Fix a common degree bound  $D := 8m|H|$ , which dominates all outer polynomials appearing below.

Given a polynomial  $A : \mathbb{F}^m \rightarrow \mathbb{F}$ , define its clause-check polynomial

$$P_{\Phi, A}(x) := \tilde{B}_{\Phi}(x)(A(x) - x_{3m+1})(A(\rho(x)) - x_{3m+2})(A(\rho^2(x)) - x_{3m+3}),$$

where, in expressions such as  $A(x)$ ,  $A(\rho(x))$ , and  $A(\rho^2(x))$ , the polynomial  $A$  is evaluated on the first  $m$  coordinates of the corresponding point of  $\mathbb{F}^{3m+3}$ . Since  $\tilde{B}_{\Phi}$  agrees with  $B_{\Phi}$  on  $H^{3m+3}$ , the restriction  $A|_{H^m}$  satisfies  $\Phi$  if and only if  $P_{\Phi, A}$  vanishes on  $H^{3m+3}$ .

We use the standard algebraic characterization of vanishing on  $H^m$ .

**Lemma 8.2.** For every field  $\mathbb{F}$ , subset  $H \subseteq \mathbb{F}$ , and integer  $r \geq 1$ , a polynomial  $P : \mathbb{F}^r \rightarrow \mathbb{F}$  of degree at most  $D$  vanishes on  $H^r$  if and only if there exist polynomials  $P_1, \dots, P_r$  of degree at most  $D - |H|$  such that  $P(x) = \sum_{j \in [r]} g_H(x_j) P_j(x)$ , where  $g_H(t) := \prod_{h \in H} (t - h)$ .

**Objects and reference charts.** An *object* is a tuple  $\Omega = (y, y', z, z')$  with  $y, y', z, z' \in \mathbb{F}^{3m+3}$  such that:

1.  $z_i = z'_i$  for all  $i \in [m]$ , and
2. the six points  $z, y, y', z', \rho(z), \rho^2(z)$  are affinely independent.

Let  $\mathcal{O}$  denote the set of all objects. For  $\Omega = (y, y', z, z') \in \mathcal{O}$  we write  $L(\Omega) := \text{aff}(z, y, y', z', \rho(z), \rho^2(z))$ , which is now always a 5-dimensional affine subspace. For any function  $Q$  we denote by  $Q|_{L(\Omega)}$  the restriction of  $Q$ 's domain to the points in  $L(\Omega)$ . We also fix the reference affine chart  $\psi_{\Omega} : \mathbb{F}^5 \rightarrow L(\Omega)$  defined by

$$\psi_{\Omega}(0) = z, \quad \psi_{\Omega}(e_1) = y, \quad \psi_{\Omega}(e_2) = y', \quad \psi_{\Omega}(e_3) = z', \quad \psi_{\Omega}(e_4) = \rho(z), \quad \psi_{\Omega}(e_5) = \rho^2(z),$$

where  $e_1, \dots, e_5$  are the standard basis vectors of  $\mathbb{F}^5$ .

Equivalently, start from the product distribution in which  $z, y, y'$  are uniform in  $\mathbb{F}^{3m+3}$  and  $z'$  is uniform among points sharing the first  $m$  coordinates with  $z$ , and then condition on the affine-independence event in item (2); this conditioned distribution is exactly the uniform distribution on  $\mathcal{O}$ . There is an absolute constant  $C_{\text{geom}} \geq 1$  such that this conditioning changes each of the marginals of  $z$ , of  $y$ , and of  $(z, z')$  by total-variation distance at most  $C_{\text{geom}}/q$  from the corresponding unconditioned uniform distribution on  $\mathbb{F}^{3m+3}$  or on the pairs sharing their first  $m$  coordinates. This is exactly the geometric regime used in the Dinur–Harsha [20] object analysis.

Two standard ingredients enter next. The *point-object theorem* converts agreement among local object restrictions into a short list of global low-degree candidates. The *local evaluation theorem* gives local access to  $\tilde{B}_{\Phi}(z)$  from the raw clause-indicator table while keeping the effect of clause swaps confined to a controlled set of left predicates. We use both only as black boxes.

For two functions  $P, Q$  over the same domains, we denote by  $P \equiv Q$  their point-wise equivalence. An object oracle should be viewed as a table of local algebraic hypotheses: for each pair  $(\Omega, \tau)$  it returns a  $k$ -tuple of degree- $D$  polynomials on the 5-flat  $L(\Omega)$ . Different objects are not assumed to be globally consistent a priori. The point-object theorem says that if, for many random triples  $(\Omega, \tau, x)$  with  $x \in L(\Omega)$ , this local tuple agrees with an ambient function  $f$  at the sampled point  $x$ , then except for a small error the entire local tuple must already equal, on all of  $L(\Omega)$ , the restriction of one of a short list of global degree- $D$  tuples. The extra seed  $\tau$  is only an external index for which local view is exposed; it does not change the distribution of the sampled object  $\Omega$  or of the sampled point  $x \in L(\Omega)$ .

**Theorem 8.3** (Point-object theorem with an external seed). There exist absolute constants  $c_{\text{PO}}, C_{\text{PO}}, C_{\text{err}} \geq 1$  such that the following holds. Let  $k \geq 1$ , let  $\mathsf{T}$  be any finite set, let  $f : \mathbb{F}^{3m+3} \rightarrow \mathbb{F}^k$  be any function, and

assume  $\delta \geq c_{\text{PO}} m \left(\frac{mD}{q}\right)^{1/4}$ . Then there exist  $t \leq \frac{C_{\text{PO}}}{\delta}$  polynomial maps  $\vec{Q}_1, \dots, \vec{Q}_t : \mathbb{F}^{3m+3} \rightarrow \mathbb{F}^k$  of coordinate-wise degree at most  $D$  such that, for every object oracle

$$O_{\text{obj}} : \mathcal{O} \times \mathbb{T} \rightarrow \{k\text{-tuples of degree-}D \text{ polynomials on } L(\Omega)\},$$

we have

$$\Pr_{\Omega \sim \mathcal{O}, \tau \sim \mathbb{T}, x \sim L(\Omega)} \left[ O_{\text{obj}}(\Omega, \tau)(x) \neq f(x) \vee \exists i \in [t], \vec{Q}_i|_{L(\Omega)} \equiv O_{\text{obj}}(\Omega, \tau) \right] \geq 1 - \delta - C_{\text{err}} \frac{1+tD}{q}.$$

*Proof.* This is the standard point-object theorem from Dinur–Harsha [20] and Harsha’s notes on the topic [32]. We simply fold the uniform seed  $\tau \in \mathbb{T}$  into the oracle’s internal randomness. This does not change the distribution of the sampled object  $\Omega$  or of the sampled point  $x \in L(\Omega)$ , so the same point-object analysis applies. We introduce the extra seed only for the later local-evaluation gadget: there the left vertices will be pairs  $(\Omega, \tau)$ , with  $\tau$  indexing the local decode/reject test attached to the distinguished point of  $\Omega$ .  $\square$

The next ingredient is a standard Reed–Muller decodable-reader gadget. We use it in a form specialized to the raw clause-indicator table  $B$ : this is a convenient repackaging of the local decode/reject procedure implicit in the Reed–Muller-based dPCP and zero-on-subcube constructions of Dinur–Harsha [20, Section 6.1.2 and Appendix A], Harsha’s notes [32, Lecture 10, Section 10.2.3], and the closely related modern presentation in [7, Section 7.3 and Appendix B.3].

Morally, this gadget supplies a locally checkable certificate that lets us recover the value  $\tilde{B}(z)$  of the low-degree extension from the raw table  $B$  without ever writing  $\tilde{B}$  itself into the label-cover instance. The auxiliary word  $W_{\text{in}}(B)$  is that certificate, its coordinate index set  $V_{\text{in}}$  will later become the family of inner right vertices, and for each pair  $(z, \tau)$  the reader  $\text{Eval}_{z, \tau}$  chooses one local decoding/checking path for the distinguished point  $z$ .

Given a table  $B : H^{3m+3} \rightarrow \mathbb{F}$  and a pair  $(z, \tau)$ , the *local view* of  $\text{Eval}_{z, \tau}$  consists of two ordered answer strings: the *raw-answer string*, formed by the answers to its predetermined raw-table queries in  $B$ , and the *auxiliary-answer string*, formed by the answers to its queries in  $W_{\text{in}}(B)$ .

**Theorem 8.4** (Uniform local evaluation of the raw clause-indicator table). There exist an absolute constant  $c_{\text{in}} \geq 1$  and an explicit construction, depending only on  $(m, q, H)$ , with the following properties.

1. It outputs a universal index set  $V_{\text{in}}$  for the coordinates of the auxiliary word with  $|V_{\text{in}}| = q^{3m+O(1)}$ .
2. It outputs a seed set  $\mathbb{T}$  of size  $q^{O(1)}$ .
3. For each table  $B : H^{3m+3} \rightarrow \mathbb{F}$  it outputs an auxiliary word  $W_{\text{in}}(B) \in \mathbb{F}^{V_{\text{in}}}$ .
4. For each point  $z \in \mathbb{F}^{3m+3}$  and seed  $\tau \in \mathbb{T}$ , it specifies a local decode/reject test  $\text{Eval}_{z, \tau}$  that reads at most  $Q_{\text{in}} := q^{c_{\text{in}}}$  symbols from the auxiliary word  $W_{\text{in}}(B)$  together with the values of  $B$  at a predetermined set of at most  $Q_{\text{in}}$  entries of the raw clause-indicator table, and outputs either reject or a decoded value in  $\mathbb{F}$ .

These tests satisfy:

- **Completeness.** Every local view of  $\text{Eval}_{z, \tau}$  induced by the pair  $(W_{\text{in}}(B), B)$  is accepted.
- **Local correctness.** Every accepted local view of  $\text{Eval}_{z, \tau}$  has a well-defined decoded value, and that value is  $\tilde{B}(z)$ .
- **Bounded locality.** Every raw table point of  $H^{3m+3}$  participates in at most  $q^{c_{\text{in}}}$  tests  $\text{Eval}_{z, \tau}$ , and every coordinate of  $V_{\text{in}}$  participates in at least one and at most  $q^{c_{\text{in}}}$  such tests.

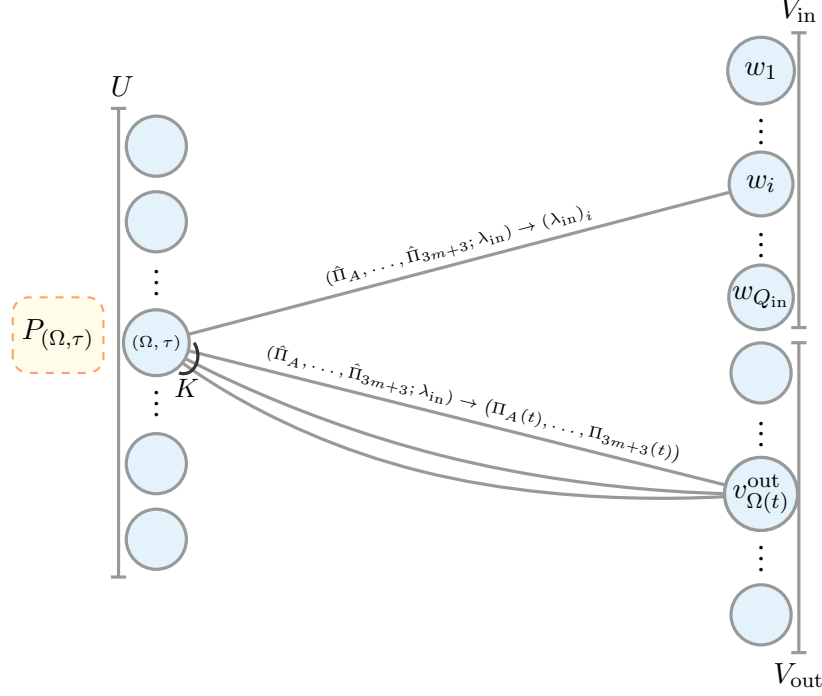


Figure 2: The stage-0 left-predicate label cover instance.

In the label-cover instance, a left vertex  $u = (\Omega, \tau)$  will carry two kinds of information: the restriction of the outer algebraic objects to the object  $\Omega$ , and the auxiliary-answer string from one accepted local view certifying the local evaluation at the distinguished point  $z$  of  $\Omega$ . The inner edges simply project the slots of that auxiliary-answer string to the corresponding coordinates of  $V_{\text{in}}$ , so the auxiliary coordinates are exactly the right-vertex locations used to distribute this local certificate across the label-cover instance.

The label-cover encoding stores only auxiliary-answer strings, not the full local views. This is enough because for fixed  $(z, \tau)$  the raw-answer string is determined by the predetermined raw-table queries and will be checked directly against  $B_\Phi$  inside the left predicate, while the coordinates in  $V_{\text{in}}$  will later serve as the inner right vertices.

To keep the auxiliary-answer string length uniform, fix for each pair  $(z, \tau)$  the ordered auxiliary-answer string of  $\text{Eval}_{z, \tau}$ : the genuine auxiliary queries appear in their natural order, and if fewer than  $Q_{\text{in}}$  such queries are used, we pad by repeating auxiliary query positions. Thus every local view determines a padded auxiliary-answer string in  $\mathbb{F}^{Q_{\text{in}}}$ , and the same auxiliary coordinate may occur in several transcript slots.

**The instance  $I_\Phi$ .** We now define the left-predicate label cover instance  $I_\Phi$ .

*Right vertices.* Let

$$V_{\text{in}}^{\text{qry}} := \{w \in V_{\text{in}} : \exists (\Omega, \tau) \in \mathcal{O} \times \mathbb{T} \text{ such that } w \text{ is queried by } \text{Eval}_{z, \tau}\},$$

where  $z$  denotes the distinguished point of  $\Omega = (y, y', z', z)$ . Discarding never-queried auxiliary coordinates does not affect completeness or soundness, and the locality bounds from [Theorem 8.4](#) remain valid on the surviving coordinates. Henceforth we relabel  $V_{\text{in}}^{\text{qry}}$  as  $V_{\text{in}}$ . By abuse of notation, we use  $V_{\text{in}}$  both for this surviving set of auxiliary-coordinate indices and for the corresponding family of “inner” right vertices in the below description of the “two-part” right vertex set  $V$ :

- For every point  $x \in \mathbb{F}^{3m+3}$  there is an *outer* right vertex  $v_x^{\text{out}}$ .
- For every auxiliary coordinate  $w \in V_{\text{in}}$  there is an *inner* right vertex  $v_w^{\text{in}}$ .

Thus  $V := V_{\text{out}} \sqcup V_{\text{in}}$  and  $|V| = q^{3m+O(1)}$ . The right alphabet is the common alphabet  $\Sigma_V := \mathbb{F}^{3m+5}$ . Outer right labels use all coordinates; inner right labels are padded into the first coordinate.

*Left vertices.* A left vertex is a pair  $u = (\Omega, \tau) \in \mathcal{O} \times \mathbb{T}$ . Hence  $U := \mathcal{O} \times \mathbb{T}$  and  $|U| = q^{11m+O(1)}$ .

*Left alphabet.* An element of the global left alphabet  $\Sigma_U$  is a tuple

$$\lambda = (\widehat{\Pi}_A, \widehat{\Pi}_P, \widehat{\Pi}_1, \dots, \widehat{\Pi}_{3m+3}; \lambda_{\text{in}}),$$

where each  $\widehat{\Pi}_*$  is a degree- $D$  polynomial on  $\mathbb{F}^5$  and  $\lambda_{\text{in}}$  is a padded auxiliary-answer string of length  $Q_{\text{in}}$  over  $\mathbb{F}$ . In words, a left label consists of (a tuple of polynomials claiming to be) the outer polynomials restricted to an object, together with one padded auxiliary-answer string. We view the outer component as a tuple of polynomials on the reference chart  $\mathbb{F}^5$ , not directly on the object's affine space  $L(\Omega)$ . Since the number of coefficients of a degree- $D$  polynomial in five variables is  $q^{O(1)}$ , we have  $|\Sigma_U| \leq q^{q^{O(1)}}$ .

For a fixed left vertex  $u = (\Omega, \tau)$ , the admissible subset  $\Sigma_u^{\text{adm}} \subseteq \Sigma_U$  consists of those tuples  $\lambda$  for which the following conditions hold.

1. **Accepted Local View:**  $\lambda_{\text{in}}$ , interpreted using the fixed  $Q_{\text{in}}$ -slot convention above and combined with the corresponding raw-answer string from  $B_\Phi$ , is an accepted local view of  $\text{Eval}_{z,\tau}$ , where  $z$  is the distinguished point of  $\Omega$ . By item (1), this local view has a well-defined decoded value; write

$$\text{Dec}(\lambda_{\text{in}}) \in \mathbb{F}$$

for that value.

2. **Dependence only on the first  $m$  coordinates at the distinguished pair:**  $\widehat{\Pi}_A(0) = \widehat{\Pi}_A(e_3)$ .
3. **Vanishing witness identity on the object:**

$$\widehat{\Pi}_P(t) = \sum_{j \in [3m+3]} g_H((\psi_\Omega(t))_j) \widehat{\Pi}_j(t) \quad \forall t \in \mathbb{F}^5.$$

4. **Clause check at the distinguished point:**

$$\widehat{\Pi}_P(0) = \text{Dec}(\lambda_{\text{in}}) (\widehat{\Pi}_A(0) - z_{3m+1}) (\widehat{\Pi}_A(e_4) - z_{3m+2}) (\widehat{\Pi}_A(e_5) - z_{3m+3}).$$

The left predicate  $P_u^\Phi : \Sigma_U \rightarrow \{0, 1\}$  is the indicator of these conditions.

*Edges and projections.* Fix  $K := q^{c_{\text{in}}+4}$ . Fix  $u = (\Omega, \tau)$ . For every  $t \in \mathbb{F}^5$  and every  $\kappa \in [K]$ , add an outer edge from  $u$  to  $v_{\psi_\Omega(t)}^{\text{out}}$ . The corresponding projection sends

$$\lambda = (\widehat{\Pi}_A, \widehat{\Pi}_P, \widehat{\Pi}_1, \dots, \widehat{\Pi}_{3m+3}; \lambda_{\text{in}})$$

to

$$(\widehat{\Pi}_A(t), \widehat{\Pi}_P(t), \widehat{\Pi}_1(t), \dots, \widehat{\Pi}_{3m+3}(t)) \in \Sigma_V.$$

Thus each outer edge is repeated with multiplicity  $K$ . This is used only to make the outer edges dominate the inner ones in the edge distribution. Let  $w_1, \dots, w_{Q_{\text{in}}} \in V_{\text{in}}$  be the ordered auxiliary query positions of  $\text{Eval}_{z,\tau}$  under the fixed padding convention. For each slot  $\ell \in [Q_{\text{in}}]$ , add one inner edge from  $u$  to  $v_{w_\ell}^{\text{in}}$ , and let the projection send  $\lambda$  to the  $\ell$ th entry of  $\lambda_{\text{in}}$ , padded with zeros to land in  $\Sigma_V$ . If the same auxiliary coordinate appears in several slots, these are parallel inner edges.

Thus every left degree is exactly  $d_u = Kq^5 + Q_{\text{in}} = q^{O(1)}$ .

**Completeness.** Assume that  $\Phi$  is satisfiable. Let  $A : \mathbb{F}^m \rightarrow \mathbb{F}$  be the low-degree extension of a satisfying assignment. Define the ambient assignment polynomial  $T_A(x) := A(x_1, \dots, x_m)$  ( $x \in \mathbb{F}^{3m+3}$ ), and let  $T_P := P_{\Phi, A}$ . Since  $A|_{H^m}$  satisfies every clause of  $\Phi$ , the polynomial  $T_P$  vanishes on  $H^{3m+3}$ . By [Lemma 8.2](#), there exist polynomials  $T_1, \dots, T_{3m+3}$  such that

$$T_P(x) = \sum_{j \in [3m+3]} g_H(x_j) T_j(x) \quad \forall x \in \mathbb{F}^{3m+3}.$$

By [Theorem 8.4](#), the pair  $(W_{\text{in}}(B_\Phi), B_\Phi)$  induces for every evaluator  $\text{Eval}_{z, \tau}$  an accepted local view whose decoded value is the correct value  $\tilde{B}_\Phi(z)$ .

Label every outer right vertex  $v_x^{\text{out}}$  by

$$(T_A(x), T_P(x), T_1(x), \dots, T_{3m+3}(x)),$$

and every inner right vertex by the corresponding padded symbol of  $W_{\text{in}}(B_\Phi)$ . For a left vertex  $u = (\Omega, \tau)$ , label it with the tuple obtained by restricting  $T_A, T_P, T_1, \dots, T_{3m+3}$  to the chart  $\psi_\Omega$  together with the accepting padded auxiliary-answer string of  $\text{Eval}_{z, \tau}$ . All admissibility conditions are satisfied, and every incident edge is satisfied. Therefore  $\text{val}(I_\Phi) = 1$ .

**Soundness.** Now assume that  $\Phi$  is unsatisfiable, and let  $\pi = (\pi_U, \pi_V)$  be an arbitrary labeling of  $I_\Phi$ . Write  $p := \text{val}_{I_\Phi}(\pi)$ . The soundness proof has five steps. We first turn the labeling into an object oracle (in the sense of [Theorem 8.3](#)) and apply [Theorem 8.3](#) to obtain a short list of global candidate tuples. We then define a set of “good tuples” satisfying three key properties ([Definition 8.5](#)) and show that any good tuple would induce a satisfying assignment to  $\Phi$ . Finally, we bound the contribution from tuples that fail each of the three properties. Summing these bounds and comparing outer and inner edges yields  $p \leq \varepsilon$ .

We begin by extracting an object oracle  $\bar{O}$  from the labeling  $\pi$ . From the outer right labels define a function  $\varphi_\pi : \mathbb{F}^{3m+3} \rightarrow \mathbb{F}^{3m+5}$  by letting  $\varphi_\pi(x)$  be the label of  $v_x^{\text{out}}$ . For the soundness analysis only, augment it by an admissibility coordinate:  $\bar{\varphi}_\pi(x) := (\varphi_\pi(x), 1) \in \mathbb{F}^{3m+6}$ .

For a left vertex  $u = (\Omega, \tau)$ , write  $b_\pi(u) := P_u^\Phi(\pi_U(u)) \in \{0, 1\}$  for the evaluation of the left-predicate  $P_u^\Phi$  associated with  $u$  on  $\pi$ . The label  $\pi_U(u)$  determines a degree- $D$  tuple of polynomials on  $L(\Omega)$ ; denote its algebraic part by

$$O_\pi^{\text{alg}}(\Omega, \tau) := (\Pi_A, \Pi_P, \Pi_1, \dots, \Pi_{3m+3}),$$

where  $\Pi_\star(x) := \hat{\Pi}_\star(\psi_\Omega^{-1}(x))$  for  $x \in L(\Omega)$ . Define the augmented object oracle

$$\bar{O}_\pi(\Omega, \tau) := (\Pi_A, \Pi_P, \Pi_1, \dots, \Pi_{3m+3}, b_\pi(u)),$$

where the last coordinate is the constant polynomial with value  $b_\pi(u)$  on  $L(\Omega)$ . Thus  $\bar{O}_\pi$  is an object oracle of the type required by [Theorem 8.3](#).

Since every left vertex has exactly  $Kq^5$  outer edges and  $Q_{\text{in}} := q^{c_{\text{in}}}$  inner edges, our choice  $K = q^{c_{\text{in}}+4}$  yields  $\frac{Q_{\text{in}}}{Kq^5 + Q_{\text{in}}} \leq q^{-4}$ . Hence, recalling that  $p := \text{val}_{I_\Phi}(\pi)$ ,

$$p \leq p_{\text{out}} + q^{-4}, \tag{2}$$

where

$$p_{\text{out}} := \Pr_{\Omega \sim \mathcal{O}, \tau \sim \mathbb{T}, x \sim L(\Omega)} [\bar{O}_\pi(\Omega, \tau)(x) = \bar{\varphi}_\pi(x)].$$

By construction,  $p_{\text{out}}$  is exactly the satisfaction probability of a uniformly random outer edge.

We now apply [Theorem 8.3](#) to  $\bar{\varphi}_\pi$  and  $\bar{O}_\pi$  with some parameter  $\delta$  to be fixed later to obtain  $t \leq \frac{C_{\text{PO}}}{\delta}$  global polynomial tuples

$$\bar{Q}_i = (Q_{i,A}, Q_{i,P}, Q_{i,1}, \dots, Q_{i,3m+3}, Q_{i,\text{adm}}) \quad (i \in [t])$$

of coordinate-wise degree at most  $D$  such that

$$p_{\text{out}} \leq \delta + C_{\text{err}} \frac{1+tD}{q} + \sum_{i \in [t]} \alpha_i, \quad (3)$$

where

$$\alpha_i := \Pr_{\Omega \sim \mathcal{O}, \tau \sim \mathbb{T}, x \sim L(\Omega)} [\bar{Q}_i|_{L(\Omega)} \equiv \bar{O}_\pi(\Omega, \tau) \text{ and } \bar{O}_\pi(\Omega, \tau)(x) = \bar{\varphi}_\pi(x)].$$

The second conjunct forces  $b_\pi(u) = 1$ , so every  $\alpha_i$  counts only admissible left labels.

For each  $i \in [t]$ , let  $\vec{Q}_i := (Q_{i,A}, Q_{i,P}, Q_{i,1}, \dots, Q_{i,3m+3})$  be the tuple obtained from  $\bar{Q}_i$  by omitting the admissibility coordinate. We next show that any such tuple contributing non-negligibly would yield a satisfying assignment to  $\Phi$ , contradicting unsatisfiability.

**Definition 8.5.** A tuple  $\vec{Q} = (Q_A, Q_P, Q_1, \dots, Q_{3m+3})$  of coordinate-wise degree- $D$  polynomials on  $\mathbb{F}^{3m+3}$  is *good* if the following hold.

1.  $Q_A$  depends only on the first  $m$  coordinates.
2. For every  $x \in \mathbb{F}^{3m+3}$ ,

$$Q_P(x) = \sum_{j \in [3m+3]} g_H(x_j) Q_j(x).$$

3. For every  $x \in \mathbb{F}^{3m+3}$ ,

$$Q_P(x) = \tilde{B}_\Phi(x) (Q_A(x) - x_{3m+1}) (Q_A(\rho(x)) - x_{3m+2}) (Q_A(\rho^2(x)) - x_{3m+3}).$$

**Claim 8.6.** Every good tuple induces a satisfying Boolean assignment to  $\Phi$ .

*Proof.* Because  $Q_A$  depends only on the first  $m$  coordinates, it induces a well-defined function on  $H^m$  via the rule  $Q_A^b(y) := Q_A(y \circ 0^{2m+3})$  for  $y \in H^m$ . Fix once and for all a map  $\beta : \mathbb{F} \rightarrow \{0, 1\}$  with  $\beta(0) = 0$  and  $\beta(1) = 1$ , and define the Boolean assignment  $a(y) := \beta(Q_A^b(y))$  for  $y \in H^m$ . Let  $x = (u, v, w, b_1, b_2, b_3) \in H^{3m+3}$  be any point which indicates a clause of  $\Phi$ . Then  $B_\Phi(x) = 1$ , hence  $\tilde{B}_\Phi(x) = 1$ . By good-property (2),  $Q_P(x) = \sum_{j \in [3m+3]} g_H(x_j) Q_j(x)$ , and each  $g_H(x_j)$  vanishes on  $H$ , so  $Q_P$  vanishes on  $H^{3m+3}$ . Therefore good-property (3) gives

$$0 = (Q_A^b(u) - b_1) (Q_A^b(v) - b_2) (Q_A^b(w) - b_3).$$

At least one factor is zero; suppose  $Q_A^b(u) = b_1$ . Since  $b_1 \in \{0, 1\}$ ,  $a(u) = \beta(Q_A^b(u)) = \beta(b_1) = b_1$ , so the literal  $x_u^{b_1}$  is satisfied. The same argument applies if the vanishing factor is the second or third one. Hence every clause of  $\Phi$  is satisfied by  $a$ .  $\square$

Since  $\Phi$  is unsatisfiable, every tuple  $\vec{Q}_i$  is bad. We now bound the corresponding quantities  $\alpha_i$ .

**Claim 8.7.** If  $\vec{Q}_i$  violates good-property (1), then  $\alpha_i \leq (1 + C_{\text{geom}}) \frac{D}{q}$ .

*Proof.* Whenever the event inside  $\alpha_i$  occurs, the second conjunct gives  $b_\pi(u) = 1$ , so the left label at  $u = (\Omega, \tau)$  is admissible. The first conjunct gives  $\vec{Q}_i|_{L(\Omega)} \equiv O_\pi^{\text{alg}}(\Omega, \tau)$ . Let  $O_\pi^{\text{alg}}(\Omega, \tau)_A$  be the part of  $O_\pi^{\text{alg}}(\Omega, \tau)$  corresponding to  $\Pi_A$ . Because every admissible left label satisfies  $\widehat{\Pi}_A(0) = \widehat{\Pi}_A(e_3)$  and  $O_\pi^{\text{alg}}(\Omega, \tau)_A(x) = \widehat{\Pi}_A(\psi_\Omega^{-1}(x))$  for  $x \in L(\Omega)$ , we have  $O_\pi^{\text{alg}}(\Omega, \tau)_A(z) = O_\pi^{\text{alg}}(\Omega, \tau)_A(z')$ . Therefore the event inside  $\alpha_i$  implies  $Q_{i,A}(z) = Q_{i,A}(z')$ . Write points of  $\mathbb{F}^{3m+3}$  as  $(s, r)$  with  $s \in \mathbb{F}^m$  and  $r \in \mathbb{F}^{2m+3}$ , and abbreviate  $Q_{i,A}(s, r) := Q_{i,A}(s \circ r)$ . Since  $\vec{Q}_i$  violates property (1), the polynomial  $R(s, r, r') := Q_{i,A}(s, r) - Q_{i,A}(s, r')$  on  $\mathbb{F}^m \times \mathbb{F}^{2m+3} \times \mathbb{F}^{2m+3}$  is nonzero and has degree at most  $D$ . By Schwartz–Zippel,  $\Pr_{s,r,r'}[R(s, r, r') = 0] \leq D/q$ . The actual marginal of  $(z, z')$  is within total-variation distance at most  $C_{\text{geom}}/q$  of the uniform distribution, by the definition of our objects, on pairs sharing the first  $m$  coordinates, namely pairs of the form  $((s, r), (s, r'))$ . Hence the same bound becomes at most  $(1 + C_{\text{geom}})D/q$ .  $\square$

**Claim 8.8.** If  $\vec{Q}_i$  violates good-property (2), then  $\alpha_i \leq (1 + C_{\text{geom}})\frac{D+|H|}{q}$ .

*Proof.* Let  $\Delta_i(x) := Q_{i,P}(x) - \sum_{j \in [3m+3]} g_H(x_j) Q_{i,j}(x)$ . This is a nonzero polynomial of degree at most  $D + |H|$ . Whenever the event inside  $\alpha_i$  occurs, the second conjunct gives  $b_\pi(u) = 1$ , so the left label at  $u = (\Omega, \tau)$  is admissible, while the first conjunct gives  $\vec{Q}_i|_{L(\Omega)} \equiv O_\pi^{\text{alg}}(\Omega, \tau)$ . Every admissible left label satisfies the witness identity on all of  $L(\Omega)$ , so  $\Delta_i$  vanishes on  $L(\Omega)$  and, in particular, at the point  $y \in L(\Omega)$ . The marginal of  $y$  is within total-variation distance at most  $C_{\text{geom}}/q$  of the uniform distribution on  $\mathbb{F}^{3m+3}$ , so Schwartz–Zippel yields the claimed bound.  $\square$

**Claim 8.9.** If  $\vec{Q}_i$  violates good-property (3), then  $\alpha_i \leq (4 + C_{\text{geom}})\frac{D}{q}$ .

*Proof.* Define

$$\Gamma_i(x) := Q_{i,P}(x) - \widetilde{B}_\Phi(x)(Q_{i,A}(x) - x_{3m+1})(Q_{i,A}(\rho(x)) - x_{3m+2})(Q_{i,A}(\rho^2(x)) - x_{3m+3}).$$

Since  $\vec{Q}_i$  violates property (3), the polynomial  $\Gamma_i$  is nonzero and has degree at most  $4D$ . Whenever the event inside  $\alpha_i$  occurs, the second conjunct gives  $b_\pi(u) = 1$ , so the left label at  $u = (\Omega, \tau)$  is admissible, while the first conjunct gives  $\vec{Q}_i|_{L(\Omega)} \equiv O_\pi^{\text{alg}}(\Omega, \tau)$ . The admissibility of the left label at  $(\Omega, \tau)$  gives the local clause identity at the distinguished point  $z$ . By [Theorem 8.4](#), the accepted local view, and hence its auxiliary-answer string, decodes exactly  $\widetilde{B}_\Phi(z)$ . Hence  $\Gamma_i(z) = 0$ . The marginal of  $z$  is within total-variation distance at most  $C_{\text{geom}}/q$  of the uniform distribution on  $\mathbb{F}^{3m+3}$ , so Schwartz–Zippel gives the stated bound.  $\square$

Combining [Claim 8.7](#), [Claim 8.8](#), and [Claim 8.9](#), we obtain for every  $i \in [t]$ ,

$$\alpha_i \leq \frac{(6 + 3C_{\text{geom}})D + (1 + C_{\text{geom}})|H|}{q} \leq C_{\text{bad}} \frac{D + |H|}{q},$$

where  $C_{\text{bad}} := 6 + 3C_{\text{geom}}$ . Substituting this into (3) and then using (2), we conclude that

$$p \leq \delta + q^{-4} + C_{\text{err}} \frac{1 + tD}{q} + C_{\text{bad}} t \frac{D + |H|}{q}.$$

Now choose  $\delta := \varepsilon/10$ . Since  $\varepsilon \geq \log^{-c} n$ ,  $D = 8m|H|$ ,  $q = (\log n)^a$  with  $a = 4c + 7$ , and all hidden constants above are absolute, the threshold  $n_0(c)$  in the theorem statement can be chosen so that every  $n \geq n_0(c)$  satisfies  $c_{\text{PO}} m \left(\frac{mD}{q}\right)^{1/4} \leq \delta$  and

$$q^{-4} + \frac{C_{\text{err}}}{q} + \frac{10C_{\text{PO}}(C_{\text{err}}D + C_{\text{bad}}(D + |H|))}{\varepsilon q} \leq \frac{9\varepsilon}{10}.$$

Hence the point-object theorem applies, the last three terms above total at most  $9\varepsilon/10$ , and therefore  $p \leq \varepsilon$ . This proves soundness.

**Swap locality and parameter bounds.** We now verify the remaining claims of [Theorem 8.1](#).

If  $\Phi$  and  $\Phi'$  differ by one clause swap, then the raw tables  $B_\Phi$  and  $B_{\Phi'}$  differ at exactly two points of  $H^{3m+3}$ . The graph  $(U, V, E)$ , the alphabets  $\Sigma_U, \Sigma_V$ , and the projections  $f_e$  depend only on  $(m, q, H)$  and on the fixed object system  $\mathcal{O}$  together with the associated affine sets  $L(\Omega)$  and distinguished points, not on the particular formula. The only formula-dependent part of the construction is the left predicate at a vertex  $u = (\Omega, \tau)$ , namely the requirement that the auxiliary-answer string of  $\text{Eval}_{z,\tau}$ , together with the corresponding raw-answer string from  $B_\Phi$ , form an accepted local view.

By [Theorem 8.4](#), each changed raw table point participates in at most  $q^{c_{\text{in}}}$  evaluator tests  $\text{Eval}_{z,\tau}$ . Fix one affected pair  $(z, \tau)$ . The number of objects with distinguished point  $z$  is at most

$$q^{2m+3} \cdot q^{3m+3} \cdot q^{3m+3} = q^{8m+9},$$

since one first chooses  $z'$  with the same first  $m$  coordinates as  $z$  and then chooses  $y, y'$  arbitrarily subject to affine independence. Therefore at most

$$2 \cdot q^{c_{\text{in}}} \cdot q^{8m+9} = 2q^{8m+c_{\text{in}}+9}$$

left predicates can change under one clause swap. This is the claimed transformation-sensitivity bound after choosing  $c_{\text{lc}}$  so that  $c_{\text{lc}} \geq c_{\text{in}} + 9$ .

For the size parameters,

$$|V| = |\mathbb{F}|^{3m+3} + |V_{\text{in}}| = q^{3m+O(1)}$$

by construction, and

$$|U| = |\mathcal{O}| \cdot |\mathbb{T}|.$$

We already observed that every left degree is exactly

$$d_u = Kq^5 + Q_{\text{in}} = q^{c_{\text{in}}+9} + q^{c_{\text{in}}}.$$

Because  $n \geq n_0(c)$ , we may also assume  $q \geq 4$ , so each crude lower bound of the form  $q^r - q^s \geq q^{r-1}$  used below is valid.

To lower-bound  $|\mathcal{O}|$ , consider the condition that  $z, \rho(z), \rho^2(z)$  are affinely independent. This is the nonvanishing of a certain  $2 \times 2$  minor of the matrix with columns  $\rho(z) - z$  and  $\rho^2(z) - z$ ; for instance, using the first coordinates of the first two variable blocks gives

$$\det \begin{pmatrix} z_{m+1} - z_1 & z_{2m+1} - z_1 \\ z_{2m+1} - z_{m+1} & z_1 - z_{m+1} \end{pmatrix},$$

which is not identically zero. By Schwartz–Zippel, there are at least

$$q^{3m+2}$$

points  $z \in \mathbb{F}^{3m+3}$  for which  $z, \rho(z), \rho^2(z)$  are affinely independent. Fix such a point  $z$ .

The number  $N_z$  of objects with distinguished point  $z$  therefore satisfies

$$(q^{2m+3} - q^2)(q^{3m+3} - q^3)(q^{3m+3} - q^4) \leq N_z \leq q^{8m+9},$$

because one chooses  $z'$  with the same first  $m$  coordinates as  $z$  outside the affine plane  $\text{aff}(z, \rho(z), \rho^2(z))$ , then chooses  $y$  outside the resulting 3-flat, and finally chooses  $y'$  outside the resulting 4-flat. In particular,

$$N_z \geq q^{8m+6}.$$

Summing this lower bound over the at least  $q^{3m+2}$  admissible choices of  $z$  gives

$$|\mathcal{O}| \geq q^{3m+2} \cdot q^{8m+6} = q^{11m+8}.$$

Since  $|\mathbb{T}| \geq 1$ , it follows that

$$|U| = |\mathcal{O}| \cdot |\mathbb{T}| \geq q^{11m+8}.$$

The crude counting bound  $|\mathcal{O}| \leq q^{11m+O(1)}$  together with  $|\mathbb{T}| = q^{O(1)}$  also gives

$$|U| = |\mathcal{O}| \cdot |\mathbb{T}| \leq q^{11m+O(1)}.$$

This lower bound on  $|U|$  will be used again later to convert counts of affected left vertices into outer-edge mass bounds.

For an outer right vertex  $v_x^{\text{out}}$ , fix  $y' := x$ . The condition that  $x, z, \rho(z), \rho^2(z)$  are affinely independent is the nonvanishing of a certain degree-3 polynomial in  $z$ ; for instance, one  $3 \times 3$  minor of the matrix with columns  $\rho(z) - z, \rho^2(z) - z$ , and  $x - z$  is

$$(x_{3m+1} - z_{3m+1}) \det \begin{pmatrix} z_{m+1} - z_1 & z_{2m+1} - z_1 \\ z_{2m+1} - z_{m+1} & z_1 - z_{m+1} \end{pmatrix},$$

which is not identically zero. By Schwartz–Zippel there are at least  $q^{3m+2}$  valid choices of  $z$ . For each such  $z$ , choose  $z'$  with the same first  $m$  coordinates as  $z$  outside the 3-flat  $\text{aff}(x, z, \rho(z), \rho^2(z))$ , and then choose  $y$  outside the resulting 4-flat. This yields at least

$$q^{3m+2} \cdot q^{2m+2} \cdot q^{3m+2} = q^{8m+6}$$

objects with  $x \in L(\Omega)$ . For the upper bound, choose  $z \in \mathbb{F}^{3m+3}$  in  $q^{3m+3}$  ways, choose  $z'$  with the same first  $m$  coordinates as  $z$  in at most  $q^{2m+3}$  ways, and choose  $y \in \mathbb{F}^{3m+3}$  in  $q^{3m+3}$  ways. Once  $x, z, z', y$  are fixed, the condition  $x \in L(\Omega)$  forces  $y'$  to lie in the 5-flat  $\text{aff}(x, z, y, z', \rho(z), \rho^2(z))$ , so there are at most  $q^5$  choices for  $y'$ . Thus at most  $q^{8m+14}$  objects  $\Omega$  satisfy  $x \in L(\Omega)$ . Multiplying by the seed factor  $|\mathbb{T}| = q^{O(1)}$  and by the parallel-edge factor  $K = q^{c_{\text{in}}+4}$  shows that

$$q^{8m+6} \leq d(v_x^{\text{out}}) \leq q^{8m+O(1)}.$$

For an inner right vertex  $v_w^{\text{in}}$ , our pruning of  $V_{\text{in}}$  in the definition of the right vertices guarantees that  $w$  is queried by at least one pair  $(z, \tau)$  arising from some left vertex. Fix such a pair. Every object with distinguished point  $z$  and the same seed  $\tau$  contributes at least one edge to  $v_w^{\text{in}}$ , so the lower bound on  $N_z$  gives

$$q^{8m+6} \leq d(v_w^{\text{in}}).$$

The bounded-locality statement in [Theorem 8.4](#) still gives at most  $q^{c_{\text{in}}}$  such pairs  $(z, \tau)$ . For each such pair and each object with distinguished point  $z$ , the slot model contributes at most  $Q_{\text{in}} = q^{c_{\text{in}}}$  parallel edges to  $v_w^{\text{in}}$ , and each pair extends to at most  $q^{8m+9}$  objects. Hence

$$d(v_w^{\text{in}}) \leq q^{8m+2c_{\text{in}}+9}.$$

After enlarging  $c_{\text{lc}}$  to absorb the bound on  $|V|$ , the lower and upper bounds on  $|U|$ , the outer-right upper bound, the inner-right upper bound, and the coefficient count in  $|\Sigma_U|$ , all of the parameter bounds in the theorem follow. Finally,

$$|\Sigma_V| = q^{3m+5}, \quad |\Sigma_U| \leq q^{q^{c_{\text{lc}}}}.$$

This completes the proof of [Theorem 8.1](#).

**Remark.** This construction is used later because it combines low soundness, explicit control of how the instance changes under clause swaps, and very large minimum right degree. A decodable lift at this stage would collapse the right degree to the decoder proof degree and would forfeit that gain.

## 8.2 Recovering Satisfying Assignments under Local Perturbations

This subsection records the recovery theorem with perturbation guarantees that we need beyond Theorem 8.1. Continue in the setting of Theorem 8.1. For a satisfiable CNF formula  $\Phi$  on  $n$  variables, write  $I_\Phi := \mathcal{T}_{\text{Robust}}(\Phi)$  for the associated left-predicate label cover instance.

**Theorem 8.10** (Recovery from large-score tuples). There exists an absolute constant  $C_{\text{hit}} \geq 1$  such that the following holds. Fix a parameter  $\Delta \in (0, 1]$ . Assume that

$$\Delta \geq C_{\text{hit}} \max \left\{ m \left( \frac{mD}{q} \right)^{1/4}, \sqrt{\frac{D + |H|}{q}}, q^{-2} \right\}.$$

Then there is an explicit randomized map

$$\mathcal{H}_{\text{lc}, \Phi, \Delta} : \{\text{labelings of } I_\Phi\} \rightarrow \{0, 1\}^{\text{Var}(\Phi)}$$

with the following properties.

1. If  $\pi$  is any labeling of  $I_\Phi$  with

$$\text{val}_{I_\Phi}(\pi) \geq \Delta,$$

then

$$\Pr[\mathcal{H}_{\text{lc}, \Phi, \Delta}(\pi) \text{ satisfies } \Phi] \geq \Delta^{C_{\text{hit}}}.$$

2. If two labelings of  $I_\Phi$  differ in one left or right coordinate, then under the natural coupling their outputs differ in expected Hamming distance at most

$$q^{C_{\text{hit}}}.$$

3. If  $\Phi$  and  $\Phi'$  differ by one clause swap, then for every common-domain labeling  $\pi$  of  $I_\Phi$  and  $I_{\Phi'}$ ,

$$\text{EMD}(\mathcal{H}_{\text{lc}, \Phi, \Delta}(\pi), \mathcal{H}_{\text{lc}, \Phi', \Delta}(\pi)) \leq q^{C_{\text{hit}}}.$$

The idea is to score global degree- $D$  tuples by their outer-edge agreement probability, note that only a few tuples can have large agreement, prove that these scores are stable under local perturbations, and then select among the surviving tuples using a random threshold. This yields the explicit randomized map from Theorem 8.10. For the proof, let  $\pi$  denote a labeling of  $I_\Phi$ . Write

$$\varphi_\pi : \mathbb{F}^{3m+3} \rightarrow \mathbb{F}^{3m+5}$$

for the outer right labeling induced by  $\pi$ , and for  $u = (\Omega, \tau)$  write

$$b_\pi(u) := P_u^\Phi(\pi_U(u)) \in \{0, 1\}, \quad O_\pi^{\text{alg}}(\Omega, \tau)$$

for the admissibility bit and the algebraic chart determined by the left label at  $u$ , exactly as in the soundness proof. We also use the augmented objects

$$\bar{\varphi}_\pi(x) := (\varphi_\pi(x), 1), \quad \bar{O}_\pi(\Omega, \tau) := (O_\pi^{\text{alg}}(\Omega, \tau), b_\pi(\Omega, \tau)),$$

which are the direct analogues of  $\bar{\varphi}_\pi$  and  $\bar{O}_\pi$  from the soundness argument. For every coordinatewise degree- $D$  tuple

$$\vec{Q} = (Q_A, Q_P, Q_1, \dots, Q_{3m+3})$$

define its outer-edge agreement probability against  $(\Phi, \pi)$ , which we will call its *score*, by

$$s_{\Phi, \pi}(\vec{Q}) := \Pr_{\Omega \sim \mathcal{O}, \tau \sim \mathbb{T}, x \sim L(\Omega)} [b_\pi(\Omega, \tau) = 1, \vec{Q}|_{L(\Omega)} \equiv O_\pi^{\text{alg}}(\Omega, \tau), \vec{Q}(x) = \varphi_\pi(x)].$$

Thus  $s_{\Phi, \pi}(\vec{Q})$  is the probability that a uniformly random outer edge is simultaneously admissible, satisfied, and explained by the global tuple  $\vec{Q}$ .

**Lemma 8.11** (Bad tuples have low outer-edge agreement). There exists an absolute constant  $C_{\text{light}} \geq 1$  such that for every labeling  $\pi$  of  $I_\Phi$  and every coordinatewise degree- $D$  tuple  $\vec{Q}$ , if  $\vec{Q}$  is not good, then

$$s_{\Phi, \pi}(\vec{Q}) \leq C_{\text{light}} \frac{D + |H|}{q}.$$

*Proof.* This is the combination of the three bad-type bounds from the soundness proof of Theorem 8.1; see Claims 8.7 to 8.9. Each proof only uses the event being bounded, so the same arguments apply here with that event replaced by the score event defining  $s_{\Phi, \pi}(\vec{Q})$ . The same computation gives

$$s_{\Phi, \pi}(\vec{Q}) \leq \frac{(6 + 3C_{\text{geom}})D + (1 + C_{\text{geom}})|H|}{q} \leq C_{\text{light}} \frac{D + |H|}{q}$$

once we take

$$C_{\text{light}} := 6 + 3C_{\text{geom}}.$$

This recovers the claimed bound.  $\square$

**Lemma 8.12** (Pairwise outer-chart overlap). There exists an absolute constant  $C_{\text{ov}} \geq 1$  such that for every labeling  $\pi$  of  $I_\Phi$  and every two distinct coordinatewise degree- $D$  tuples  $\vec{Q} \neq \vec{Q}'$ , we have

$$\Pr_{\Omega \sim \mathcal{O}, \tau \sim \mathbb{T}, x \sim L(\Omega)} [b_\pi(\Omega, \tau) = 1, \vec{Q}|_{L(\Omega)} \equiv O_\pi^{\text{alg}}(\Omega, \tau), \vec{Q}'|_{L(\Omega)} \equiv O_\pi^{\text{alg}}(\Omega, \tau)] \leq C_{\text{ov}} \frac{D}{q}.$$

*Proof.* Choose a coordinate on which the degree- $D$  polynomials for  $\vec{Q}$  and  $\vec{Q}'$  on this coordinate differ, and let  $R$  be the nonzero difference polynomial on  $\mathbb{F}^{3m+3}$ . If the event whose probability we are bounding in the lemma statement occurs, then  $R$  vanishes on all of  $L(\Omega)$ , hence in particular at the point  $y \in L(\Omega)$  from the object definition. The marginal of  $y$  is within total-variation distance  $C_{\text{geom}}/q$  of the uniform distribution on  $\mathbb{F}^{3m+3}$ , so Schwartz–Zippel gives

$$\Pr[R(y) = 0] \leq (1 + C_{\text{geom}}) \frac{D}{q}.$$

This proves the claim with  $C_{\text{ov}} := 1 + C_{\text{geom}}$ .  $\square$

**Lemma 8.13** (Only  $O(1/\theta)$  tuples can have large outer-edge agreement). There exists an absolute constant  $C_{\text{pack}} \geq 1$  such that the following holds. Fix  $\theta \in (0, 1]$  with

$$\theta^2 \geq 20C_{\text{ov}} \frac{D}{q}.$$

Then, for every labeling  $\pi$  of  $I_\Phi$ , the number of tuples  $\vec{Q}$  with

$$s_{\Phi,\pi}(\vec{Q}) \geq \theta$$

is at most

$$\frac{C_{\text{pack}}}{\theta}.$$

*Proof.* Let  $\mathcal{S}$  be the set of tuples with score at least  $\theta$ , and write  $k := |\mathcal{S}|$ . For  $\vec{Q} \in \mathcal{S}$ , let  $E_{\vec{Q}}$  denote the score event in the definition of  $s_{\Phi,\pi}(\vec{Q})$ , so  $\Pr[E_{\vec{Q}}] \geq \theta$ . By Lemma 8.12, every two distinct tuples in  $\mathcal{S}$  satisfy

$$\Pr[E_{\vec{Q}} \wedge E_{\vec{Q}'}] \leq C_{\text{ov}} \frac{D}{q} \leq \frac{\theta^2}{20}.$$

Applying the first two terms of inclusion–exclusion gives

$$1 \geq \Pr\left[\bigcup_{\vec{Q} \in \mathcal{S}} E_{\vec{Q}}\right] \geq k\theta - \binom{k}{2} \frac{\theta^2}{20}.$$

If  $k \geq 6/\theta$ , then the right-hand side is at least

$$6 - \frac{36}{40} > 1,$$

a contradiction. Thus  $k < 6/\theta$ , and the lemma follows with  $C_{\text{pack}} := 6$ .  $\square$

**Lemma 8.14** (Perturbation bounds for outer-edge agreement). There exists an absolute constant  $C_{\text{drift}} \geq 1$  such that the following holds. Let  $\vec{Q}$  be any coordinatewise degree- $D$  tuple.

1. If two labelings of  $I_\Phi$  differ in one left or right coordinate, then

$$|s_{\Phi,\pi}(\vec{Q}) - s_{\Phi,\tilde{\pi}}(\vec{Q})| \leq q^{-3m+C_{\text{drift}}}.$$

2. If  $\Phi$  and  $\Phi'$  differ by one clause swap and  $\pi$  is a common labeling of  $I_\Phi$  and  $I_{\Phi'}$ , then

$$|s_{\Phi,\pi}(\vec{Q}) - s_{\Phi',\pi}(\vec{Q})| \leq q^{-3m+C_{\text{drift}}}.$$

*Proof.* The score event for  $\vec{Q}$  is determined by a uniformly random outer edge. Every left vertex contributes exactly  $Kq^5$  outer edges, so the total number of outer edges is  $|U|Kq^5$ . Changing one left coordinate affects only outer edges incident to that left vertex, whose total mass is exactly

$$\frac{Kq^5}{|U|Kq^5} = \frac{1}{|U|}.$$

By the lower bound on  $|U|$  in item (4) of Theorem 8.1, this is at most

$$q^{-11m+c_{\text{lc}}} = q^{-11m+O(1)}.$$

Changing one inner right coordinate does not affect any outer edge at all. Changing one outer right coordinate can affect only the outer edges incident to the corresponding outer right vertex. The outer-right degree upper bound and the same lower bound on  $|U|$  from item (4) of Theorem 8.1 show that this mass is at most

$$\frac{q^{8m+O(1)}}{|U|Kq^5} = q^{-3m+O(1)}.$$

This proves item (1). For item (2), a one-clause swap changes only the left predicates on at most  $2q^{8m+c_{lc}}$  left vertices by item (3) of Theorem 8.1. Since the score uses the formula only through the bit  $b_\pi(u)$ , the score can change only when the sampled outer edge is incident to one of those vertices. Using again that each left vertex contributes exactly  $Kq^5$  outer edges and that item (4) of Theorem 8.1 gives  $|U| \geq q^{11m-c_{lc}}$ , this has total mass at most

$$\frac{2q^{8m+c_{lc}}Kq^5}{|U|Kq^5} = \frac{2q^{8m+c_{lc}}}{|U|} = q^{-3m+O(1)}.$$

After enlarging  $C_{\text{drift}}$ , both bounds take the stated form.  $\square$

*Proof of Theorem 8.10.* Fix  $\Delta$  as in the statement and set

$$\delta := \Delta/20.$$

Because  $\Delta \geq 20c_{\text{PO}}m((mD)/q)^{1/4}$  after enlarging  $C_{\text{hit}}$ , the Point-object theorem (Theorem 8.3) applies with this choice of  $\delta$ . Moreover, the lower bound on  $\Delta$  also gives

$$q^{-4} + C_{\text{err}} \frac{1+tD}{q} + C_{\text{light}} \frac{D+|H|}{q} \leq \Delta/20$$

for all sufficiently large  $n$ , uniformly over the later choice of  $t \leq C_{\text{PO}}/\delta$ . Now fix a labeling  $\pi$  with  $\text{val}_{I_\Phi}(\pi) \geq \Delta$ . By (2), the outer-edge satisfaction probability obeys

$$p_{\text{out}} \geq \Delta - q^{-4} \geq 19\Delta/20.$$

Applying Theorem 8.3 to the augmented pair  $(\bar{\varphi}_\pi, \bar{O}_\pi)$  from the soundness proof, with parameter  $\delta$ , yields a list of size

$$t \leq \frac{20C_{\text{PO}}}{\Delta}$$

and tuples  $\vec{Q}_1, \dots, \vec{Q}_t$  whose scores satisfy

$$\sum_{i=1}^t s_{\Phi, \pi}(\vec{Q}_i) \geq p_{\text{out}} - \delta - C_{\text{err}} \frac{1+tD}{q} \geq \frac{4\Delta}{5}.$$

Therefore some tuple in the list has score at least

$$\frac{4\Delta/5}{20C_{\text{PO}}/\Delta} = \frac{\Delta^2}{25C_{\text{PO}}}.$$

After enlarging  $C_{\text{hit}}$  once more, the bad-tuple bound from Lemma 8.11 is strictly smaller than

$$\theta_- := \frac{\Delta^2}{200C_{\text{PO}}},$$

so every tuple with score at least  $\theta_-$  is good. Fix the threshold interval

$$I_\Delta := [\theta_-, \theta_+], \quad \theta_+ := 2\theta_-,$$

and let  $a_{\text{def}} \in \{0, 1\}^{\text{Var}(\Phi)}$  be an arbitrary fixed “default” assignment. Given a labeling  $\pi$ , the map  $\mathcal{H}_{\text{lc}, \Phi, \Delta}(\pi)$  samples

$$\theta \sim I_\Delta \quad \text{and} \quad (\Omega, \tau, x) \sim (\mathcal{O}, \mathbb{T}, L(\Omega))$$

uniformly. If the sampled left label is inadmissible or the sampled outer edge is unsatisfied, it outputs  $a_{\text{def}}$ . Otherwise, it considers the set of tuples  $\vec{Q}$  such that

$$s_{\Phi, \pi}(\vec{Q}) \geq \theta \quad \text{and} \quad \vec{Q}|_{L(\Omega)} \equiv O_{\pi}^{\text{alg}}(\Omega, \tau),$$

and outputs the Boolean assignment to  $\Phi$  induced from the lexicographically first such tuple via Claim 8.6; if this set is empty, it again outputs  $a_{\text{def}}$ . Because every score- $\theta$  tuple is good, every nondefault output of this procedure satisfies  $\Phi$ . The tuple found above has score at least  $\theta_+$ , so whenever the sampled outer edge lands in its score event, the output is nondefault and therefore satisfying. Hence

$$\Pr[\mathcal{H}_{\text{lc}, \Phi, \Delta}(\pi) \text{ satisfies } \Phi] \geq \theta_+ \geq \Delta^{C_{\text{hit}}}$$

after enlarging  $C_{\text{hit}}$  if necessary. This proves item (1).

For item (2), couple the two runs of  $\mathcal{H}_{\text{lc}, \Phi, \Delta}$ , one for  $\pi$  and one for  $\pi'$ , by using the same random threshold  $\theta$  and the same sampled outer edge  $(\Omega, \tau, x)$ . Let

$$\nu_{\text{fix}} := q^{-3m + C_{\text{drift}}}.$$

By Lemma 8.14, every tuple score changes by at most  $\nu_{\text{fix}}$  under a one-coordinate perturbation of the labeling. Under the coupling, the outputs can differ only if either the sampled outer edge is itself affected by the changed coordinate, or else some tuple crosses the threshold  $\theta$ . The first event has probability at most  $\nu_{\text{fix}}$  by the proof of Lemma 8.14. For the second event, note that only tuples with score at least  $\theta/2$  can matter. Because  $\theta \in I_{\Delta}$  and  $\theta_- = \Delta^2/(200C_{\text{PO}})$ , one further enlargement of the constant  $C_{\text{hit}}$  in the hypothesis ensures that

$$(\theta/2)^2 \geq 20C_{\text{ov}} \frac{D}{q}.$$

Therefore Lemma 8.13, applied with threshold  $\eta = \theta/2$ , shows that there are at most  $2C_{\text{pack}}/\theta$  such tuples. A union bound over these tuples therefore shows that the probability that some tuple crosses the threshold is at most

$$O\left(\frac{\nu_{\text{fix}}}{\theta^2}\right).$$

If neither event occurs, the set of tuples of score at least  $\theta/2$  consistent with the sampled chart is unchanged and the output is identical. Since every output assignment lives on  $|\text{Var}(\Phi)| = |H|^m \leq q^m$  coordinates, the expected Hamming distance is at most

$$q^m \cdot O\left(\nu_{\text{fix}} + \frac{\nu_{\text{fix}}}{\theta^2}\right) \leq q^{C_{\text{hit}}}$$

after enlarging  $C_{\text{hit}}$ . This proves item (2).

For item (3), couple the two runs in the same way. A one-clause swap changes only the admissibility bit on at most  $2q^{8m + c_{\text{lc}}}$  left vertices, so by Lemma 8.14 every tuple score changes by at most

$$\nu_{\text{swap}} := q^{-3m + C_{\text{drift}}}.$$

Exactly the same argument as above—using again that  $(\theta/2)^2 \geq 20C_{\text{ov}}D/q$  and hence the bound of Lemma 8.13 at threshold  $\theta/2$ —shows that the outputs can differ only with probability

$$O\left(\nu_{\text{swap}} + \frac{\nu_{\text{swap}}}{\theta^2}\right),$$

and multiplying by  $|\text{Var}(\Phi)| \leq q^m$  gives the bound  $q^{C_{\text{hit}}}$  after one final enlargement of  $C_{\text{hit}}$ . This proves item (3).  $\square$

## 9 Alphabet Reduction

This section shrinks the right alphabet of the label cover instances obtained from the left-predicate label cover under clause swaps in Section 8.1, while keeping enough locality to preserve the later sensitivity lower bound. Throughout this section, we refer to the input label cover instance to the alphabet-reduction step as the source instance. The main subtlety is that the recovery map for this step depends on the source instance. To recover the label of a right vertex  $v$  in the source instance, we look at the labels suggested by the neighboring left vertices through the maps  $f_{(u,v)}$  on edges incident to  $v$ , and sample from the resulting empirical distribution. Because this distribution depends both on the assignment to the reduced instance and on the maps  $f_{(u,v)}$  in the source instance, we must separately track perturbations of the assignment and swaps of the source instance itself.

After recalling the required coding-theoretic facts, we define the transformation  $\mathcal{T}_{\text{AR}}$  and its instance-specific recovery maps  $\mathcal{R}_{\text{AR},I}$ , then we prove the resulting bounds on size, degrees, alphabets, soundness, and the three locality quantities needed later in Section 11: transformation under source swaps, recovery under assignment perturbations, and recovery under source swaps. If the source instance carries left predicates in the sense of Definition 2.1, then the alphabet-reduction gadget simply transports them unchanged; to avoid clutter we suppress them from the tuple notation.

### 9.1 The Alphabet Reduction Gadget

A mapping  $\mathcal{C} : \Sigma \rightarrow \hat{\Sigma}^k$  is called a *code* with minimum relative distance  $1 - \delta$  if for every  $a \neq b \in \Sigma$  the strings  $\mathcal{C}(a)$  and  $\mathcal{C}(b)$  differ in at least  $(1 - \delta)k$  coordinates.

**Proposition 9.1** (Fact 5.5 of [20]). Suppose  $\mathcal{C} : \Sigma \rightarrow \hat{\Sigma}^k$  is a code with (relative) distance at least  $1 - \delta$ , and let  $\eta > 2\sqrt{\delta}$ . Then, for every word  $w \in \hat{\Sigma}^k$  there are at most  $2/\eta$  symbols  $a \in \Sigma$  such that  $\mathcal{C}(a)$  agrees with  $w$  on at least an  $\eta$  fraction of the coordinates.

**Proposition 9.2** (Remark 5.6 of [20]). For every  $0 < \delta < 1$  and alphabet  $\Sigma$ , there exists a code  $\mathcal{C} : \Sigma \rightarrow \hat{\Sigma}^k$  with relative distance  $1 - \delta$  where  $|\hat{\Sigma}| = O(\delta^{-2})$  and  $k = O(\delta^{-2} \log |\Sigma|)$ .

Later constructions will require a deterministic instantiation of such a code for a fixed source alphabet  $\Sigma$  and a fixed distance parameter  $\delta$ . To handle this, we will fix one such code once and for all. No algorithmic explicitness beyond the existence statement above is needed for that use.

Suppose  $\mathcal{C} : \Sigma_V \rightarrow \hat{\Sigma}_V^k$  is a code with (relative) distance  $1 - \eta^3$  for some  $\eta > 0$  satisfying the guarantees of Proposition 9.2 for  $\delta = \eta^3$ . We now describe our alphabet reduction transformation, which we denote by  $\mathcal{T}_{\text{AR}}$ . The transformation maps label cover instances  $I = (U, V, E, \Sigma_U, \Sigma_V, F)$  to label cover instances  $I' = (U, V \times [k], E', \Sigma_U, \hat{\Sigma}_V, F')$ , where the set of edges  $E'$  and the set of projections  $F'$  are defined as follows:  $E' = \{(u, (v, i)) \mid (u, v) \in E, i \in [k]\}$  and for each  $e = (u, (v, i)) \in E'$ , the function  $f'_e : \Sigma_U \rightarrow \hat{\Sigma}_V$  is defined as  $f'_e(a) = \mathcal{C}(f_{(u,v)}(a))_i$ .

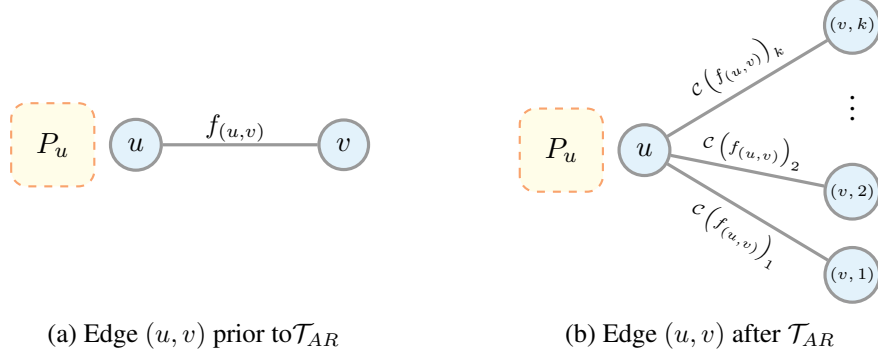


Figure 3: The transformation of the alphabet reduction procedure  $\mathcal{T}_{AR}$  on one edge.

For an instance  $I = (U, V, E, \Sigma_U, \Sigma_V, F)$  we write  $\bar{d}_U := \frac{|E|}{|U|}$ ,  $\bar{d}_V := \frac{|E|}{|V|}$ ,  $\Delta_U := \max_{u \in U} \deg(u)$ , and  $\delta_V := \min_{v \in V} \deg(v)$  for the average left and right degrees, the maximum left degree, and the minimum right degree, respectively.

For each fixed source instance  $I$  and assignment  $\pi'$  for  $I' = \mathcal{T}_{AR}(I)$ , define the randomized recovery map  $\mathcal{R}_{AR, I}(\pi')$ , which outputs a random assignment  $\pi$  for  $I$ . When the source instance is clear from context we suppress the subscript  $I$ .  $\mathcal{R}_{AR, I}(\pi')$  is defined as follows: for each  $v \in V$  and  $b \in \Sigma_V$ , define  $p_v(b) := \Pr_{e=(u,v) \text{ incident to } v} [f_{(u,v)}(\pi'(u)) = b]$ . Here the probability is over a uniformly random edge incident to  $v$  (counting multiplicity). We set

$$\pi(u) = \pi'(u) \quad \text{for every } u \in U.$$

For each  $v \in V$ , we then sample  $\pi(v)$  independently from the distribution  $p_v$ .<sup>13</sup>

**Lemma 9.3.** Let  $\mathcal{I}$  be a swap-closed family of label cover instances  $I = (U, V, E, \Sigma_U, \Sigma_V, F)$  with average left degree  $\bar{d}_U$ , average right degree  $\bar{d}_V$ , maximum left degree  $\Delta_U$ , and minimum right degree  $\delta_V$ , and let  $\mathcal{I}' = \mathcal{T}_{AR}(\mathcal{I})$ . Then for any  $\varepsilon > 0$  and  $\eta \in (0, 1/4)$ , the right alphabet of instances in  $\mathcal{I}'$  is  $\Sigma_V$  and satisfies  $|\hat{\Sigma}_V| = O(\eta^{-6})$ . Moreover, writing

$$s_{AR}(\varepsilon, \eta) := \frac{\varepsilon}{\sqrt{\eta}} + 2\sqrt{\eta} \quad \text{and} \quad k := O(\eta^{-6} \log |\Sigma_V|),$$

the following hold.

1. If  $I \in \mathcal{I}$  is satisfiable, then  $\mathcal{T}_{AR}(I)$  is satisfiable.
2. If  $I \in \mathcal{I}$  and an assignment  $\pi'$  for  $\mathcal{T}_{AR}(I)$  satisfies  $\text{val}_{\mathcal{T}_{AR}(I)}(\pi') \geq s_{AR}(\varepsilon, \eta)$ , then the recovered assignment  $\pi = \mathcal{R}_{AR, I}(\pi')$  satisfies  $\mathbf{E}[\text{val}_I(\pi)] \geq \varepsilon$ .
3. For all  $I, \tilde{I} \in \mathcal{I}$ , we have

$$\text{SwapDist}(\mathcal{T}_{AR}(I), \mathcal{T}_{AR}(\tilde{I})) \leq k \cdot \text{SwapDist}(I, \tilde{I}).$$

4. For every fixed  $I \in \mathcal{I}$  and every two assignments  $\pi', \tilde{\pi}'$  for  $\mathcal{T}_{AR}(I)$ , we have

$$\text{EMD}(\mathcal{R}_{AR, I}(\pi'), \mathcal{R}_{AR, I}(\tilde{\pi}')) \leq \left(1 + \frac{\Delta_U}{\delta_V}\right) d_H(\pi', \tilde{\pi}').$$

<sup>13</sup>This randomized recovery is sufficient for our sensitivity analysis; choosing a maximizer instead would only sharpen the constants.

5. If  $I, \tilde{I} \in \mathcal{I}$  differ by one swap, then for every assignment  $\pi'$  on the common domain of  $\mathcal{T}_{\text{AR}}(I)$  and  $\mathcal{T}_{\text{AR}}(\tilde{I})$ , we have  $\text{EMD}(\mathcal{R}_{\text{AR},I}(\pi'), \mathcal{R}_{\text{AR},\tilde{I}}(\pi')) \leq \frac{1}{\delta_V}$ .

*Proof.* The size bound on  $\hat{\Sigma}_V$  follows from Proposition 9.2 applied with  $\delta = \eta^3$ . For completeness, if  $I$  is satisfiable with witness  $\pi$ , then the assignment  $\pi'$  defined by  $\pi'(u) = \pi(u)$  for  $u \in U$  and  $\pi'(v, i) = \mathcal{C}(\pi(v))_i$  for  $(v, i) \in V \times [k]$  satisfies every edge of  $\mathcal{T}_{\text{AR}}(I)$ .

We next analyze Item 2. It can be easily checked that the fraction of edges satisfied by the assignment  $\pi'$  is  $\text{val}_{I'}(\pi') = \sum_{v \in V} \frac{\deg(v)}{|E|} \alpha_v$ , where  $\alpha_v := \sum_{b \in \Sigma_V} p_v(b) \cdot \Pr_{i \sim [k]}[\pi'(v, i) = \mathcal{C}(b)_i]$ . Likewise, the expected fraction of edges satisfied by the recovered assignment  $\pi$  is  $\mathbf{E}[\text{val}_I(\pi)] = \sum_{v \in V} \frac{\deg(v)}{|E|} \nu_v$ , where  $\nu_v := \sum_{b \in \Sigma_V} p_v(b)^2$ .

It suffices to show that for each  $v \in V$ , we have  $\alpha_v \leq \nu_v/\sqrt{\eta} + 2\sqrt{\eta}$ . Fix a vertex  $v \in V$ . Define  $\text{list}(v) = \{b \in \Sigma_V \mid \Pr_{i \sim [k]}[\pi'(v, i) = \mathcal{C}(b)_i] \geq \eta\}$ . From Proposition 9.1, we have that  $l := |\text{list}(v)| \leq 2\eta^{-1}$  since  $\eta > 2\eta^{3/2}$  for  $\eta < 1/4$ .

We have

$$\begin{aligned} \alpha_v &= \sum_{b \in \text{list}(v)} p_v(b) \cdot \Pr_i[\pi'(v, i) = \mathcal{C}(b)_i] + \sum_{b \notin \text{list}(v)} p_v(b) \cdot \Pr_i[\pi'(v, i) = \mathcal{C}(b)_i] \\ &\leq \sqrt{\nu_v \cdot \sum_{b \in \text{list}(v)} \Pr_i[\pi'(v, i) = \mathcal{C}(b)_i]} + \eta \\ &\leq \sqrt{\nu_v \cdot \left( \Pr_i[\exists b \in \text{list}(v), \pi'(v, i) = \mathcal{C}(b)_i] + \sum_{b_1 \neq b_2 \in \text{list}(v)} \Pr_i[\mathcal{C}(b_1)_i = \mathcal{C}(b_2)_i] \right)} + \eta \\ &\leq \sqrt{\nu_v \cdot \left( 1 + \binom{l}{2} \eta^3 \right)} + \eta \leq (1 + \eta)\sqrt{\nu_v} + \eta \leq \frac{\nu_v}{\sqrt{\eta}} + 2\sqrt{\eta}. \end{aligned}$$

The first inequality bounds the contribution of  $\text{list}(v)$  by Cauchy–Schwarz, using  $\sum_b p_v(b)^2 = \nu_v$  and  $\Pr_i[\pi'(v, i) = \mathcal{C}(b)_i] \leq 1$ , while the second term is at most  $\eta$  because  $b \notin \text{list}(v)$  implies  $\Pr_i[\pi'(v, i) = \mathcal{C}(b)_i] < \eta$ . The next line applies the union bound together with inclusion–exclusion; the third uses that  $\mathcal{C}$  has distance at least  $1 - \eta^3$ . For  $\eta \in (0, 1/4)$  we have  $\binom{l}{2} \eta^3 \leq 2\eta$  and  $(1 + \eta)\sqrt{\nu_v} \leq \nu_v/\sqrt{\eta} + \sqrt{\eta}$ ; adding the final  $\eta$  term and using  $\eta \leq \sqrt{\eta}$  yields  $\alpha_v \leq \nu_v/\sqrt{\eta} + 2\sqrt{\eta}$ . Averaging over  $v$  therefore gives  $\text{val}_{I'}(\pi') \leq \frac{\mathbf{E}[\text{val}_I(\pi)]}{\sqrt{\eta}} + 2\sqrt{\eta}$ , and hence  $\text{val}_{I'}(\pi') \geq s_{\text{AR}}(\varepsilon, \eta)$  implies  $\mathbf{E}[\text{val}_I(\pi)] \geq \varepsilon$ .

Item 3 follows because a single source projection swap on an edge  $(u, v)$  changes exactly the  $k$  transformed projections indexed by  $(u, (v, i))$  for  $i \in [k]$ , while a source left-predicate swap changes exactly one transformed left predicate. Applying this step-by-step along a shortest swap path from  $I$  to  $\tilde{I}$  gives the stated bound.

For Item 4, by decomposing the Hamming disagreement into single-coordinate changes and using the triangle inequality for EMD, it suffices to treat the case  $d_H(\pi', \tilde{\pi}') = 1$ . If  $\pi'$  and  $\tilde{\pi}'$  disagree on a right vertex, then the recovery map never reads that label, so  $\mathcal{R}_{\text{AR},I}(\pi') = \mathcal{R}_{\text{AR},I}(\tilde{\pi}')$  identically.

Suppose instead that  $\pi'$  and  $\tilde{\pi}'$  disagree on a left vertex  $u \in U$  with degree  $\deg(u)$ . Then the recovered assignments disagree on  $u$  itself. For any neighbor  $v$  of  $u$ , changing  $\pi'(u)$  modifies the empirical distribution  $p_v$  by moving mass  $1/\deg(v)$  from one label to another, so there is a coupling under which the contribution of  $v$  to the expected Hamming distance is at most  $1/\deg(v) \leq 1/\delta_V$ . Summing over the  $\deg(u)$  neighbors

Table 2: Parameters after the alphabet-then-degree reduction with  $0 < \varepsilon < 1/16$ ,  $\eta = \varepsilon/4$ ,  $k = \Theta(\varepsilon^{-6} \log |\Sigma_V|)$ , and  $d = \Theta(\varepsilon^{-1})$ . Degrees are averages unless indicated otherwise.

Parameter	$I$	alphabet $\rightarrow \sigma$	degree $\rightarrow d$
# left vertices	$n$	$n$	$n$
# right vertices	$m$	$\Theta(\varepsilon^{-6} m \log  \Sigma_V )$	$\Theta(\varepsilon^{-6} m \bar{d}_V \log  \Sigma_V )$
Average left degree	$\bar{d}_U$	$\Theta(\varepsilon^{-6} \bar{d}_U \log  \Sigma_V )$	$\Theta(\varepsilon^{-7} \bar{d}_U \log  \Sigma_V )$
Average right degree	$\bar{d}_V$	$\bar{d}_V$	$\Theta(\varepsilon^{-1})$
Left alphabet	$\Sigma_U$	$\Sigma_U$	$\Sigma_U$
Right alphabet	$\Sigma_V$	$\sigma$ ( $ \sigma  = \Theta(\varepsilon^{-6})$ )	$\sigma$
Soundness	$\varepsilon$	$3\sqrt{\varepsilon}$	$4\sqrt{\varepsilon}$

of  $u$  shows that

$$\text{EMD}(\mathcal{R}_{\text{AR},I}(\pi'), \mathcal{R}_{\text{AR},I}(\tilde{\pi}')) \leq 1 + \frac{\deg(u)}{\delta_V} \leq 1 + \frac{\Delta_U}{\delta_V}.$$

Finally, we prove Item 5. Let  $I$  and  $\tilde{I}$  differ by one swap, and fix an assignment  $\pi'$  on the common transformed domain. If the swap changes a left predicate, then every distribution  $p_v$  is unchanged, so the recovered assignments are identical. Otherwise the swap changes exactly one source projection, say on an edge  $(u, v)$ . Then the recovered left labels are still identical, and for every  $w \neq v$  the distributions  $p_w$  agree. At the single affected right vertex  $v$ , the swap changes the empirical distribution by moving mass  $1/\deg(v)$  from one label to another, so the total-variation distance between the two distributions is exactly  $1/\deg(v)$  and hence the one-coordinate earth mover's distance is at most  $1/\deg(v)$ . Therefore

$$\text{EMD}(\mathcal{R}_{\text{AR},I}(\pi'), \mathcal{R}_{\text{AR},\tilde{I}}(\pi')) \leq \frac{1}{\deg(v)} \leq \frac{1}{\delta_V}. \quad \square$$

The point of Lemma 9.3 is that the alphabet-reduction step yields both a bound for how the recovery changes under assignment perturbations and a separate bound for how the recovery changes under source swaps. Because  $\mathcal{R}_{\text{AR},I}$  depends on  $I$  through the distributions  $p_v$ , these two bounds are not interchangeable: the first only controls perturbations of assignments on one transformed instance, whereas Item 5 controls what happens when the source instance itself changes.

## 9.2 Reduced Label Cover

In this section, we summarize the combined effect of the applying alphabet reduction and then degree reduction on a label cover instance. For the soundness thresholds recorded below, we henceforth assume  $0 < \varepsilon < 1/16$ , so that  $3\sqrt{\varepsilon}$  and  $4\sqrt{\varepsilon}$  both lie in  $(0, 1)$ . Let  $I = (U, V, E, \Sigma_U, \Sigma_V, F)$  be a label cover instance with  $|U| = n$ ,  $|V| = m$ , average left degree  $\bar{d}_U$ , average right degree  $\bar{d}_V$ , maximum right degree  $\Delta_V$ , maximum left degree  $\Delta_U$ , and minimum right degree  $\delta_V$ . Table 2 tracks how these parameters change when we first apply the alphabet-reduction map induced by a code  $\mathcal{C} : \Sigma_V \rightarrow \sigma^k$  of relative distance  $1 - \eta^3$ , and then run the degree-reduction step based on  $[D_v, d, \lambda]$ -expanders tailored to the right-degree  $D_v$  of each vertex  $v \in V$ .

We instantiate  $\mathcal{C}$  using Lemma 9.3 (alphabet reduction) with  $\eta = \varepsilon/4$  (which in turn appeals to Proposition 9.2), yielding a right alphabet  $\sigma$  of size  $\Theta(\varepsilon^{-6})$  and block length  $k = \Theta(\varepsilon^{-6} \log |\Sigma_V|)$ , while the alphabet-reduction soundness threshold becomes  $3\sqrt{\varepsilon}$ . For the degree reduction, we take  $\lambda = \sqrt{\varepsilon}$ , choose  $d = \Theta(\varepsilon^{-1})$ , and for each right-degree value  $D$  that occurs in the common source graph we fix once and

for all the explicit  $[D, \Theta(\varepsilon^{-1}), \sqrt{\varepsilon}]$ -expander returned by Proposition 4.2; every right vertex of degree  $D$  uses that same expander. With these fixed choices, the degree-reduction recovery depends only on the transformed assignment and the fixed expanders, so the only source-instance dependence in the combined alphabet/degree-reduction recovery comes from the alphabet-reduction step.

**Lemma 9.4** (Alphabet-then-degree reduction). Let  $0 < \varepsilon < 1/16$  and  $\mathcal{I}$  be a swap-closed family of label cover instances  $I = (U, V, E, \Sigma_U, \Sigma_V, F)$  with  $|U| = n$ ,  $|V| = m$ , average left degree  $\bar{d}_U$ , average right degree  $\bar{d}_V$ , maximum right degree  $\Delta_V$ , maximum left degree  $\Delta_U$ , and minimum right degree  $\delta_V$ . For each  $I \in \mathcal{I}$ , let  $\text{red}_\varepsilon^\circ(I)$  denote the output obtained by first applying the alphabet-reduction map with  $\eta = \varepsilon/4$  and then, if  $\delta_V > c/\varepsilon$ , applying the degree-reduction map with  $d = \lceil c/\varepsilon \rceil$ ; otherwise we stop after alphabet reduction. Then every instance  $\text{red}_\varepsilon^\circ(I) = (U', V', E', \Sigma'_L, \Sigma'_R, F')$  satisfies

$$|U'| = n, \quad |V'| = O(\varepsilon^{-6} m \bar{d}_V \log |\Sigma_V|), \quad \bar{d}_{U'} = O(\bar{d}_U \varepsilon^{-7} \log |\Sigma_V|),$$

has maximum right degree  $O(\max\{\Delta_V, \varepsilon^{-1}\})$ , and has right alphabet  $\sigma$  of size at most  $O(\varepsilon^{-6})$ . Moreover, if  $R_{\varepsilon, I}^{\text{red}}$  denotes the natural recovery map from assignments on  $\text{red}_\varepsilon^\circ(I)$  back to assignments on  $I$ , then the following hold.

1. If  $I$  is satisfiable, then  $\text{red}_\varepsilon^\circ(I)$  is satisfiable.
2. If an assignment  $\pi'$  for  $\text{red}_\varepsilon^\circ(I)$  satisfies  $\text{val}_{\text{red}_\varepsilon^\circ(I)}(\pi') \geq 4\sqrt{\varepsilon}$ , then  $\mathbf{E}[\text{val}_I(R_{\varepsilon, I}^{\text{red}}(\pi'))] \geq \varepsilon$ .
3. For all  $I, \tilde{I} \in \mathcal{I}$ , we have

$$\text{SwapDist}(\text{red}_\varepsilon^\circ(I), \text{red}_\varepsilon^\circ(\tilde{I})) \leq O(\varepsilon^{-7} \log |\Sigma_V|) \cdot \text{SwapDist}(I, \tilde{I}).$$

4. For every fixed  $I \in \mathcal{I}$  and every two assignments  $\pi', \tilde{\pi}'$  on  $\text{red}_\varepsilon^\circ(I)$ , we have

$$\text{EMD}(R_{\varepsilon, I}^{\text{red}}(\pi'), R_{\varepsilon, I}^{\text{red}}(\tilde{\pi}')) \leq \left(1 + \frac{\Delta_U}{\delta_V}\right) d_H(\pi', \tilde{\pi}').$$

5. If  $I, \tilde{I} \in \mathcal{I}$  differ by one swap, then for every assignment  $\pi'$  on the common domain of  $\text{red}_\varepsilon^\circ(I)$  and  $\text{red}_\varepsilon^\circ(\tilde{I})$ , we have

$$\text{EMD}(R_{\varepsilon, I}^{\text{red}}(\pi'), R_{\varepsilon, \tilde{I}}^{\text{red}}(\pi')) \leq \frac{1}{\delta_V}.$$

*Proof.* Fix  $0 < \varepsilon < 1/16$  and set  $\eta := \varepsilon/4$ . For every instance  $I \in \mathcal{I}$  let  $I^{(1)} := \mathcal{T}_{\text{AR}}(I)$ . Write  $\mathcal{I}^{(1)} := \mathcal{T}_{\text{AR}}(\mathcal{I})$ . Since all instances in  $\mathcal{I}$  share the same underlying graph, they have the same value of  $\delta_V$ , so the branch condition  $\delta_V \leq c/\varepsilon$  versus  $\delta_V > c/\varepsilon$  is uniform across the family. The code has block length  $k = \Theta(\eta^{-6} \log |\Sigma_V|) = \Theta(\varepsilon^{-6} \log |\Sigma_V|)$  and target alphabet size  $|\sigma| = \Theta(\eta^{-6}) = \Theta(\varepsilon^{-6})$ . Consequently, every  $I^{(1)}$  satisfies:  $|U^{(1)}| = |U| = n$ ,  $|V^{(1)}| = mk$ , average left degree  $\bar{d}_U^{(1)} = k\bar{d}_U$ , average right degree  $\bar{d}_V^{(1)} = \bar{d}_V$ , maximum right degree  $\Delta_V^{(1)} = \Delta_V$ , minimum right degree  $\delta_V^{(1)} = \delta_V$ , and right alphabet  $\sigma$ .

By Lemma 9.3, the map  $\mathcal{T}_{\text{AR}}$  with recovery  $\mathcal{R}_{\text{AR}, I}$  has soundness threshold  $s_{\text{AR}}(\varepsilon, \eta) = 3\sqrt{\varepsilon}$ . It also has a factor- $k$  bound for how the transformation changes under source swaps, an assignment-perturbation constant  $(1 + \Delta_U/\delta_V)$ , and a source-swap recovery constant  $1/\delta_V$ .

If  $\delta_V \leq c/\varepsilon$  for a sufficiently large absolute constant  $c$ , we simply set

$$\begin{aligned} \text{red}_\varepsilon^\circ(I) &:= I^{(1)}, \\ R_{\varepsilon, I}^{\text{red}} &:= \mathcal{R}_{\text{AR}, I}. \end{aligned}$$

Then all five items follow immediately from the preceding paragraph and from the size, degree, and alphabet bounds recorded above, since  $k = O(\varepsilon^{-6} \log |\Sigma_V|) \leq O(\varepsilon^{-7} \log |\Sigma_V|)$ .

Suppose now that  $\delta_V > c/\varepsilon$  and proceed to the degree-reduction branch. Let  $d := \lceil c/\varepsilon \rceil$ ; by choice of the branch we have  $d \leq \delta_V$ . The family  $\mathcal{I}^{(1)}$  has a common underlying graph and common alphabets, and the expander choices for each right-degree value in that graph were fixed above. Thus the degree-reduction gadget from Section 7 applies uniformly to the instances  $I^{(1)}$ . Let  $I^{(2)} := \mathcal{T}_{\text{DR}}(I^{(1)}; d)$ , and let  $R_{\text{DR}}$  be the degree-reduction recovery map. Define

$$\begin{aligned} \text{red}_\varepsilon^\circ(I) &:= I^{(2)}, \\ R_{\varepsilon, I}^{\text{red}} &:= \mathcal{R}_{\text{AR}, I} \circ R_{\text{DR}}. \end{aligned}$$

Each instance  $I^{(2)}$  satisfies

$$|U^{(2)}| = n, \quad |V^{(2)}| = k|E| = km\bar{d}_V, \quad \bar{d}_V^{(2)} = dk\bar{d}_U.$$

Moreover, it has right degree  $d = \Theta(\varepsilon^{-1})$  and right alphabet  $\sigma$ . Hence the same size, degree, and alphabet bounds claimed in the statement hold in this branch as well.

We next verify the five items. Completeness is preserved by both alphabet reduction and degree reduction, proving Item 1. For Item 2, the proof of Lemma 7.1 gives an additive soundness loss of  $O(d^{-1/2})$  when recovering from  $I^{(2)}$  to  $I^{(1)}$ . Since  $d = \Theta(\varepsilon^{-1})$ , this loss is  $O(\sqrt{\varepsilon})$ ; choosing the absolute constant  $c$  large enough ensures that every assignment of value at least  $4\sqrt{\varepsilon}$  on  $I^{(2)}$  recovers via  $R_{\text{DR}}$  to an assignment of value at least  $3\sqrt{\varepsilon}$  on  $I^{(1)}$ . Applying the alphabet-reduction soundness guarantee then yields  $\mathbf{E}[\text{val}_I(R_{\varepsilon, I}^{\text{red}}(\pi'))] \geq \varepsilon$ .

For Item 3, alphabet reduction contributes a factor  $k$ , and degree reduction contributes a further factor  $d$ . Therefore  $\text{SwapDist}(\text{red}_\varepsilon^\circ(I), \text{red}_\varepsilon^\circ(\tilde{I})) \leq O(kd) \cdot \text{SwapDist}(I, \tilde{I}) = O(\varepsilon^{-7} \log |\Sigma_V|) \cdot \text{SwapDist}(I, \tilde{I})$ .

For Item 4, the proof of Lemma 7.1 gives assignment-perturbation constant 1 for  $R_{\text{DR}}$ . Composing that with Item 4 of Lemma 9.3 yields  $\text{EMD}(R_{\varepsilon, I}^{\text{red}}(\pi'), R_{\varepsilon, I}^{\text{red}}(\tilde{\pi}')) \leq \left(1 + \frac{\Delta_U}{\delta_V}\right) d_{\text{H}}(\pi', \tilde{\pi}')$ .

Finally, for Item 5 we use the same fixed expanders on both  $I^{(1)}$  and  $\tilde{I}^{(1)}$ . Under these fixed choices,  $R_{\text{DR}}$  depends only on the transformed assignment and on the fixed expanders, so it contributes nothing to the bound for how the recovery changes under source swaps. Therefore the only source-swap contribution comes from the alphabet-reduction recovery map. Item 5 of Lemma 9.3 gives

$$\text{EMD}(R_{\varepsilon, I}^{\text{red}}(\pi'), R_{\varepsilon, \tilde{I}}^{\text{red}}(\pi')) \leq \frac{1}{\delta_V}.$$

This completes the proof.  $\square$

In particular, Lemma 9.4 records the alphabet-then-degree reduction as an output map together with its soundness threshold, the bound for how the transformation changes under source swaps, the bound for how the recovery changes under assignment perturbations, and the bound for how the recovery changes under source swaps. These are the data used later in the framework of Section 11, which separately tracks how the recovery map changes under assignment perturbations and under source swaps.

## 10 Decodable PCPs and Composition

This section develops the decodable PCP ingredient used in the low-soundness label cover construction and proves the composition theorem that is applied at each stage of the iteration in Section 11. Starting from a

suitable outer label cover instance and a decodable PCP (dPCP) decoder, we construct the next label cover instance in the iteration and record the completeness, soundness, and stability properties needed there. In this section, the “source” instance is the outer label cover instance fed into the composition, so “source swaps” mean swaps on that outer instance. We state the result for left-predicate label cover instances so that the same theorem applies before and after swaps. In the special case of ordinary label cover, where all predicates are identically 1, the argument is identical.

## 10.1 Decodable PCPs

For a finite alphabet  $\Sigma$ , we let  $\text{CircuitSAT}_\Sigma$  denote the satisfiability problem for circuits whose inputs range over  $\Sigma$ : given a circuit  $C : \Sigma^k \rightarrow \{0, 1\}$ , the instance is a Yes instance if there exists  $y \in \Sigma^k$  with  $C(y) = 1$  and a No instance otherwise. Any such vector  $y$  is called a *satisfying assignment* for  $C$ . Throughout this section, the *decision complexity* of a circuit means the size of its given circuit representation. In particular, we do not minimize over equivalent circuits, and we use the same size measure for the local decoding maps and admissible-neighborhood circuits below.

Informally, a PCP decoder is given a circuit  $C$ , an index  $j \in [k]$ , and oracle access to an auxiliary proof string  $\pi \in \sigma^{m(n)}$ , where  $n$  denotes the decision complexity of  $C$ . The proof is simply an additional string over the proof alphabet  $\sigma$ ; it is not part of the input circuit, and for our purposes we only need the existence of such a string in the completeness case. Using  $r(n)$  random coins, the decoder chooses  $q(n)$  locations of  $\pi$  to query and then applies a local decoding rule of complexity  $s(n)$  to the queried symbols, with the goal of recovering the target symbol  $y_j$ . Thus  $m(n)$  is the proof length, while  $q(n)$ ,  $r(n)$ , and  $s(n)$  are the query complexity, randomness complexity, and local-decoder complexity, respectively.

**Definition 10.1** (PCP decoder). Fix a proof alphabet  $\sigma$ . A PCP decoder for  $\text{CircuitSAT}_\Sigma$  is a probabilistic polynomial-time algorithm  $\mathcal{D}$  that, on input a circuit  $C : \Sigma^k \rightarrow \{0, 1\}$  whose decision complexity is  $n$  and an index  $j \in [k]$ , uses  $r(n)$  random coins and outputs:

1. A tuple of query locations  $I = (i_1, \dots, i_q) \in [m(n)]^q$  for some  $q = q(n)$  and proof length  $m(n)$ , and
2. A deterministic local decoding map  $f : \sigma^q \rightarrow \Sigma \cup \{\perp\}$  whose decision complexity is at most  $s(n)$ .

After reading the proof symbols at positions  $I$ , the decoder returns  $f(\pi_I)$ , where  $\pi \in \sigma^{m(n)}$  is the purported proof and  $\pi_I := (\pi_{i_1}, \dots, \pi_{i_q})$  is the queried substring.

**Definition 10.2** (Decodable PCP). Let  $\delta : \mathbb{Z}_{>0} \rightarrow [0, 1]$  and  $L : \mathbb{Z}_{>0} \rightarrow \mathbb{Z}_{>0}$ . A PCP decoder  $\mathcal{D}$  for  $\text{CircuitSAT}_\Sigma$  is a decodable PCP (dPCP) with soundness error  $\delta$  and list size  $L$  if the following conditions hold for every circuit  $C : \Sigma^k \rightarrow \{0, 1\}$ .

- **Completeness:** For every satisfying assignment  $y \in \Sigma^k$  there exists a proof  $\pi \in \sigma^{m(n)}$  such that, when  $j \in [k]$  is chosen uniformly at random and the randomness of  $\mathcal{D}$  is taken over  $(I, f)$ , we have  $\Pr_{j,I,f}[f(\pi_I) = y_j] = 1$ .
- **List-decoding soundness:** For every proof  $\pi \in \sigma^{m(n)}$  there exists a list  $Y = \{y^{(1)}, \dots, y^{(\ell)}\}$  of satisfying assignments for  $C$  with  $\ell \leq L(n)$  such that

$$\Pr_{j,I,f}[f(\pi_I) \in \{\perp, y_j^{(1)}, \dots, y_j^{(\ell)}\}] \geq 1 - \delta(n). \quad (4)$$

- **Robust soundness:** Define  $\text{Good}_Y(f, j) := \{w \in \sigma^q : f(w) \in \{\perp, y_j^{(1)}, \dots, y_j^{(\ell)}\}\}$ . For  $w \in \sigma^q$  and  $S \subseteq \sigma^q$ , let  $d(w, S) := \min_{w' \in S} \frac{1}{q} \cdot d_H(w, w')$  be the (normalized) Hamming distance of  $w$

to  $S$ . The decoder  $\mathcal{D}$  is *robust* with error  $\delta$  if the list from the previous item can be chosen so that for every proof  $\pi$ ,  $\mathbf{E}_{j,I,f} [d(\pi_I, \text{Goody}_Y(f, j))] \leq \delta(n)$ . Equivalently, on average (over the decoder's randomness and the choice of  $j$ ), one must modify at most a  $\delta(n)$ -fraction of the queried symbols to make the local view decode to  $\perp$  or to a value consistent with the list.

**Lemma 10.3** (Theorem 6.5 of [20]). There exist constants  $a_1, a_2, \alpha, \gamma > 0$  such that for every  $\delta \geq n^{-\alpha}$  and input alphabet  $\Sigma$  of size at most  $n^\gamma$ ,  $\text{CircuitSAT}_\Sigma$  has a dPCP decoder with soundness error  $\delta$  and list size  $L \leq 2/\delta$ , query complexity  $n^{1/8}$ , proof alphabet size at most  $n^\gamma$ , proof length at most  $n^{a_1}$ , local-decoder decision complexity at most  $n^{a_1}$ , and randomness complexity  $a_2 \log n$ .

## 10.2 Composition

We next adapt the composition framework of Dinur and Harsha [20] to left-predicate label cover instances. The key point is that the inner decoder should see only the set of admissible tuples of neighboring right labels at each outer left vertex, rather than the full left alphabet  $\Sigma_U$ . This is captured by the following notion of local description complexity.

**Definition 10.4** (Admissible Neighborhood Tuples and Local Description Complexity). Let

$$I = (U, V, E, \Sigma_U, \Sigma_V, \{P_u\}_{u \in U}, F = \{f_e\}_{e \in E})$$

be a left-predicate label cover instance. For every  $u \in U$ , fix an ordering

$$N_I(u) = (v_1, \dots, v_{\deg_I(u)})$$

of the neighbors of  $u$ . A tuple  $(b_1, \dots, b_{\deg_I(u)}) \in \Sigma_V^{\deg_I(u)}$  is *admissible at  $u$*  if there exists  $a \in \Sigma_U$  such that  $P_u(a) = 1$  and  $f_{(u, v_t)}(a) = b_t$  for every  $t \in [\deg_I(u)]$ . The *admissible neighborhood set* at  $u$  is the set of all admissible tuples, denoted by

$$\mathcal{N}_I(u) \subseteq \Sigma_V^{\deg_I(u)}.$$

We say that  $I$  has *local description complexity at most  $S$*  if, for every  $u \in U$ , there exists a circuit  $C_u : \Sigma_V^{\deg_I(u)} \rightarrow \{0, 1\}$  of decision complexity at most  $S$  such that  $C_u^{-1}(1) = \mathcal{N}_I(u)$ . A family has local description complexity at most  $S$  if every instance in the family does.

At a high level, the composition procedure proceeds as follows; an illustration is given in Figure 4 and a dPCP decoder  $\mathcal{D}$ . Fix an outer instance  $I = (U, V, E, \Sigma_U, \Sigma_V, \{P_u\}_{u \in U}, F)$ . For each outer right vertex  $v \in V$  and decoder randomness string  $r \in \{0, 1\}^r$ , the composition creates one composed left vertex  $(v, r)$ , and for each outer left vertex  $u \in U$  and proof position  $t \in [m]$ , it creates one composed right vertex  $(u, t)$ . A label of  $(v, r)$  is a  $d_V \times m$  matrix, where  $d_V$  is the right-degree of  $V$ , over the proof alphabet  $\sigma$ : the columns are proof positions, and the rows index the at most  $d_V$  neighbors of  $v$ . If  $u_1, \dots, u_{t(v)}$  are the actual neighbors of  $v$  in the outer instance, where  $t(v) := \deg_I(v)$ , then the rows of this matrix which are indexed by  $u_1, \dots, u_{t(v)}$  are the *active rows*, and the remaining rows are *unused rows*, which are added only so that every assignment has the same size. For every active row  $i$  and every proof position  $t \in [m]$ , we include an edge from  $(v, r)$  to  $(u_i, t)$ , and a projection constraint on this edge. To define the left-predicate for  $(v, t)$ , the decoder  $\mathcal{D}$  is run on  $(C_{u_i}, j_i)$  with randomness  $r$ , where  $C_{u_i}^{-1}(1) = \mathcal{N}_I(u_i)$  and  $v$  is the  $j_i$ -th neighbor of  $u_i$ . The output selects a few columns  $t_{i,1}, \dots, t_{i,q}$  from each active row of the matrix and uses these to define the left predicate for  $(v, r)$ . This means that the query pattern affects the left predicates, but not the edges of the resulting label cover instance.

---

**Algorithm 2:** Composition  $I \circledast \mathcal{D}$ 

---

**Input:** A left-predicate label cover instance  $I$  together with circuits  $\{C_u\}_{u \in U}$  deciding admissible neighborhood tuples, and a dPCP decoder  $\mathcal{D}$  for  $\text{CircuitSAT}_{\Sigma_V}$

**Output:** The composed instance  $I'$

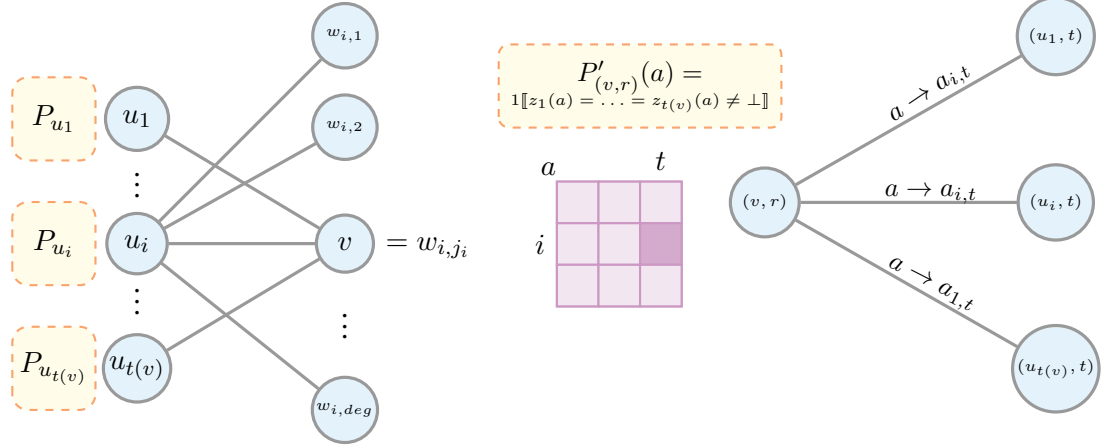
```
1  $U' \leftarrow V \times \{0, 1\}^r$ ;  
2  $V' \leftarrow U \times [m]$ ;  
3  $\Sigma'_U \leftarrow \sigma^{d_V m}$ ;  $\Sigma'_V \leftarrow \sigma$ ;  
4  $E' \leftarrow \emptyset$ ;  $F' \leftarrow \emptyset$ ;  
5 Interpret  $a \in \Sigma'_U$  as a  $d_V \times m$  matrix over  $\sigma$ ;  
6  $(a_{i,t})_{i \in [d_V], t \in [m]}$ ;  
7 for  $v \in V$  do  
8   Let  $t(v) := \deg_I(v)$  ;  
9   Let  $u_1, \dots, u_{t(v)}$  be the neighbors of  $v$  in  $I$ ;  
10  for  $i \in [t(v)]$  do  
11    Let  $N_I(u_i) = (w_{i,1}, \dots, w_{i, \deg_I(u_i)})$  be the fixed ordering of the neighbors of  $u_i$ ;  
12    Let  $j_i \in [\deg_I(u_i)]$  be such that  $w_{i,j_i} = v$ ;  
13  for  $r \in \{0, 1\}^r$  do  
14    for  $i \in [t(v)]$  do  
15      Run  $\mathcal{D}$  on input  $(C_{u_i}, j_i)$  using randomness  $r$ ;  
16      Obtain query tuple  $I_{i,r} = (t_{i,1}, \dots, t_{i,q}) \in [m]^q$ ;  
17      and local decoder  $g_{i,r} : \sigma^q \rightarrow \Sigma_V \cup \{\perp\}$ ;  
18    for  $i \in [t(v)]$  do  
19       $z_i(a) \leftarrow g_{i,r}(a_{i,t_{i,1}}, \dots, a_{i,t_{i,q}})$ ;  
20    If  $t(v) = 0$ , set  $P'_{(v,r)}(a) = 1$  for every  $a$ ;  
21    Otherwise, set  $P'_{(v,r)}(a) = 1$  iff  $z_1(a) = \dots = z_{t(v)}(a) \neq \perp$ ;  
22    for  $i \in [t(v)]$  do  
23      for  $t \in [m]$  do  
24        Add edge  $e = ((v, r), (u_i, t))$  to  $E'$ ;  
25        Define the projection  $f'_e : \Sigma'_U \rightarrow \Sigma'_V$  by  $f'_e(a) = a_{i,t}$ ;  
26        add  $f'_e$  to  $F'$ ;  
27 return  $I'$ 
```

---

The explicit construction is spelled out in Algorithm 2, and is shown pictorially for a vertex  $v$  in Figure 4. For a fixed outer underlying graph, the vertex and edge sets of the label cover instance obtained by applying composition depend only on that graph and on the decoder parameters, not on the query tuples. Instance by instance, the only part of the composed verifier that changes is the family of left predicates. Equivalently, this is a composition on a common padded graph: as in the overview above, the label given to each left vertex is a matrix where active rows are the rows corresponding to actual neighbors, unused rows are padding, and the coordinate-projection edges are fixed in advance across all instances in the family.

The following theorem provides the guarantees that we obtain from composition.

**Theorem 10.5** (Composition on a common padded graph). Let  $\mathcal{I}$  be a family of left-predicate label cover



(a) The vertices on which the composed-instance-transformation for right-vertex  $v$  depends on. For each  $i \in [\text{deg}(v)]$ , we run the decoder  $\mathcal{D}((C_{u_i}, j_i), r)$  to get  $(t_{i,1}, \dots, t_{i,q}, g_{i,r})$  after composition. and denote  $z_i(a) := g_{i,r}(a_{1,t_{i,1}}, \dots, a_{1,t_{i,q}})$ . (b) The resulting label cover instance at  $(v, t)$  after composition. The matrix represents the assignment to vertex  $(v, r)$ .

Figure 4: The conversion of a vertex  $v$  and its neighborhood under composition.

instances with local description complexity at most  $S$

$$I = (U, V, E, \Sigma_U, \Sigma_V, \{P_u\}_{u \in U}, F = \{f_e\}_{e \in E})$$

whose left degree is at most  $d_U$ , whose right degree is at most  $d_V$ , and whose right alphabet is  $\Sigma_V$ . Fix any target outer soundness parameter  $\Delta \in [0, 1]$ . Let  $N$  be a common upper bound, after padding, on the decision complexities (equivalently, circuit sizes) of the circuits deciding admissible neighborhood tuples from Definition 10.4. Thus, the earlier local-description parameter  $S$  enters only through the existence of this common bound  $N$ : once  $N$  is fixed, the decoder parameters depend only on  $N$  rather than on  $S$  separately. Suppose  $\text{CircuitSAT}_{\Sigma_V}$  at decision complexity  $N$  admits a dPCP decoder  $\mathcal{D}$  with query complexity  $q$ ,  $r$  bits of randomness, list size  $L$ , robust soundness error  $\delta$ , proof alphabet  $\sigma$ , proof length  $m$ , and local-decoder decision complexity  $s(N)$ . Let  $d$  be any upper bound on the corresponding *decoder query degree*, i.e. on

$$\max_{I \in \mathcal{I}} \max_{u \in U} \max_{t \in [m]} \left| \{(j, r, \ell) : \text{on input } (C_u, j) \text{ with randomness } r, \text{ slot } \ell \text{ queries } t\} \right|.$$

This counts only decoder calls in which the proof position  $t$  is actually queried. It should be compared with the graph degree of the right vertex  $(u, t)$  in Algorithm 2, where edges are added to *all* vertices  $t \in [m]$  for every active row. Thus  $(u, t)$  has graph degree  $\text{deg}_I(u)2^r$ , while  $d$  counts only the subset of those incidences coming from decoder calls that truly read proof location  $t$ .

For every  $I \in \mathcal{I}$  there is a composed left-predicate label cover instance  $I \otimes \mathcal{D}$ , and we write  $\mathcal{I} \otimes \mathcal{D} := \{I \otimes \mathcal{D} : I \in \mathcal{I}\}$ . Then the map  $\mathcal{T}_{\text{Comp}} : I \mapsto I \otimes \mathcal{D}$  and the instance-specific recovery maps  $\mathcal{R}_{\text{Comp}, I}$  defined below satisfy the following properties. Items (3)–(5) track, respectively, the composition-layer transformation parameter under source swaps, the fixed-source recovery Lipschitz parameter, and the source-swap recovery drift parameter. Define

$$C_{T, \text{Comp}} := d_U 2^r, \quad C_{R, \text{Comp}} := d(1 + d_V) 2^{-r}, \quad D_{\text{Comp}} := d_U(1 + d_V) + 1.$$

Table 3: Parameter evolution in the composition  $I \mapsto I \otimes \mathcal{D}$ .

Parameter	Outer instance $I$	Decoder $\mathcal{D}$	Composed $I \otimes \mathcal{D}$
# left vertices	$ U $	–	$ V  \cdot 2^r$
# right vertices	$ V $	proof length $m$	$ U  \cdot m$
Left degree	$\leq d_U$	proof length $m$	$\leq d_V m$
Right degree	$\leq d_V$	randomness $2^r$	$\leq d_U 2^r$
Left alphabet	$\Sigma_U$	–	$\sigma^{d_V m}$
Right alphabet	$\Sigma_V$	$\sigma$	$\sigma$
Local description complexity	at most $S$	local maps of size $s(N)$	$\text{poly}(d_V, m, s(N))$
Soundness	$\Delta$	$\delta$ and list size $L$	$\delta + L\Delta$

1. **Completeness.** If  $I$  is satisfiable, then  $I \otimes \mathcal{D}$  is satisfiable.
2. **Soundness implication.** If a (possibly random) assignment  $\pi'$  to  $I \otimes \mathcal{D}$  satisfies

$$\mathbf{E}[\text{val}_{I \otimes \mathcal{D}}(\pi')] \geq \delta + L\Delta,$$

then

$$\mathbf{E}[\text{val}_I(\mathcal{R}_{\text{Comp}, I}(\pi'))] \geq \Delta.$$

3. **Transformation under source swaps.** For all  $I, \tilde{I} \in \mathcal{I}$ ,

$$\text{SwapDist}(\mathcal{T}_{\text{Comp}}(I), \mathcal{T}_{\text{Comp}}(\tilde{I})) \leq C_{T, \text{Comp}} \text{SwapDist}(I, \tilde{I}).$$

4. **Recovery under assignment perturbations.** For every fixed outer instance  $I \in \mathcal{I}$  and every two assignments  $\pi', \tilde{\pi}'$  to  $\mathcal{T}_{\text{Comp}}(I)$ ,

$$\text{EMD}(\mathcal{R}_{\text{Comp}, I}(\pi'), \mathcal{R}_{\text{Comp}, I}(\tilde{\pi}')) \leq C_{R, \text{Comp}} d_H(\pi', \tilde{\pi}').$$

5. **Recovery under source swaps.** If  $I$  and  $\tilde{I}$  differ by one outer swap, then for every assignment  $\pi'$  to the common composed domain,

$$\text{EMD}(\mathcal{R}_{\text{Comp}, I}(\pi'), \mathcal{R}_{\text{Comp}, \tilde{I}}(\pi')) \leq D_{\text{Comp}},$$

where  $D_{\text{Comp}} = d_U(1 + d_V) + 1$ .

6. **Structural bounds.** For every  $I \in \mathcal{I}$ ,

$$|U(I \otimes \mathcal{D})| = |V(I)| \cdot 2^r, \quad |V(I \otimes \mathcal{D})| = |U(I)| \cdot m,$$

every left degree is at most  $d_V m$ , every right degree is at most  $d_U 2^r$ , the right alphabet is  $\sigma$ , the left alphabet is  $\sigma^{d_V m}$ , and the local description complexity is at most  $\text{poly}(d_V, m, s(N))$ .

Table 3 is stated after the decoder has been fixed, so the dependence on the input-side local description complexity has already been absorbed into the choice of the common bound  $N$ ; this is why the composed local description complexity is recorded in terms of  $d_V$ ,  $m$ , and  $s(N)$  rather than carrying  $S$  separately.

*Proof.* Write  $\mathcal{T}_{\text{Comp}}$  for Algorithm 2. Fix once and for all a total order  $\prec$  on  $\Sigma_U$  and a default symbol  $b_0 \in \Sigma_V$ . For each fixed outer instance  $I \in \mathcal{I}$ , let  $\mathcal{R}_{\text{Comp}, I}$  denote the following recovery map. Given a

labeling  $\pi' : U' \cup V' \rightarrow \Sigma'_U \cup \Sigma'_V$ , sample  $r \in \{0, 1\}^r$  uniformly and recompute the query tuples  $I_{i,r}$  and functions  $g_{i,r}$ . For each  $v \in V$ , set

$$\pi_V(v) := \begin{cases} b_0 & \text{if } \deg_I(v) = 0, \\ g_{1,r}(w_{1,r}) & \text{if } g_{1,r}(w_{1,r}) = \cdots = g_{t(v),r}(w_{t(v),r}) \neq \perp, \\ b_0 & \text{otherwise,} \end{cases}$$

where  $u_1, \dots, u_{t(v)}$  are the neighbors of  $v$ ,  $I_{i,r} = (t_{i,1}, \dots, t_{i,q})$  is the query tuple output by  $\mathcal{D}$  on  $(C_{u_i}, j_i)$  using randomness  $r$ , and

$$w_{i,r} := (\pi'_{(u_i, t_{i,1})}, \dots, \pi'_{(u_i, t_{i,q})}) \in \sigma^q$$

is the local view of the proof at those query locations. For each  $u \in U$ , let  $S_u$  be the set of all labels  $a \in \Sigma_U$  such that  $P_u(a) = 1$  and  $f_{(u, v_\ell)}(a) = \pi_V(v_\ell)$  for all  $\ell \in [\deg_I(u)]$ . Recover the left label by

$$\pi_U(u) := \min_{\prec} S_u,$$

if  $S_u$  is nonempty, and otherwise set  $\pi_U(u)$  to the  $\prec$ -least element of  $\Sigma_U$ . We write  $\mathcal{R}_{\text{Comp}, I}(\pi') = (\pi_U, \pi_V)$ .

The same Dinur–Harsha decoding/list-decoding proof from [20, Theorem 4.1] applies to the circuits  $\{C_u\}_{u \in U}$  deciding admissible neighborhood tuples. The only semantic change is that a satisfying assignment for  $C_u$  now means an admissible neighborhood labeling around  $u$ . Consequently, whenever the decoded neighborhood labels around a vertex  $u$  are consistent with a satisfying assignment of  $C_u$ , the recovered outer label  $\pi_U(u)$  is admissible and satisfies every incident projection. This proves items (1) and (2).

**Transformation under source swaps.** A single outer swap changes either one projection  $f_{(u,v)}$  or one left predicate  $P_u$ . In either case, only the circuit  $C_u$  deciding admissible neighborhood tuples changes. This circuit appears once for each neighbor of  $u$ , so there are at most  $\deg_I(u) \leq d_U$  affected composed left vertices for each randomness string. Since the composed edge set and the composed projections are fixed coordinate projections, no composed projection changes. Hence a single outer swap changes at most

$$d_U 2^r$$

composed left predicates, which proves item (3).

**Recovery under assignment perturbations.** Couple  $\mathcal{R}_{\text{Comp}, I}(\pi')$  and  $\mathcal{R}_{\text{Comp}, I}(\tilde{\pi}')$  by using the same sampled randomness  $r$ . Write

$$\mathcal{R}_{\text{Comp}, I}(\pi') = (\pi_U, \pi_V), \quad \mathcal{R}_{\text{Comp}, I}(\tilde{\pi}') = (\tilde{\pi}_U, \tilde{\pi}_V)$$

under this coupling. If  $\pi'$  and  $\tilde{\pi}'$  differ at one left vertex  $(v, r_0) \in U'$ , then

$$\mathcal{R}_{\text{Comp}, I}(\pi') = \mathcal{R}_{\text{Comp}, I}(\tilde{\pi}')$$

identically, because the recovery map never reads composed left labels at all.

If  $\pi'$  and  $\tilde{\pi}'$  differ at one right vertex  $(u, t) \in V'$ , then the common padded graph may still contain many edges incident to  $(u, t)$ ; what matters for the recovery map is only the subset of decoder calls that actually query slot  $t$ . For each randomness string  $\rho \in \{0, 1\}^r$ , let

$$Q_{u,t}(\rho) := \left\{ j \in [\deg_I(u)] : \begin{array}{l} \text{the decoder run on } (C_u, j) \text{ with randomness } \rho \\ \text{queries slot } t \text{ in at least one position} \end{array} \right\}.$$

If the sampled randomness equals  $\rho$ , then only the recovered right labels  $\pi_V(v)$  with  $v \in N_I(u)$  and neighbor index  $j(v) \in Q_{u,t}(\rho)$  can change, where  $j(v)$  denotes the position of  $v$  in the fixed ordering  $N_I(u)$ . By definition of the query-relevant proof-degree bound,

$$\sum_{\rho \in \{0,1\}^r} |Q_{u,t}(\rho)| \leq d.$$

Therefore under the shared-randomness coupling,

$$\mathbf{E} [d_H(\pi_V, \tilde{\pi}_V)] \leq 2^{-r} \sum_{\rho \in \{0,1\}^r} |Q_{u,t}(\rho)| \leq d \cdot 2^{-r}.$$

As above, each changed recovered right label can affect at most  $d_V$  recovered left labels, hence

$$\mathbf{E} [d_H(\pi_U, \tilde{\pi}_U) + d_H(\pi_V, \tilde{\pi}_V)] \leq \frac{d(1 + d_V)}{2^r}.$$

Let  $h := d_H(\pi', \tilde{\pi}')$  be the Hamming distance between the two assignments, and choose a sequence

$$\pi' = \pi^{(0)}, \pi^{(1)}, \dots, \pi^{(h)} = \tilde{\pi}'$$

in which each consecutive pair differs in exactly one coordinate. Applying the Hamming-distance-1 bound to each step and then using the triangle inequality for EMD gives

$$\text{EMD}(\mathcal{R}_{\text{Comp},I}(\pi'), \mathcal{R}_{\text{Comp},I}(\tilde{\pi}')) \leq \sum_{s=1}^h \text{EMD}(\mathcal{R}_{\text{Comp},I}(\pi^{(s-1)}), \mathcal{R}_{\text{Comp},I}(\pi^{(s)})) \leq \frac{d(1 + d_V)}{2^r} h.$$

Thus item (4) holds with

$$C_{R,\text{Comp}} = \frac{d(1 + d_V)}{2^r}.$$

**Recovery under source swaps.** Let  $I$  and  $\tilde{I}$  differ by one outer swap, and couple  $\mathcal{R}_{\text{Comp},I}(\pi')$  and  $\mathcal{R}_{\text{Comp},\tilde{I}}(\pi')$  by using the same sampled randomness  $r$ . Write

$$\mathcal{R}_{\text{Comp},I}(\pi') = (\pi_U, \pi_V), \quad \mathcal{R}_{\text{Comp},\tilde{I}}(\pi') = (\tilde{\pi}_U, \tilde{\pi}_V)$$

under this coupling. Only one circuit deciding admissible neighborhood tuples changes, say at an outer left vertex  $u$ . Therefore the recovered right labels can differ only at right vertices  $v$  adjacent to  $u$ , of which there are at most  $d_U$ . Each changed recovered right label can affect at most  $d_V$  recovered left labels, because the recovered left label at a vertex depends only on the labels of its right neighbors. In addition, the recovered left label at the swapped vertex  $u$  can change directly even if all recovered right labels stay fixed, since the defining feasible set for  $\pi_U(u)$  also depends on the local predicate  $P_u$  and the incident projections  $f_{(u,v)}$ . Hence the total number of changed recovered labels is at most

$$d_U + d_U d_V + 1 = d_U(1 + d_V) + 1,$$

and therefore

$$\text{EMD}(\mathcal{R}_{\text{Comp},I}(\pi'), \mathcal{R}_{\text{Comp},\tilde{I}}(\pi')) \leq d_U(1 + d_V) + 1,$$

proving item (5).

**Local description complexity.** Fix a composed left vertex  $(v, r)$  and let  $t(v) = \deg_I(v)$ . A tuple of symbols on the active rows of  $(v, r)$ , namely the rows indexed by the actual neighbors  $u_1, \dots, u_{t(v)}$  of  $v$  in Algorithm 2, extends to an admissible left label in the composed instance if and only if the decoder outputs on the queried coordinates of those rows all agree on the same non- $\perp$  symbol. Thus the admissible neighborhood set at  $(v, r)$  is recognized by a circuit that evaluates at most  $t(v) \leq d_V$  local maps  $g_{i,r}$ , each on coordinates among the  $m$  positions of one row, and checks equality of their outputs. Its decision complexity is therefore  $\text{poly}(d_V, m, s(N))$ .

The remaining structural bounds follow directly from Algorithm 2: the left side is  $V \times \{0, 1\}^r$ , the right side is  $U \times [m]$ , each composed left vertex  $(v, r)$  has exactly  $t(v)m = \deg_I(v)m \leq d_V m$  neighbors, and each composed right vertex  $(u, t)$  has exactly  $\deg_I(u)2^r \leq d_U 2^r$  neighbors. In particular, if  $I$  has no isolated vertices on either side, then neither does  $I \circledast \mathcal{D}$ . This proves item (6).  $\square$

## 11 Iteration by Stages

This section proves Theorem 1.6 by applying the Dinur–Harsha iteration stage-by-stage. Beginning with a stage-0 family, each subsequent stage consists of applying composition

$$\mathcal{I}_i \xrightarrow{\circledast \mathcal{D}_i} \mathcal{J}_{i+1},$$

followed by the fixed alphabet-then-degree reduction

$$\mathcal{J}_{i+1} \xrightarrow{\text{red}_{\varepsilon_{i+1}}^\circ} \mathcal{I}_{i+1},$$

where the stage parameter moves from  $\varepsilon_i$  to  $\varepsilon_{i+1} = 20\varepsilon_i^{1/4}$ . Throughout each iteration we track stage families on one common underlying graph together with maps from the hard base family and constants  $C_T^{(i)}, C_R^{(i)}, D_i$ , which control, respectively, how the target instance changes under one source swap, how the recovery changes when a labeling on a fixed target instance is perturbed, and how the recovery itself drifts under a source swap. Section 11.1 isolates the lemmas that show how these constants evolve from one step to the next, Section 11.2 describes the fixed alphabet/degree reduction on a common family, the next two subsections initialize stage 0 and formalize the one stage update, and the final subsection completes the proof of Theorem 1.6. Because each of these constants grows only polynomially in  $1/\varepsilon_i$  or  $1/\varepsilon_{i+1}$ , the recursion  $\varepsilon_{i+1} = 20\varepsilon_i^{1/4}$  keeps the cumulative loss polylogarithmic and the final sensitivity lower bound polynomial.

Symbol	Meaning
$N_{\text{base}}$	Size of the bounded-degree 2CSP family $\Phi_{N_{\text{base}}}$ from the Fleming–Yoshida lower bound
$n_0$	Initial target size used to initialize the iteration
$n_{\text{LC}}$	Final left-side size of the label cover instance in the proof of Theorem 1.6
$\varepsilon_i$	Stage parameter at stage $i$
$\chi_i$	Recovery threshold at stage $i$ , defined by $\chi_i = 9\sqrt{\varepsilon_i}$
$\xi_i$	Decoder soundness parameter at stage $i$ , defined by $\xi_i = \varepsilon_i^{1/4}$
$\mathcal{I}_i, \mathcal{J}_{i+1}$	Stage- $i$ family after reduction and intermediate family after composition
$C_T^{(i)}, C_R^{(i)}, D_i$	Bounds for transformation under source swaps, recovery under assignment perturbations, and recovery under source swaps

A *source swap* means one swap on the input side of the reduction currently under discussion: globally this is a swap in  $\Phi_{N_{\text{base}}}$ , while inside one stage it may mean a swap in the preceding label cover family. The stage families inherited from Section 8 still carry left predicates, so we keep that language throughout the iteration to maintain a common underlying graph under swaps; in the ordinary label cover special case, all predicates are identically 1.

## 11.1 How the Constants Evolve

Theorem 10.5 is the concrete one-layer composition statement. The next lemmas extract from it the update rules that are used later: how the transformation changes under source swaps, how the recovery changes under assignment perturbations, and how the recovery changes under source swaps propagate through one more fixed layer. In these lemmas, the input side of the current reduction step is the source family and the image side is the target family; we reserve “output” for assignments or distributions produced by algorithms or recovery maps. We begin with the following immediate specialization of Lemma 2.3: once a swap-closed family lives on a common underlying graph, a one-swap bound for the recovery drift extends to arbitrary swap distance by the same shortest-path triangle-inequality argument.

**Lemma 11.1** (Pathwise extension of a one-swap recovery bound). Let  $\mathcal{J}$  be a swap-closed family of instances whose outputs all live on one common underlying graph. For each  $I \in \mathcal{J}$ , let  $R_I$  be a recovery map from assignments on that common domain to some output space. Assume that for every one-swap pair  $I, \tilde{I} \in \mathcal{J}$  and every assignment  $\pi$  on the common domain,

$$\text{EMD}(R_I(\pi), R_{\tilde{I}}(\pi)) \leq D_{\text{single}}.$$

Then for every  $I, \tilde{I} \in \mathcal{J}$  and every assignment  $\pi$  on the common domain,

$$\text{EMD}(R_I(\pi), R_{\tilde{I}}(\pi)) \leq D_{\text{single}} \cdot \text{SwapDist}(I, \tilde{I}).$$

*Proof.* For the fixed assignment  $\pi$ , define an algorithm  $A_\pi$  on  $\mathcal{J}$  by

$$A_\pi(I) := R_I(\pi).$$

Because all instances in  $\mathcal{J}$  share one common underlying graph, the same assignment  $\pi$  is a valid input to every recovery map  $R_I$ . The hypothesis says exactly that for every one-swap pair  $I, \tilde{I} \in \mathcal{J}$ ,

$$\text{EMD}(A_\pi(I), A_\pi(\tilde{I})) \leq D_{\text{single}}.$$

Applying Lemma 2.3 to  $A_\pi$  therefore gives

$$\text{EMD}(R_I(\pi), R_{\tilde{I}}(\pi)) \leq D_{\text{single}} \cdot \text{SwapDist}(I, \tilde{I}),$$

as claimed. □

The next lemma gives the update rule that will be used twice in each stage update. Start with a map  $\Phi \mapsto \text{Inst}(\Phi)$  from the base family to some intermediate instances, together with recovery maps  $\text{Rec}_\Phi$  back to the base instance. The constants  $(C_T, C_R, D)$  record, respectively, how far the target instance can move under one source swap, how sensitively the recovery depends on perturbing the assignment while the source instance is fixed, and how much the recovery itself can change when the source instance changes but the

assignment is held fixed. Now suppose that, after producing the intermediate instances, we apply one further transformation step

$$I \mapsto T(I)$$

to each intermediate instance  $I$ , together with a recovery map  $R_I$  from assignments to  $T(I)$  back to assignments to  $I$ . Assume that this additional step comes with its own analogous constants  $(C_{T,\text{layer}}, C_{R,\text{layer}}, D_{\text{layer}})$ . The lemma says that after composing the two steps, the new instance map  $\Phi \mapsto \text{Inst}^+(\Phi)$  and the composed recoveries  $\text{Rec}_\Phi^+$  again satisfy the same three tracking properties, with the expected update rule: the transformation and assignment-sensitivity constants multiply, while the part measuring how the recovery map changes under source swaps contributes an additive term after passing through the old Lipschitz constant  $C_R$ .

**Lemma 11.2** (Composing reduction data with one additional layer). Let  $\Phi_N$  be a source family, and let  $\mathcal{K} := \{\text{Inst}(\Phi) : \Phi \in \Phi_N\}$  be the corresponding target family. Suppose we are given recoveries  $\text{Rec}_\Phi$ , one for each source instance  $\Phi$ , and constants  $C_T, C_R, D$  such that the following hold. Items (1) and (3) below are one-source-swap bounds, whereas item (2) is a fixed-instance Lipschitz bound for arbitrary assignment perturbations.

1. For every one-swap pair  $\Phi, \tilde{\Phi} \in \Phi_N$ ,

$$\text{SwapDist}(\text{Inst}(\Phi), \text{Inst}(\tilde{\Phi})) \leq C_T;$$

2. For every fixed  $\Phi$  and every two assignments  $x, \tilde{x}$  on  $\text{Inst}(\Phi)$ ,

$$\text{EMD}(\text{Rec}_\Phi(x), \text{Rec}_\Phi(\tilde{x})) \leq C_R \cdot d_H(x, \tilde{x});$$

3. For every one-swap pair  $\Phi, \tilde{\Phi} \in \Phi_N$ , the two target instances  $\text{Inst}(\Phi)$  and  $\text{Inst}(\tilde{\Phi})$  live on a common domain, and for every assignment  $x$  on that common domain,

$$\text{EMD}(\text{Rec}_\Phi(x), \text{Rec}_{\tilde{\Phi}}(x)) \leq D.$$

Larger source distances will later be handled by Lemma 11.1, so here we record only the one-swap bounds on the source side. Suppose moreover that we are given another transformation step on  $\mathcal{K}$ : for each intermediate instance  $I \in \mathcal{K}$ , it produces a new instance  $T(I)$  together with a recovery map  $R_I$  from assignments on  $T(I)$  back to assignments on  $I$ . Assume that this step has associated constants  $C_{T,\text{layer}}, C_{R,\text{layer}}, D_{\text{layer}}$  such that:

- (a) for every  $I, \tilde{I} \in \mathcal{K}$ ,

$$\text{SwapDist}(T(I), T(\tilde{I})) \leq C_{T,\text{layer}} \text{SwapDist}(I, \tilde{I});$$

- (b) for each fixed  $I \in \mathcal{K}$  and every two assignments  $y, \tilde{y}$  on  $T(I)$ ,

$$\text{EMD}(R_I(y), R_I(\tilde{y})) \leq C_{R,\text{layer}} d_H(y, \tilde{y});$$

- (c) for every one-swap pair  $\Phi, \tilde{\Phi} \in \Phi_N$ , the two instances  $T(\text{Inst}(\Phi))$  and  $T(\text{Inst}(\tilde{\Phi}))$  live on a common domain, and for every assignment  $y$  on that common domain,

$$\text{EMD}(R_{\text{Inst}(\Phi)}(y), R_{\text{Inst}(\tilde{\Phi})}(y)) \leq D_{\text{layer}}.$$

Define

$$\text{Inst}^+(\Phi) := T(\text{Inst}(\Phi)), \quad \text{Rec}_\Phi^+ := \text{Rec}_\Phi \circ R_{\text{Inst}(\Phi)}.$$

Then the composed tracking constants satisfy

$$C_T^+ = C_T C_{T,\text{layer}}, \quad C_R^+ = C_R C_{R,\text{layer}}, \quad D^+ \leq D + C_R D_{\text{layer}},$$

where for every one-swap pair  $\Phi, \tilde{\Phi} \in \Phi_N$ ,

$$\text{SwapDist}(\text{Inst}^+(\Phi), \text{Inst}^+(\tilde{\Phi})) \leq C_T^+,$$

for every fixed  $\Phi$  and every two assignments  $y, \tilde{y}$  on  $\text{Inst}^+(\Phi)$ ,

$$\text{EMD}(\text{Rec}_\Phi^+(y), \text{Rec}_\Phi^+(\tilde{y})) \leq C_R^+ d_H(y, \tilde{y}),$$

and for every one-swap pair  $\Phi, \tilde{\Phi}$  and every assignment  $y$  on the common domain of  $\text{Inst}^+(\Phi)$  and  $\text{Inst}^+(\tilde{\Phi})$ ,

$$\text{EMD}(\text{Rec}_\Phi^+(y), \text{Rec}_{\tilde{\Phi}}^+(y)) \leq D^+.$$

*Proof.* The bound for how the transformation changes under source swaps is immediate:

$$\begin{aligned} \text{SwapDist}(\text{Inst}^+(\Phi), \text{Inst}^+(\tilde{\Phi})) &= \text{SwapDist}(T(\text{Inst}(\Phi)), T(\text{Inst}(\tilde{\Phi}))) \\ &\leq C_{T,\text{layer}} \text{SwapDist}(\text{Inst}(\Phi), \text{Inst}(\tilde{\Phi})) \\ &\leq C_T C_{T,\text{layer}}. \end{aligned}$$

Likewise, for fixed  $\Phi$  and assignments  $y, \tilde{y}$  on  $\text{Inst}^+(\Phi)$ , couple  $R_{\text{Inst}(\Phi)}(y)$  and  $R_{\text{Inst}(\Phi)}(\tilde{y})$  optimally and apply the bound for how  $\text{Rec}_\Phi$  changes under assignment perturbations pointwise under that coupling. This gives

$$\begin{aligned} \text{EMD}(\text{Rec}_\Phi^+(y), \text{Rec}_\Phi^+(\tilde{y})) &= \text{EMD}\left(\text{Rec}_\Phi(R_{\text{Inst}(\Phi)}(y)), \text{Rec}_\Phi(R_{\text{Inst}(\Phi)}(\tilde{y}))\right) \\ &\leq C_R \text{EMD}\left(R_{\text{Inst}(\Phi)}(y), R_{\text{Inst}(\Phi)}(\tilde{y})\right) \\ &\leq C_R C_{R,\text{layer}} d_H(y, \tilde{y}). \end{aligned}$$

For the bound for how the recovery changes under source swaps, fix a one-swap pair  $\Phi, \tilde{\Phi} \in \Phi_N$  and an assignment  $y$  on the common domain of  $\text{Inst}^+(\Phi)$  and  $\text{Inst}^+(\tilde{\Phi})$ . Let  $(\mathbf{X}, \tilde{\mathbf{X}})$  be an optimal coupling of the two distributions

$$R_{\text{Inst}(\Phi)}(y) \quad \text{and} \quad R_{\text{Inst}(\tilde{\Phi})}(y).$$

Because  $\text{Inst}(\Phi)$  and  $\text{Inst}(\tilde{\Phi})$  live on a common domain by assumption, we may view  $\mathbf{X}$  and  $\tilde{\mathbf{X}}$  as assignments on that same domain. Assumption (c) gives

$$\mathbf{E}[d_H(\mathbf{X}, \tilde{\mathbf{X}})] = \text{EMD}(R_{\text{Inst}(\Phi)}(y), R_{\text{Inst}(\tilde{\Phi})}(y)) \leq D_{\text{layer}}.$$

Conditioning on  $\mathbf{X} = x$  and  $\tilde{\mathbf{X}} = \tilde{x}$ , choose an optimal coupling of  $\text{Rec}_\Phi(x)$  and  $\text{Rec}_\Phi(\tilde{x})$ , and an optimal coupling of  $\text{Rec}_\Phi(\tilde{x})$  and  $\text{Rec}_{\tilde{\Phi}}(\tilde{x})$ . Both of these couplings have the same middle marginal, namely the distribution  $\text{Rec}_\Phi(\tilde{x})$ . Hence we may glue them to obtain a joint distribution on  $(\mathbf{Z}, \mathbf{Z}', \tilde{\mathbf{Z}})$  whose

$(\mathbf{Z}, \mathbf{Z}')$ -marginal is an optimal coupling of  $\text{Rec}_\Phi(x)$  and  $\text{Rec}_\Phi(\tilde{x})$ , and whose  $(\mathbf{Z}', \tilde{\mathbf{Z}})$ -marginal is an optimal coupling of  $\text{Rec}_\Phi(\tilde{x})$  and  $\text{Rec}_{\tilde{\Phi}}(\tilde{x})$ . Averaging this joint coupling over  $(\mathbf{X}, \tilde{\mathbf{X}})$  yields

$$\begin{aligned} \text{EMD}(\text{Rec}_\Phi^+(y), \text{Rec}_{\tilde{\Phi}}^+(y)) &\leq \mathbf{E} \left[ \text{EMD}(\text{Rec}_\Phi(\mathbf{X}), \text{Rec}_\Phi(\tilde{\mathbf{X}})) + \text{EMD}(\text{Rec}_\Phi(\tilde{\mathbf{X}}), \text{Rec}_{\tilde{\Phi}}(\tilde{\mathbf{X}})) \right] \\ &\leq C_R \mathbf{E} [d_{\text{H}}(\mathbf{X}, \tilde{\mathbf{X}})] + D \\ &\leq D + C_R D_{\text{layer}}. \end{aligned}$$

□

**Lemma 11.3** (Sensitivity pullback when the recovery map depends on the source instance). Let  $\mathcal{X}$  be a source family and  $\mathcal{Y}$  a target family, each equipped with a swap distance and Hamming distance on their outputs. For each  $I \in \mathcal{X}$ , let  $\text{Inst}(I) \in \mathcal{Y}$  and let  $\text{Rec}_I$  be a recovery map from assignments on the target instance  $\text{Inst}(I)$  to assignments on  $I$ . Assume there are constants  $C_T, C_R, D$  such that for every one-swap pair  $I, \tilde{I} \in \mathcal{X}$ :

1.  $\text{SwapDist}(\text{Inst}(I), \text{Inst}(\tilde{I})) \leq C_T$ ;
2. For every fixed  $I$  and every two assignments  $y, \tilde{y}$  on the target instance  $\text{Inst}(I)$ ,  $\text{EMD}(\text{Rec}_I(y), \text{Rec}_I(\tilde{y})) \leq C_R d_{\text{H}}(y, \tilde{y})$ ;
3. The two target instances  $\text{Inst}(I)$  and  $\text{Inst}(\tilde{I})$  live on a common target domain, and for every assignment  $y$  on that domain,  $\text{EMD}(\text{Rec}_I(y), \text{Rec}_{\tilde{I}}(y)) \leq D$ .

Let  $\mathcal{Y}' \supseteq \{\text{Inst}(I) : I \in \mathcal{X}\}$  be swap-closed, and let  $A$  be an algorithm on  $\mathcal{Y}'$ . Define

$$B(I) := \text{Rec}_I(A(\text{Inst}(I))).$$

Then for every one-swap pair  $I, \tilde{I} \in \mathcal{X}$ ,

$$\text{EMD}(B(I), B(\tilde{I})) \leq C_T C_R \text{SwapSens}(A, \mathcal{Y}') + D.$$

Consequently,

$$\text{SwapSens}(B, \mathcal{X}) \leq C_T C_R \text{SwapSens}(A, \mathcal{Y}') + D.$$

*Proof.* Set  $S := \text{SwapSens}(A, \mathcal{Y}')$ . Fix a one-swap pair  $I, \tilde{I} \in \mathcal{X}$ . Then

$$\begin{aligned} \text{EMD}(B(I), B(\tilde{I})) &\leq \text{EMD}(\text{Rec}_I(A(\text{Inst}(I))), \text{Rec}_I(A(\text{Inst}(\tilde{I})))) \\ &\quad + \text{EMD}(\text{Rec}_I(A(\text{Inst}(\tilde{I}))), \text{Rec}_{\tilde{I}}(A(\text{Inst}(\tilde{I}))))). \end{aligned}$$

For the first term, because  $\mathcal{Y}'$  is swap-closed and contains both target instances, a shortest swap path

$$J^{(0)} = \text{Inst}(I), J^{(1)}, \dots, J^{(k)} = \text{Inst}(\tilde{I}), \quad k := \text{SwapDist}(\text{Inst}(I), \text{Inst}(\tilde{I})),$$

lies entirely in  $\mathcal{Y}'$ . By the definition of

$$S = \text{SwapSens}(A, \mathcal{Y}'),$$

each step satisfies

$$\text{EMD}(A(J^{(t)}), A(J^{(t+1)})) \leq S \quad \text{for every } t \in \{0, \dots, k-1\}.$$

Hence

$$\text{EMD}(A(\text{Inst}(I)), A(\text{Inst}(\tilde{I}))) \leq kS \leq C_T S.$$

Couple  $A(\text{Inst}(I))$  and  $A(\text{Inst}(\tilde{I}))$  optimally and apply the bound for how the recovery changes under assignment perturbations pointwise under that coupling. This gives

$$\text{EMD}\left(\text{Rec}_I(A(\text{Inst}(I))), \text{Rec}_I(A(\text{Inst}(\tilde{I})))\right) \leq C_R \text{EMD}(A(\text{Inst}(I)), A(\text{Inst}(\tilde{I}))) \leq C_T C_R S.$$

For the second term, sample  $\mathbf{y} \sim A(\text{Inst}(\tilde{I}))$  and couple  $\text{Rec}_I(\mathbf{y})$  with  $\text{Rec}_{\tilde{I}}(\mathbf{y})$  optimally. Assumption (3) then gives

$$\text{EMD}\left(\text{Rec}_I(A(\text{Inst}(\tilde{I}))), \text{Rec}_{\tilde{I}}(A(\text{Inst}(\tilde{I})))\right) \leq D.$$

Combining the two bounds proves the claim.  $\square$

**Corollary 11.4** (Lower-bound consequence). In the setting of Lemma 11.3, if  $\text{SwapSens}(B, \mathcal{X}) \geq L$ , then  $\text{SwapSens}(A, \mathcal{Y}') \geq \frac{L-D}{C_T C_R}$ . In particular, if  $D = o(L)$ , then  $\text{SwapSens}(A, \mathcal{Y}') = \Omega\left(\frac{L}{C_T C_R}\right)$ .

*Proof.* Rearranging the bound from Lemma 11.3 gives the claim.  $\square$

## 11.2 The Fixed Alphabet/Degree Reduction on a Common Family

The following definitions and lemmas describe how  $\text{red}_\varepsilon^\circ$ —alphabet-then-degree reduction from Lemma 9.4—is used in a single stage update. Later stage statements will also require every left and right vertex to remain non-isolated, so in addition to the usual structural bounds, we observe at the end of this subsection that if the instance began with no isolated vertices, then after alphabet-then-degree reduction, the resultant instance does not have any isolated vertices.

we record that this fixed cleanup step preserves that property. For each common stage family, we fix the reduction once-and-for-all so that the map  $I \mapsto \text{red}_\varepsilon^\circ(I)$  is fixed and independent of which instance in the family is chosen. The recovery maps still depend on the source instance, however, because the alphabet-reduction recovery map uses the source instance through the distributions and projections from Section 9.

For the alphabet-then-degree reduction of Section 9, it is convenient to separate the output of a single instance from the swap closure used in the sensitivity statement.

**Definition 11.5** (Output of the alphabet-then-degree reduction). Fix  $\varepsilon > 0$ . For a left-predicate label cover instance  $I$ , let  $\text{red}_\varepsilon^\circ(I)$  denote the particular instance obtained from  $I$  by the constructions used in the proof of Lemma 9.4, before taking the swap closure. In other words, we first apply the alphabet-reduction map from Section 9, and then apply the degree-reduction map exactly when the input instance has minimum right degree  $\delta_V > c/\varepsilon$ ; if  $\delta_V \leq c/\varepsilon$ , we stop after alphabet reduction. For a family  $\mathcal{I}$ , write

$$\text{red}_\varepsilon^\circ(\mathcal{I}) := \{\text{red}_\varepsilon^\circ(I) : I \in \mathcal{I}\}.$$

Every swap performed before the alphabet-then-degree reduction can be pushed through the fixed gadgets to a corresponding swap after the reduction, coordinate by coordinate. Consequently,

$$\text{red}_\varepsilon^\circ(\text{SwapClo}(\mathcal{I})) \subseteq \text{SwapClo}(\text{red}_\varepsilon^\circ(\mathcal{I})).$$

The reverse inclusion need not hold, because the right-hand side also allows additional swaps among the new codeword coordinates and among the cloud copies  $(v, i)$  created when degree reduction replaces one original right vertex by a cloud of right vertices; these extra swaps need not descend from a single pre-reduction swap.

Recall that in the proof of Lemma 9.4, the alphabet-then-degree reduction, we branch on whether the minimum right degree  $\delta_V$  of the instance is small or large. If it is large, then we apply degree reduction, and otherwise we do not. The following lemma states that the same choice is made for every instance in a family sharing the same structure.

**Lemma 11.6** (The reduction takes the same branch on a common family). Let  $\varepsilon > 0$ , and let  $\mathcal{J}$  be a family of left-predicate label cover instances with a common underlying graph and common alphabets. Then the proof of Lemma 9.4 takes the same branch for every  $I \in \mathcal{J}$ .

*Proof.* In the proof of Lemma 9.4, we branch on whether the minimum right degree  $\delta_V$  of the input instance satisfies

$$\delta_V \leq c/\varepsilon \quad \text{or} \quad \delta_V > c/\varepsilon,$$

where  $c$  is the absolute constant fixed there. Because all instances in  $\mathcal{J}$  share the same underlying graph, they have the same right-degree profile and hence the same value of  $\delta_V$ . Therefore the same branch is taken for every  $I \in \mathcal{J}$ .  $\square$

**Definition 11.7** (Fixing the auxiliary choices for the reduction). In Section 11 we apply  $\text{red}_\varepsilon^\circ$  only to stage families  $\mathcal{J}$  for which Lemma 11.6 holds. For such a family we additionally fix:

1. A code  $\mathcal{C} : \Sigma_V \rightarrow \widehat{\Sigma}^k$  promised by Proposition 9.2 for the common source right alphabet  $\Sigma_V$  and the parameter  $\eta = \varepsilon/4$ ;
2. If the degree-reduction branch is taken, then for each right-degree value  $D$  in the common source graph and for  $d = \lceil c/\varepsilon \rceil$ , the expander  $H_D$  returned by the explicit construction of Proposition 4.2, and we use  $H_{D_v}$  at every right vertex  $v$ ;
3. The neighbor orderings, default symbols, and tie-breaking conventions required by the reduction gadgets.

These choices affect only the reduction itself. They make the map  $I \mapsto \text{red}_\varepsilon^\circ(I)$  deterministic on the stage family under discussion, but the alphabet/degree-reduction recovery maps still vary with the source instance. We write  $R_{\varepsilon, I}^{\text{red}}$  for the recovery map associated with a source instance  $I$ . This recovery still depends on  $I$  itself: already the alphabet-reduction recovery  $\mathcal{R}_{\text{AR}}$  uses the source projections of  $I$  through the distributions  $p_v$  from Section 9.

By Lemma 11.6, any two instances  $I, \tilde{I}$  in such a family with the same underlying graph and the same alphabets take the same branch in the reduction. In the case when we stop after alphabet-reduction, the target graph is the common graph  $V \times [k]$  determined by the fixed code length  $k$ , and the target alphabet is the common code alphabet. In the degree-reduction case, the fixed expanders  $H_D$  depend only on the common right-degree profile and on  $d = \lceil c/\varepsilon \rceil$ . Thus the outputs  $\text{red}_\varepsilon^\circ(I)$  and  $\text{red}_\varepsilon^\circ(\tilde{I})$  always have a common target graph and common alphabets. In particular, on any common stage family  $\mathcal{J}$  with the fixed auxiliary choices from Definition 11.7, the whole image family  $\text{red}_\varepsilon^\circ(\mathcal{J})$  has a common target graph and common alphabets.

**Lemma 11.8** (Local description complexity survives the alphabet-then-degree reduction). Let  $\varepsilon > 0$  and let  $\mathcal{I}$  be a family of left-predicate label cover instances with common right alphabet  $\Sigma_V$ , local description complexity at most  $S$ , left degree at most  $\Delta_U$ , and right degree at most  $\Delta_V$ . Then  $\text{red}_\varepsilon^\circ(\mathcal{I})$  has local description complexity at most

$$\text{poly}(S, \varepsilon^{-1}, \Delta_U, \Delta_V, |\Sigma_V|).$$

In particular, if  $S, \Delta_U, \Delta_V, |\Sigma_V| \leq \varepsilon^{-O(1)}$ , then the same is true for the local description complexity of  $\text{red}_\varepsilon^\circ(\mathcal{I})$ .

We track this parameter because the next composition step feeds the circuits deciding admissible neighborhood tuples into the dPCP decoder, so their decision complexity must remain polynomially bounded throughout the iteration. Here this is an interface condition for the next stage, not an end in itself: after cleanup we still need these predicates to admit circuits of size polynomial in the current stage parameter, so that they can be padded to a common bound  $M_i$  and fed back into Lemma 10.3 and theorem 10.5 at the next stage.

*Proof.* We inspect the two gadgets used in the proof of Lemma 9.4.

**Alphabet reduction.** Let  $I^{(1)} = \mathcal{T}_{\text{AR}}(I)$  be the alphabet-reduced instance built from the fixed code  $C : \Sigma_V \rightarrow \widehat{\Sigma}^k$  from Section 9, where  $k = O(\varepsilon^{-6} \log |\Sigma_V|)$ . Fix a left vertex  $u$  of  $I$  with ordered neighborhood  $N_I(u) = (v_1, \dots, v_t)$ , and let  $C_u$  be a circuit deciding admissible tuples of neighboring right labels for  $u$ , of size at most  $S$ . Fix an injective binary encoding

$$\text{enc} : \Sigma_V \hookrightarrow \{0, 1\}^\ell, \quad \ell := \lceil \log_2 |\Sigma_V| \rceil.$$

Under this encoding, each finite-alphabet gate appearing in  $C_u$  can be replaced by a Boolean subcircuit implementing the same local truth table. Thus  $C_u$  has an equivalent Boolean realization

$$\widehat{C}_u : \{0, 1\}^{t\ell} \rightarrow \{0, 1\}$$

of size  $\text{poly}(S, t, |\Sigma_V|) = \text{poly}(S, \Delta_U, |\Sigma_V|)$ . In  $I^{(1)}$ , the neighbors of  $u$  are the blocks

$$(v_1, 1), \dots, (v_1, k), \dots, (v_t, 1), \dots, (v_t, k).$$

A tuple of block labels extends across  $u$  in  $I^{(1)}$  if and only if two conditions hold. First, for every  $r \in [t]$ , the  $r$ th block is a codeword  $C(b_r)$  for some  $b_r \in \Sigma_V$ . Second, the decoded tuple  $(b_1, \dots, b_t)$  lies in  $\mathcal{N}_I(u)$ . For each block index  $r \in [t]$ , each block label  $x_r \in \widehat{\Sigma}^k$ , and each symbol  $a \in \Sigma_V$ , compare  $x_r$  with the codeword  $C(a)$ . Let  $\chi_{r,a}(x_r)$  be the indicator that  $x_r = C(a)$ , implemented by  $k$  coordinate-equality tests followed by an AND gate. Because  $C$  is injective, for each fixed  $r$  at most one  $\chi_{r,a}(x_r)$  can equal 1. Hence image-membership for the  $r$ th block is checked by the OR

$$\bigvee_{a \in \Sigma_V} \chi_{r,a}(x_r).$$

On the valid branch, the encoded bits of the unique decoded symbol are recovered by

$$(\widehat{b}_r)_j := \bigvee_{a \in \Sigma_V : \text{enc}(a)_j = 1} \chi_{r,a}(x_r), \quad j \in [\ell].$$

This uses size  $\text{poly}(k, |\Sigma_V|)$  for each block. Running this decoder independently on the  $t \leq \Delta_U$  blocks produces a bit string

$$(\widehat{b}_1, \dots, \widehat{b}_t) \in \{0, 1\}^{t\ell},$$

and the new circuit accepts exactly when every block passes its image-membership test and the bit string  $(\widehat{b}_1, \dots, \widehat{b}_t)$  is accepted by  $\widehat{C}_u$ . This gives a Boolean circuit deciding admissible tuples of neighboring right labels for  $u$  in  $I^{(1)}$ , of size  $\text{poly}(S, \varepsilon^{-1}, \Delta_U, |\Sigma_V|)$ .

**Degree reduction.** If the proof of Lemma 9.4 stops after alphabet reduction, there is nothing more to prove. Otherwise, let  $I^{(2)} = \text{TDR}(I^{(1)}; d)$  be the degree-reduced instance, with  $d = \Theta(\varepsilon^{-1})$ . The left vertices are unchanged. Each original right vertex  $v$  is replaced by a cloud of vertices  $(v, i)$ , and every original edge incident to  $u$  is replaced by several edges whose projections are exact copies of the corresponding original projection. Therefore, a tuple of labels on the new neighborhood of  $u$  extends across  $u$  in  $I^{(2)}$  if and only if the labels on all copies belonging to the same original neighbor  $v$  are equal, and the resulting collapsed tuple belongs to the admissible neighborhood set of  $u$  in  $I^{(1)}$ . A circuit can check this by performing equality tests inside each cloud-copy group and then evaluating the circuit for  $u$  in  $I^{(1)}$ . Its size is  $\text{poly}(S, \varepsilon^{-1}, \Delta_U, \Delta_V, |\Sigma_V|)$ .

This proves the stated bound for  $\text{red}_\varepsilon^\circ(I)$ , uniformly over all  $I \in \mathcal{I}$ .  $\square$

In Section 11, every application of Lemma 11.8 has input right alphabet size  $|\Sigma_V| \leq \varepsilon^{-O(1)}$  at the relevant stage, so the polynomial dependence on  $|\Sigma_V|$  still yields local description complexity  $\varepsilon^{-O(1)}$  throughout the iteration.

Finally, we observe that applying alphabet-then-degree reduction preserves the property of not having isolated vertices. We record this because the stage invariants below require every left and right vertex to have degree at least 1. Together with the analogous preservation fact for composition from Theorem 10.5, this lets us carry non-isolatedness through each stage of the iteration, until the final exact-size padding step where isolated left vertices are added deliberately. To see that this is the case, observe that in both branches of  $\text{red}_\varepsilon^\circ$ , the left vertices are unchanged, and each new right vertex is a copy of an old right vertex attached to copies of its incident edges. Consequently, if the input instance has no isolated vertices on either side, then neither does its reduced image. We record this because the stage invariants below require every left and right vertex to have degree at least 1. Together with the analogous preservation fact for composition from Theorem 10.5, this lets us carry non-isolatedness through each stage of the iteration, until the final exact-size padding step where isolated left vertices are added deliberately.

### 11.3 Initialization of the Left-Predicate Label Cover under Clause Swaps

We initialize the stage-0 reduction data from the left-predicate label cover under clause swaps supplied by Section 8.1. The point of this subsection is not merely to produce one family  $\mathcal{I}_0$ , but to put the proof into the exact form needed to start the iterative update from Proposition 11.12: we need a satisfiable stage-0 family on a common target graph and common alphabets, with local description complexity, degrees, alphabets, and the three constants controlling swaps and perturbations already polynomial in  $1/\varepsilon_0$ , together with a direct recovery that turns value at least  $9\sqrt{\varepsilon_0}$  on the target side into source value at least  $1 - \eta_{\text{FY}}$ .

The construction has three steps. Starting from the robust left-predicate family constructed in Section 8, we first perform one preliminary cleanup at scale  $\tau_0 = \varepsilon_0^8$  so that the family is ready for the Dinur–Harsha composition. We then compose once to obtain the actual stage-0 family  $\mathcal{I}_0$ . Finally, we build the stage-0 recovery by running a randomized recovery procedure repeatedly and feeding the resulting source assignments into a threshold selector. Unlike the later stages, the initial family  $\mathcal{G}_0$ —obtained by running the robust PCP transformation from Section 8 on the initial family of CNF formulas—still has very large right degree, so feeding it directly into the first composition would result in a large left-degree through the  $d_V \cdot m$  bound from Theorem 10.5. We therefore perform a preliminary cleanup before the first composition, after which the usual stage transition

$$\mathcal{I}_i \xrightarrow{\otimes \mathcal{D}_i} \mathcal{J}_{i+1} \xrightarrow{\text{red}_{\varepsilon_{i+1}}^\circ} \mathcal{I}_{i+1}.$$

resumes unchanged. This subsection carries out that preliminary cleanup step, producing the stage-0 family

$\mathcal{I}_0$  together with the structural bounds and the three swap/perturbation constants, all already polynomial in  $1/\varepsilon_0$ .

The remaining ingredient in the initialization is a single recovery map back to the source family; we want this map to have strong perturbation bounds. One run of the sampler from Theorem 8.10, which selects among tuples with large outer-edge agreement probability, already gives an inverse-polynomial-probability success event, with both the assignment-perturbation bound and the source-swap bound for the recovery map bounded by  $\varepsilon_0^{-O(1)}$ . We therefore repeat this randomized extraction polynomially many times and then apply a randomized threshold selection map on source assignments. Because this map only compares exact source values against a random threshold in a constant-width interval, its assignment-perturbation constant is linear in the total candidate Hamming perturbation and its source-swap constant is linear in the number of candidates. The next lemma records this map in the form needed for Proposition 11.10.

**Lemma 11.9** (Random-threshold selection map for source assignments). There exists an absolute constant  $C_{\text{sel}} \geq 1$  with the following property. Fix constants  $0 < \eta_1 < \eta_2 < 1$  and  $I := [1 - \eta_2, 1 - \eta_1]$ . For a source instance  $\Phi \in \Phi_{N_{\text{base}}}$  and a list of assignments  $\mathbf{a} = (a^{(1)}, \dots, a^{(T)})$  on the variable set of  $\Phi$ , let  $\text{Sel}_{\Phi}^I(\mathbf{a})$  be the randomized map that samples  $\theta \sim I$  uniformly and returns the first index  $j \in [T]$  with  $\text{val}_{\Phi}(a^{(j)}) \geq \theta$ , or a fixed default source assignment  $a_{\text{def}}$  if no such index exists. Then:

1. If some candidate in the list satisfies  $\Phi$ , then every output of  $\text{Sel}_{\Phi}^I(\mathbf{a})$  has source value at least  $1 - \eta_2$ .
2. For every fixed  $\Phi \in \Phi_{N_{\text{base}}}$  and every two candidate lists

$$\mathbf{a} = (a^{(1)}, \dots, a^{(T)}), \quad \tilde{\mathbf{a}} = (\tilde{a}^{(1)}, \dots, \tilde{a}^{(T)}),$$

we have

$$\text{EMD}(\text{Sel}_{\Phi}^I(\mathbf{a}), \text{Sel}_{\Phi}^I(\tilde{\mathbf{a}})) \leq C_{\text{sel}} \sum_{j=1}^T d_{\text{H}}(a^{(j)}, \tilde{a}^{(j)}).$$

3. For every one-swap pair  $\Phi, \tilde{\Phi} \in \Phi_{N_{\text{base}}}$  and every fixed candidate list  $\mathbf{a} = (a^{(1)}, \dots, a^{(T)})$  on the common source domain,

$$\text{EMD}(\text{Sel}_{\Phi}^I(\mathbf{a}), \text{Sel}_{\tilde{\Phi}}^I(\mathbf{a})) \leq C_{\text{sel}} T.$$

*Proof.* Because the Fleming–Yoshida source family has constant degree and constant alphabet, there exists an absolute constant  $C_{\text{src}} \geq 1$  such that for every source instance  $\Phi \in \Phi_{N_{\text{base}}}$  and every two assignments  $a, \tilde{a}$  on its variable set,  $|\text{val}_{\Phi}(a) - \text{val}_{\Phi}(\tilde{a})| \leq C_{\text{src}} \frac{d_{\text{H}}(a, \tilde{a})}{|\text{Var}(\Phi)|}$ . Likewise, if  $\Phi$  and  $\tilde{\Phi}$  differ by one source swap, then every fixed assignment changes its source value by at most  $C_{\text{src}}/|\text{Var}(\Phi)|$ , because only  $O(1)$  constraints change and  $|E(\Phi)| = \Theta(|\text{Var}(\Phi)|)$ .

Item (1) is immediate: if one candidate satisfies  $\Phi$ , then for every threshold  $\theta \in I$  there exists at least one index whose value is at least  $\theta$ , and the returned assignment therefore has value at least  $\theta \geq 1 - \eta_2$ .

For item (2), couple the two runs by using the same threshold  $\theta \sim I$ . Let  $L := \sum_{j=1}^T d_{\text{H}}(a^{(j)}, \tilde{a}^{(j)})$ . If, for every  $j \in [T]$ , the threshold  $\theta$  does not lie between  $\text{val}_{\Phi}(a^{(j)})$  and  $\text{val}_{\Phi}(\tilde{a}^{(j)})$ , then the indicator of the event “candidate  $j$  is above threshold” is the same for both lists, hence the selected index is identical. On this event the output Hamming distance is at most  $L$ . By the value-Lipschitz bound above, the probability of a threshold accident for a fixed  $j$  is at most

$$\frac{C_{\text{src}}}{(\eta_2 - \eta_1)|\text{Var}(\Phi)|} d_{\text{H}}(a^{(j)}, \tilde{a}^{(j)}),$$

and summing over  $j$  gives total accident probability at most

$$\frac{C_{\text{src}}}{(\eta_2 - \eta_1)|\text{Var}(\Phi)|} L.$$

Since every source assignment lives on  $|\text{Var}(\Phi)|$  coordinates, the expected Hamming distance contributed by the accident event is at most

$$\frac{C_{\text{src}}}{\eta_2 - \eta_1} L.$$

Therefore

$$\text{EMD}(\text{Sel}_{\Phi}^I(\mathbf{a}), \text{Sel}_{\tilde{\Phi}}^I(\tilde{\mathbf{a}})) \leq \left(1 + \frac{C_{\text{src}}}{\eta_2 - \eta_1}\right) L,$$

proving item (2) after enlarging  $C_{\text{sel}}$ .

For item (3), again couple the two runs by using the same threshold  $\theta$ . Now the candidate list itself is identical, so the outputs can differ only if some candidate value crosses the threshold between  $\Phi$  and  $\tilde{\Phi}$ . The one-swap value perturbation bound gives accident probability at most

$$\frac{C_{\text{src}} T}{(\eta_2 - \eta_1)|\text{Var}(\Phi)|}.$$

Multiplying by the maximum possible Hamming distance  $|\text{Var}(\Phi)|$  proves

$$\text{EMD}(\text{Sel}_{\Phi}^I(\mathbf{a}), \text{Sel}_{\tilde{\Phi}}^I(\mathbf{a})) \leq \frac{C_{\text{src}}}{\eta_2 - \eta_1} T,$$

and item (3) follows after one final enlargement of  $C_{\text{sel}}$ .  $\square$

**Proposition 11.10** (Initialization from the left-predicate label cover under clause swaps). There exist absolute constants  $c_{\text{init}}, c_{\text{str}}, c_{\text{size}}, c_{\text{sens}}, \delta > 0$  such that the following holds for every sufficiently large target size  $n_0$ . Let  $\varepsilon_0 := (\log n_0)^{-c_{\text{init}}}$ . Then there exists a family  $\mathcal{I}_0$  of left-predicate label cover instances with local description complexity at most  $\varepsilon_0^{-c_{\text{str}}}$  satisfying the following properties.

1. The local description complexity, the left degree, the right degree, and the right alphabet size are all at most  $\varepsilon_0^{-c_{\text{str}}}$ , the left alphabet size is at most  $\exp(\varepsilon_0^{-c_{\text{str}}})$ , and every left and right vertex has degree at least 1.
2. The numbers of left and right vertices satisfy

$$|U|, |V| \leq n_0 \varepsilon_0^{-c_{\text{size}}}.$$

3. The swap sensitivity of the swap closure satisfies

$$\text{SwapSens}_{9\sqrt{\varepsilon_0}}(\text{SwapClo}(\mathcal{I}_0)) \geq \Omega\left(\frac{n_0^\delta}{\varepsilon_0^{-c_{\text{sens}}}}\right).$$

Moreover,  $\mathcal{I}_0$  is obtained by the reduction chain

$$\Phi_{N_{\text{base}}} \longrightarrow \text{CNF}_{N_{\text{base}}} \longrightarrow \mathcal{G}_0 := \mathcal{T}_{\text{Robust}}(\text{CNF}_{N_{\text{base}}}) \longrightarrow \hat{\mathcal{G}}_0 := \text{red}_{\tau_0}^{\circ}(\mathcal{G}_0) \longrightarrow \mathcal{I}_0 := \text{Boot}_{\varepsilon_0}(\hat{\mathcal{G}}_0),$$

where  $\tau_0 := \varepsilon_0^8$ ,  $N_{\text{base}} := \lfloor n_0^{1/B^*} \rfloor$  for a sufficiently large absolute constant  $B^*$ ,  $\Phi_{N_{\text{base}}}$  is the bounded-degree satisfiable 2CSP hard family from [28],  $\text{CNF}_{N_{\text{base}}}$  is the 3CNF family obtained from  $\Phi_{N_{\text{base}}}$  by replacing each source constraint with a constant-size gadget,  $\text{Boot}_{\varepsilon_0}(\hat{\mathcal{G}}_0)$  denotes the stage-0 bootstrap family obtained by one Dinur–Harsha composition of  $\hat{\mathcal{G}}_0$ , and the direct initialization reduction from  $\Phi_{N_{\text{base}}}$  to  $\mathcal{I}_0$  obtained from this chain has total transformation and recovery losses polynomial in  $1/\varepsilon_0$ .

*Proof. Setup / preliminary cleanup scale.* Fix a sufficiently large absolute constant  $B_*$  and set  $N_{\text{base}} := \lfloor n_0^{1/B_*} \rfloor$ . By Fleming and Yoshida [28], there exists an absolute constant  $\eta_{\text{FY}} \in (0, 1)$  and a swap-closed family  $\Phi_{N_{\text{base}}}$  of satisfiable bounded-degree 2CSP instances on  $\Theta(N_{\text{base}})$  variables, with constant alphabet and constant degree, such that any algorithm attaining value at least  $1 - \eta_{\text{FY}}$  on  $\Phi_{N_{\text{base}}}$  has swap sensitivity  $\Omega(N_{\text{base}})$ .

Encode each binary constraint of  $\Phi \in \Phi_{N_{\text{base}}}$  by a constant-size satisfiability-preserving 3CNF gadget, using  $O(1)$  auxiliary variables and  $O(1)$  clauses per constraint. Because the original alphabet and degree are constant, the resulting family  $\text{CNF}_{N_{\text{base}}}$  is swap closed, has  $\Theta(N_{\text{base}})$  variables and clauses, and remains satisfiable. We write  $T_{\text{CNF}} : \Phi_{N_{\text{base}}} \rightarrow \text{CNF}_{N_{\text{base}}}$  for this encoding map and  $R_{\Phi \leftarrow F}$  for the projection that simply forgets the auxiliary gadget variables and returns an assignment to the underlying 2CSP instance. A single constraint swap in  $\Phi$  changes only  $O(1)$  clauses of  $T_{\text{CNF}}(\Phi)$ . Moreover, every satisfying Boolean assignment of  $F = T_{\text{CNF}}(\Phi)$  maps under  $R_{\Phi \leftarrow F}$  to an assignment of  $\Phi$  of value 1.

Apply Theorem 8.1 to each  $F \in \text{CNF}_{N_{\text{base}}}$ , choosing the constant in that theorem large enough that the target soundness  $\Delta_{\text{lc}} := \varepsilon_0^{16}$  lies above its threshold. Denote the resulting left-predicate family by  $\mathcal{G}_0 := \mathcal{T}_{\text{Robust}}(\text{CNF}_{N_{\text{base}}})$ . By the parameter bounds in Theorem 8.1, the family  $\mathcal{G}_0$  has completeness 1, soundness at most  $\Delta_{\text{lc}}$ , left degree  $\varepsilon_0^{-O(1)}$ , right alphabet size  $\varepsilon_0^{-O(1)}$ , left alphabet size at most  $\exp(\varepsilon_0^{-O(1)})$ , minimum right degree at least  $N_{\text{base}}^{8-o(1)}$ , and at most  $N_{\text{base}}^{O(1)}$  vertices on each side. Moreover, by construction of  $\mathcal{T}_{\text{Robust}}$ , every left vertex of every instance in  $\mathcal{G}_0$  has degree at least 1, and every right vertex has degree at least  $N_{\text{base}}^{8-o(1)}$ , so no vertex is isolated. Moreover, by the bounded-locality clause of Theorem 8.4, for each left vertex of  $\mathcal{G}_0$  the corresponding admissible-neighborhood circuit depends only on the chart component of the left label at  $(\Omega, \tau)$  and one local view of  $\text{Eval}_{z, \tau}$ , and that local view uses only  $q^{\text{cin}} = \varepsilon_0^{-O(1)}$  raw or auxiliary symbols. Hence  $\mathcal{G}_0$  has local description complexity at most  $\varepsilon_0^{-O(1)}$ .

Define

$$\tau_0 := \varepsilon_0^8.$$

Since  $\tau_0^{-1} = \varepsilon_0^{-8} = (\log n_0)^{O(1)}$  while every instance in  $\mathcal{G}_0$  has minimum right degree at least  $N_{\text{base}}^{8-o(1)}$ , we have

$$N_{\text{base}}^{8-o(1)} > c/\tau_0$$

for all sufficiently large  $n_0$ . The family  $\mathcal{G}_0$  has a common underlying graph and common alphabets by Theorem 8.1, so Lemma 11.6 shows that all instances in  $\mathcal{G}_0$  take the degree-reduction branch of  $\text{red}_{\tau_0}^\circ$ . Fix the auxiliary code and expander choices for this family as in Definition 11.7, and define

$$\widehat{\mathcal{G}}_0 := \text{red}_{\tau_0}^\circ(\mathcal{G}_0).$$

Because the reduction is fixed on the whole family, all instances in  $\widehat{\mathcal{G}}_0$  have a common target graph and common alphabets.

Let

$$k_0 = \Theta(\tau_0^{-6} \log |\Sigma_V|), \quad d_0 = \lceil c/\tau_0 \rceil.$$

Since  $\tau_0 = \varepsilon_0^8$  and the source right alphabet size in  $\mathcal{G}_0$  is already  $\varepsilon_0^{-O(1)}$ , both  $k_0$  and  $d_0$  are  $\varepsilon_0^{-O(1)}$ . The proof of Lemma 9.4 therefore yields that every instance in  $\widehat{\mathcal{G}}_0$  has left degree  $\varepsilon_0^{-O(1)}$ , right degree  $d_0 = \varepsilon_0^{-O(1)}$ , right alphabet size  $\varepsilon_0^{-O(1)}$ , left alphabet size at most  $\exp(\varepsilon_0^{-O(1)})$ , and vertex counts  $N_{\text{base}}^{O(1)} \varepsilon_0^{-O(1)}$ . By construction of the alphabet-then-degree reduction, non-isolatedness on both sides is preserved from  $\mathcal{G}_0$  to  $\widehat{\mathcal{G}}_0$ .

We also record directly that  $\widehat{\mathcal{G}}_0$  has local description complexity at most  $\varepsilon_0^{-O(1)}$ . Fix a left vertex  $u$  of an instance in  $\mathcal{G}_0$ , and let  $C_u$  be an admissible-neighborhood circuit of size  $\varepsilon_0^{-O(1)}$  for that vertex. In the alphabet-reduction step at parameter  $\tau_0$ , each incident right label is replaced by a block of length  $k_0$ , so admissibility at  $u$  is checked by verifying that each of the  $\deg_{\mathcal{G}_0}(u) \leq \varepsilon_0^{-O(1)}$  blocks is a codeword and then evaluating  $C_u$  on the decoded tuple. This gives a circuit of size  $\varepsilon_0^{-O(1)}$  because both  $\deg_{\mathcal{G}_0}(u)$  and  $k_0$  are  $\varepsilon_0^{-O(1)}$ . In the subsequent degree-reduction step, each original incident block contributes only  $d_0 = \varepsilon_0^{-O(1)}$  copies at  $u$ . Admissibility in the final cleaned instance is therefore checked by equality tests within each copy group followed by the alphabet-reduced circuit, and the resulting circuit size is still  $\varepsilon_0^{-O(1)}$ .

Choose a sufficiently large absolute constant  $A$  and pad all circuits deciding admissible neighborhood tuples in  $\widehat{\mathcal{G}}_0$  to a common decision complexity  $M_0 := \varepsilon_0^{-A}$ , so that the common right alphabet size of  $\widehat{\mathcal{G}}_0$  is at most  $M_0^A$  and  $\varepsilon_0/2 \geq M_0^{-\alpha}$ . Let  $\mathcal{D}_0$  be the Dinur–Harsha decoder from Lemma 10.3 at decision complexity  $M_0$  with inner soundness parameter  $\delta_{\text{boot}} := \varepsilon_0/2$ .

**Construction of the initialized family.** Define  $\mathcal{I}_0 := \text{Boot}_{\varepsilon_0}(\widehat{\mathcal{G}}_0) := \widehat{\mathcal{G}}_0 \circledast \mathcal{D}_0$ . Because  $\widehat{\mathcal{G}}_0$  already has a common target graph and common alphabets, Theorem 10.5 preserves that property when passing to  $\mathcal{I}_0$ . The decoder parameters satisfy  $q, m, 2^r, |\sigma|, s(M_0) \leq M_0^{O(1)} = \varepsilon_0^{-O(1)}$ . Because every left vertex of  $\widehat{\mathcal{G}}_0$  already has degree at most  $\varepsilon_0^{-O(1)}$ , the trivial proof-degree bound gives

$$d \leq \varepsilon_0^{-O(1)} q \cdot 2^r.$$

With this bound, Theorem 10.5 shows that  $\mathcal{I}_0$  has local description complexity  $\varepsilon_0^{-O(1)}$ , left degree  $\varepsilon_0^{-O(1)}$ , right degree  $\varepsilon_0^{-O(1)}$ , right alphabet size  $\varepsilon_0^{-O(1)}$ , left alphabet size at most  $\exp(\varepsilon_0^{-O(1)})$ , and vertex counts  $N_{\text{base}}^{O(1)} \varepsilon_0^{-O(1)}$ . Because every vertex of every instance in  $\widehat{\mathcal{G}}_0$  is non-isolated, the exact degree formulas from Algorithm 2 show that every vertex of every instance in  $\mathcal{I}_0$  is non-isolated as well. Enlarging  $c_{\text{str}}$  if necessary, all local-description-complexity, degree, alphabet, and non-isolatedness bounds in item (1) follow. Since  $N_{\text{base}}^{O(1)} \varepsilon_0^{-O(1)} \leq n_0 \varepsilon_0^{-c_{\text{size}}}$  for sufficiently large  $B_*$  after enlarging  $c_{\text{size}}$ , item (2) follows as well. The gain from this first composition is therefore not a dramatic improvement in the displayed size bounds by itself, but that it puts the initialized family into the same form used later in the iteration, with target recovery threshold  $9\sqrt{\varepsilon_0}$  and with the three constants controlling swaps and perturbations already polynomial in  $1/\varepsilon_0$ .

**Recovery from repeated independent trials.** For the direct initialization reduction, define the map

$$\text{Inst}_0(\Phi) := \text{Boot}_{\varepsilon_0} \left( \text{red}_{\tau_0}^{\circ} \left( \mathcal{T}_{\text{Robust}}(T_{\text{CNF}}(\Phi)) \right) \right) \in \mathcal{I}_0.$$

For each  $\Phi \in \Phi_{N_{\text{base}}}$ , write

$$F := T_{\text{CNF}}(\Phi), \quad G_F := \mathcal{T}_{\text{Robust}}(F), \quad \widehat{G}_F := \text{red}_{\tau_0}^{\circ}(G_F).$$

Set

$$R_{0,F}^{\text{red}} := R_{\tau_0,G_F}^{\text{red}}, \quad \Delta_0 := \Delta_{\text{lc}}/2,$$

and let  $\mathcal{R}_{\text{Comp},0,F}$  denote the recovery map from Theorem 10.5 for the outer instance  $\widehat{G}_F$ . One trial of the recovery will first move from  $\text{Inst}_0(\Phi)$  back through the composition and cleanup layers to a labeling of  $G_F$ , then use Theorem 8.10 to produce, with inverse-polynomial probability, a satisfying Boolean assignment of  $F$ , and finally project away the gadget variables to obtain an assignment of  $\Phi$  of value 1. The only reason the

final guarantee drops from value 1 to  $1 - \eta_{FY}$  is that this single trial succeeds with only inverse-polynomial probability, so we repeat it independently and then apply the selector to amplify that success event. Define the single-trial candidate map

$$\text{Hit}_{0,\Phi}^{(1)} := R_{\Phi \leftarrow F} \circ \mathcal{H}_{\text{lc},F,\Delta_0} \circ R_{0,F}^{\text{red}} \circ \mathcal{R}_{\text{Comp},0,F}.$$

Fix

$$\eta_1 := \eta_{FY}/4, \quad \eta_2 := \eta_{FY}/2, \quad I_0 := [1 - \eta_2, 1 - \eta_1],$$

and let  $\text{Sel}_{\Phi}^{I_0}$  be the selection map on source assignments from Lemma 11.9.

Let  $A$  be any algorithm for  $\text{SwapClo}(\mathcal{I}_0)$  that attains expected value at least  $9\sqrt{\varepsilon_0}$  on every satisfiable instance in  $\text{SwapClo}(\mathcal{I}_0)$ , and let  $S := \text{SwapSens}(A, \text{SwapClo}(\mathcal{I}_0))$ . For a fixed source instance  $\Phi \in \Phi_{N_{\text{base}}}$ , let  $\mathbf{y} := A(\text{Inst}_0(\Phi))$ . One trial of the candidate map first applies the composition recovery  $\mathcal{R}_{\text{Comp},0,F}$  to move from  $\text{Inst}_0(\Phi) = \widehat{G}_F \otimes \mathcal{D}_0$  back to  $\widehat{G}_F$ , then applies  $R_{0,F}^{\text{red}}$  to move from  $\widehat{G}_F$  back to  $G_F$ , and finally runs the sampler from Theorem 8.10 at threshold  $\Delta_0$ . The candidate list map  $\text{Cand}_{0,\Phi}$  will consist of  $T$  independent copies of this single trial, and the final recovery will be  $\text{Rec}_{0,\Phi} := \text{Sel}_{\Phi}^{I_0} \circ \text{Cand}_{0,\Phi}$ .

We first lower bound the success probability of a single trial. Because  $\text{Inst}_0(\Phi)$  is satisfiable and lies in  $\text{SwapClo}(\mathcal{I}_0)$ , the output  $\mathbf{y}$  of  $A$  has expected value at least  $9\sqrt{\varepsilon_0}$ . Set the outer threshold for the composition layer to

$$4\sqrt{\tau_0} = 4\varepsilon_0^4.$$

Since  $L \leq 2/\delta_{\text{boot}} \leq 4/\varepsilon_0$ , we have

$$\delta_{\text{boot}} + L \cdot 4\sqrt{\tau_0} \leq \frac{\varepsilon_0}{2} + \frac{4}{\varepsilon_0} \cdot 4\varepsilon_0^4 \leq \varepsilon_0$$

for all sufficiently small  $\varepsilon_0$ . Because  $\varepsilon_0 \leq 9\sqrt{\varepsilon_0}$  for all sufficiently small  $\varepsilon_0$ , Theorem 10.5 applied to the outer instance  $\widehat{G}_F$  yields a random labeling

$$\widehat{\boldsymbol{\pi}}_F := \mathcal{R}_{\text{Comp},0,F}(\mathbf{y})$$

of  $\widehat{G}_F$  satisfying

$$\mathbf{E}[\text{val}_{\widehat{G}_F}(\widehat{\boldsymbol{\pi}}_F)] \geq 4\sqrt{\tau_0}.$$

Now apply the cleanup recovery map  $R_{0,F}^{\text{red}}$  to  $\widehat{\boldsymbol{\pi}}_F$ . The soundness statement for  $\text{red}_{\tau_0}^{\circ}$  from Section 11.2 then gives a random labeling

$$\boldsymbol{\pi}_F := R_{0,F}^{\text{red}}(\widehat{\boldsymbol{\pi}}_F)$$

of  $G_F$  satisfying

$$\mathbf{E}[\text{val}_{G_F}(\boldsymbol{\pi}_F)] \geq \tau_0.$$

Since  $\tau_0 = \varepsilon_0^8 \geq \Delta_{\text{lc}} = \varepsilon_0^{16}$  for all sufficiently small  $\varepsilon_0$ , the same labeling also satisfies

$$\mathbf{E}[\text{val}_{G_F}(\boldsymbol{\pi}_F)] \geq \Delta_{\text{lc}}.$$

Since  $\text{val}_{G_F}(\boldsymbol{\pi}_F) \in [0, 1]$  and  $\Delta_0 = \Delta_{\text{lc}}/2$ , we have  $\Pr[\text{val}_{G_F}(\boldsymbol{\pi}_F) \geq \Delta_0] \geq \Delta_0$ . For all sufficiently large  $n_0$ , the threshold  $\Delta_0$  satisfies the hypotheses of Theorem 8.10; indeed,  $\Delta_0 = \varepsilon_0^{16}$  dominates both the point-object threshold and the light-bad-tuple threshold once the exponent from Theorem 8.1 and the initialization constant  $c_{\text{init}}$  are fixed large enough. Conditional on the event  $\text{val}_{G_F}(\boldsymbol{\pi}_F) \geq \Delta_0$ , Theorem 8.10 therefore returns a satisfying Boolean assignment to  $F$  with probability at least  $\Delta_0^{C_{\text{hit}}}$ . Projecting away the auxiliary

gadget variables via  $R_{\Phi \leftarrow F}$  then gives an assignment of  $\Phi$  of value 1. The later target  $1 - \eta_{FY}$  comes only after averaging over repeated independent trials and the selector, since a single trial succeeds with only inverse-polynomial probability. Hence one independent trial of  $\text{Hit}_{0,\Phi}^{(1)}$  succeeds with probability at least  $p_0 := \Delta_0^{C_{\text{hit}}+1} = \varepsilon_0^{O(1)}$ .

Choose  $T := \left\lceil C p_0^{-1} \log \frac{4}{\eta_{FY}} \right\rceil$  for a sufficiently large absolute constant  $C$ . Let

$$\text{Cand}_{0,\Phi}(\mathbf{y}) := (b^{(1)}, \dots, b^{(T)})$$

be the  $T$ -tuple of independent outputs of  $\text{Hit}_{0,\Phi}^{(1)}$  on the common input labeling  $\mathbf{y}$ . Then

$$\Pr[\exists j \in [T] \text{ with } \text{val}_{\Phi}(b^{(j)}) = 1] \geq 1 - (1 - p_0)^T \geq 1 - \eta_{FY}/4.$$

By item (1) of Lemma 11.9, whenever this event occurs the selector  $\text{Sel}_{\Phi}^{I_0}$  outputs an assignment of source value at least  $1 - \eta_2 = 1 - \eta_{FY}/2$ . Therefore

$$\mathbf{E}[\text{val}_{\Phi}(\text{Rec}_{0,\Phi}(\mathbf{y}))] \geq (1 - \eta_{FY}/4)(1 - \eta_{FY}/2) > 1 - \eta_{FY}.$$

Thus the algorithm  $B(\Phi) := \text{Rec}_{0,\Phi}(A(\text{Inst}_0(\Phi)))$  attains value at least  $1 - \eta_{FY}$  on every source instance  $\Phi \in \Phi_{N_{\text{base}}}$ .

**Update of  $C_T, C_R, D$ .** The transformation map  $\Phi \mapsto \text{Inst}_0(\Phi)$  has instance-distance constant  $C_T^{(0)} = \varepsilon_0^{-O(1)}$ . Indeed, one source swap in  $\Phi$  changes only  $O(1)$  clauses of  $T_{\text{CNF}}(\Phi)$ . By Theorems 8.1 and 8.4, the robust label-cover step turns those clause changes into only  $\varepsilon_0^{-O(1)}$  left-predicate changes. The cleanup  $\text{red}_{\tau_0}^{\circ}$  then multiplies that by at most  $\varepsilon_0^{-O(1)}$ , and item (3) of Theorem 10.5 contributes one more factor  $\varepsilon_0^{-O(1)}$ .

For the recovery side, let  $L_{\text{hit}}$  be the constant for how the single-sample map  $\text{Hit}_{0,\Phi}^{(1)}$  changes under assignment perturbations on a fixed source instance  $\Phi$ . To bound  $L_{\text{hit}}$ , follow one sample through the maps in the order

$$\mathcal{R}_{\text{Comp},0,F}, \quad R_{0,F}^{\text{red}}, \quad \mathcal{H}_{\text{lc},F,\Delta_0}, \quad R_{\Phi \leftarrow F}.$$

The first map changes by at most  $\varepsilon_0^{-O(1)}$  under assignment perturbations by Theorem 10.5, the second changes by at most  $\varepsilon_0^{-O(1)}$  by Section 11.2, the sampler from Theorem 8.10 has assignment-perturbation constant  $q^{O(1)} = \varepsilon_0^{-O(1)}$  by item (2) of that proposition, and changing one Boolean input coordinate of the final projection changes only  $O(1)$  source coordinates. Hence  $L_{\text{hit}} = \varepsilon_0^{-O(1)}$ . The  $T$ -tuple map  $\text{Cand}_{0,\Phi}$  therefore has assignment-perturbation constant at most  $T L_{\text{hit}}$  with respect to the sum of the coordinatewise Hamming distances on the candidate list. Applying item (2) of Lemma 11.9 yields

$$C_R^{(0)} \leq C_{\text{sel}} T L_{\text{hit}} = \varepsilon_0^{-O(1)}.$$

For the bound for how the recovery changes under source swaps, let  $\Phi, \tilde{\Phi} \in \Phi_{N_{\text{base}}}$  differ by one source swap, and write

$$\tilde{F} := T_{\text{CNF}}(\tilde{\Phi}), \quad G_{\tilde{F}} := \mathcal{T}_{\text{Robust}}(\tilde{F}), \quad \widehat{G}_{\tilde{F}} := \text{red}_{\tau_0}^{\circ}(G_{\tilde{F}}).$$

Again follow one sample through the maps in order. On the common cleaned family  $\widehat{\mathcal{G}}_0$ , the composition recovery changes by at most  $\varepsilon_0^{-O(1)}$  under one source swap. On the common raw family  $\mathcal{G}_0$ , the cleanup

recovery changes by at most  $\varepsilon_0^{-O(1)}$  under one source swap. Finally, one source swap in  $\Phi$  changes  $F$  by only  $O(1)$  clause swaps, so item (3) of Theorem 8.10 contributes another  $\varepsilon_0^{-O(1)}$  single-sample source-swap bound. Composing these bounds with the projection  $R_{\Phi \leftarrow F}$ , which changes only  $O(1)$  source coordinates when one Boolean input coordinate changes, gives a single-sample source-swap constant  $D_{\text{hit}} = \varepsilon_0^{-O(1)}$  for  $\text{Hit}_{0,\Phi}^{(1)}$ . Consequently, the candidate-list map  $\text{Cand}_{0,\Phi}$  has source-swap constant at most  $TD_{\text{hit}}$  under the sum metric on the list coordinates. Using item (2) of Lemma 11.9 on a fixed source instance and then item (3) on the selector itself yields

$$D_0 \leq C_{\text{sel}} TD_{\text{hit}} + C_{\text{sel}} T = \varepsilon_0^{-O(1)}.$$

**Pulling back the sensitivity lower bound.** The direct initialization reduction therefore satisfies the hypotheses of Corollary 11.4. Here the source family is  $\Phi_{N_{\text{base}}}$ , the target family is  $\text{SwapClo}(\mathcal{I}_0)$ , and the constants are  $C_T^{(0)}, C_R^{(0)}, D_0$ . Since  $B$  attains value at least  $1 - \eta_{\text{FY}}$  on  $\Phi_{N_{\text{base}}}$ , the Fleming–Yoshida lower bound gives  $\text{SwapSens}(B, \Phi_{N_{\text{base}}}) \geq \Omega(N_{\text{base}})$ . Applying Corollary 11.4 yields

$$S \geq \frac{\Omega(N_{\text{base}}) - D_0}{C_T^{(0)} C_R^{(0)}}.$$

Because  $D_0 = \varepsilon_0^{-O(1)} = (\log n_0)^{O(1)} = o(N_{\text{base}})$  after enlarging  $B_*$  if necessary, this implies  $S \geq \Omega\left(\frac{N_{\text{base}}}{\varepsilon_0^{-O(1)}}\right)$ . Because  $A$  was arbitrary, we conclude that

$$\text{SwapSens}_{9\sqrt{\varepsilon_0}}(\text{SwapClo}(\mathcal{I}_0)) \geq \Omega\left(\frac{N_{\text{base}}}{\varepsilon_0^{-O(1)}}\right).$$

Since  $N_{\text{base}} = n_0^{1/B_* + o(1)}$ , after shrinking  $\delta$  below  $1/B_*$  and enlarging  $c_{\text{sens}}$ , this is exactly item (3).  $\square$

## 11.4 One-Stage Update of the Reduction Data

We now formalize stage- $i$  reduction data and the one-stage update. Fix the absolute constants from Proposition 11.10, and denote them by

$$c_{\text{init}}, c_{\text{str}}, c_{\text{size}}, c_{\text{sens}}, \delta.$$

For every sufficiently large base size  $n_0$ , set

$$\varepsilon_0 := (\log n_0)^{-c_{\text{init}}}.$$

Let  $\mathcal{I}_0$  be the initialized family supplied by Proposition 11.10. Throughout the rest of this subsection,  $n_0$  denotes the initial target size of this family. Let  $N_{\text{base}}$  denote the corresponding base 2CSP size. The update uses

$$\xi_i := \varepsilon_i^{1/4} \quad \text{and} \quad \varepsilon_{i+1} := 20\xi_i.$$

**Definition 11.11** (Stage- $i$  reduction data). Fix  $i \geq 0$ . Let  $\mathcal{I}_i$  be a family of left-predicate label cover instances whose instances all have a common target graph and common alphabets. We say that the tuple

$$(\mathcal{I}_i, \text{Inst}_i, \{\text{Rec}_{i,\Phi}\}_{\Phi \in \Phi_{N_{\text{base}}}}, C_T^{(i)}, C_R^{(i)}, D_i)$$

is stage- $i$  reduction data on  $\mathcal{I}_i$  if:

1.  $\text{Inst}_i(\Phi)$  is satisfiable for every  $\Phi \in \Phi_{N_{\text{base}}}$ .
2. If a labeling of  $\text{Inst}_i(\Phi)$  has expected value at least  $9\sqrt{\varepsilon_i}$ , then  $\text{Rec}_{i,\Phi}$  recovers an assignment to  $\Phi$  of expected value at least  $1 - \eta_{\text{FY}}$ .
3. For every one-swap pair  $\Phi, \tilde{\Phi} \in \Phi_{N_{\text{base}}}$ ,

$$\text{SwapDist}(\text{Inst}_i(\Phi), \text{Inst}_i(\tilde{\Phi})) \leq C_T^{(i)}.$$

4. For every fixed  $\Phi \in \Phi_{N_{\text{base}}}$  and every two assignments  $\pi, \tilde{\pi}$  on  $\text{Inst}_i(\Phi)$ ,

$$\text{EMD}(\text{Rec}_{i,\Phi}(\pi), \text{Rec}_{i,\Phi}(\tilde{\pi})) \leq C_R^{(i)} d_{\text{H}}(\pi, \tilde{\pi}).$$

5. For every one-swap pair  $\Phi, \tilde{\Phi} \in \Phi_{N_{\text{base}}}$ , the two target instances  $\text{Inst}_i(\Phi)$  and  $\text{Inst}_i(\tilde{\Phi})$  have a common target graph and common alphabets, and for every assignment  $\pi$  on that common target domain,

$$\text{EMD}(\text{Rec}_{i,\Phi}(\pi), \text{Rec}_{i,\tilde{\Phi}}(\pi)) \leq D_i.$$

The initialized family from Proposition 11.10 therefore supplies stage-0 reduction data. The next table summarizes the formal stage transition proved below. In using this update, we repeatedly rely on two facts: once the relevant swap closure has a common target graph and common alphabets, Lemma 11.1 upgrades a single-swap bound for how the recovery changes under source swaps to a uniform bound over the family, and on each stage family the alphabet-then-degree reduction from Section 9 is fixed after the auxiliary code and expander choices are frozen even though the recovery map  $R_{\varepsilon,I}^{\text{red}}$  still depends on the source instance.

Table 4: Summary of the stage- $i$  update.

Family	Threshold update	Structure	Tracking
$\mathcal{I}_i$	Target value $\chi_i = 9\sqrt{\varepsilon_i}$	Common target graph and alphabets; no isolated vertices; local description complexity, degrees, and right alphabet $\leq \varepsilon_i^{-c_{\text{str}}}$ ; size $\leq n_0 \prod_{t=0}^i \varepsilon_t^{-c_{\text{size}}}$	Stage- $i$ reduction data $(C_T^{(i)}, C_R^{(i)}, D_i)$
$\mathcal{J}_{i+1} = \mathcal{I}_i \circledast \mathcal{D}_i$	Decoder soundness $\xi_i = \varepsilon_i^{1/4}$ ; value $\geq \varepsilon_{i+1} = 20\xi_i$ recovers to $\chi_i$	Common target graph and alphabets persist; no isolated vertices persist; all composition blow-ups are $\varepsilon_i^{-O(1)}$	Intermediate reduction data after composition
$\mathcal{I}_{i+1} = \text{red}_{\varepsilon_{i+1}}^{\circ}(\mathcal{J}_{i+1})$	Next-stage target value $9\sqrt{\varepsilon_{i+1}}$	Common target graph and alphabets persist; no isolated vertices persist; local description complexity, degrees, and right alphabet $\leq \varepsilon_{i+1}^{-c_{\text{str}}}$ ; size $\leq n_0 \prod_{t=0}^{i+1} \varepsilon_t^{-c_{\text{size}}}$	Updated reduction data $(C_T^{(i+1)}, C_R^{(i+1)}, D_{i+1})$

The next proposition advances such data by one stage through composition followed by the fixed alphabet-then-degree reduction, and Lemma 11.13 then iterates this update from stage 0.

**Proposition 11.12** (One-stage update of the reduction data). There exists an absolute constant  $c_0 > 0$  with the following property. Fix  $i \geq 0$  and suppose  $\varepsilon_i < c_0$ . Assume that  $\mathcal{I}_i$  is a family of left-predicate label

cover instances whose local description complexity, left degree, right degree, and right alphabet size are all at most  $\varepsilon_i^{-c_{\text{str}}}$ , whose left alphabet size is at most  $\exp(\varepsilon_i^{-c_{\text{str}}})$ , whose vertex counts satisfy

$$|U|, |V| \leq n_0 \cdot \prod_{t=0}^i \varepsilon_t^{-c_{\text{size}}}.$$

Assume further that every left and right vertex has degree at least 1. Assume also that all instances in  $\mathcal{I}_i$  are defined on a common graph and domain, and that

$$(\mathcal{I}_i, \text{Inst}_i, \{\text{Rec}_{i,\Phi}\}_{\Phi \in \Phi_{N_{\text{base}}}}, C_T^{(i)}, C_R^{(i)}, D_i)$$

is stage- $i$  reduction data on  $\mathcal{I}_i$ , with

$$C_T^{(i)}, C_R^{(i)}, D_i \leq \prod_{t=0}^i \varepsilon_t^{-O(1)}.$$

Choose a sufficiently large absolute constant  $A$  and define

$$M_i := \varepsilon_i^{-A}, \quad \xi_i := \varepsilon_i^{1/4}, \quad \varepsilon_{i+1} := 20\xi_i.$$

Let  $\mathcal{D}_i$  be the Dinur–Harsha decoder from Lemma 10.3 at decision complexity  $M_i$  and inner soundness parameter  $\xi_i$ , and define

$$\mathcal{J}_{i+1} := \mathcal{I}_i \otimes \mathcal{D}_i = \{I \otimes \mathcal{D}_i : I \in \mathcal{I}_i\}, \quad \mathcal{I}_{i+1} := \text{red}_{\varepsilon_{i+1}}^\circ(\mathcal{J}_{i+1}).$$

Then there exist a map

$$\text{Inst}_{i+1} : \Phi_{N_{\text{base}}} \rightarrow \mathcal{I}_{i+1},$$

recoveries, one for each source instance,

$$\text{Rec}_{i+1,\Phi}, \quad \Phi \in \Phi_{N_{\text{base}}},$$

and constants

$$C_T^{(i+1)}, C_R^{(i+1)}, D_{i+1}$$

such that

$$(\mathcal{I}_{i+1}, \text{Inst}_{i+1}, \{\text{Rec}_{i+1,\Phi}\}_{\Phi \in \Phi_{N_{\text{base}}}}, C_T^{(i+1)}, C_R^{(i+1)}, D_{i+1})$$

is stage- $(i+1)$  reduction data on  $\mathcal{I}_{i+1}$ , and

$$C_T^{(i+1)}, C_R^{(i+1)}, D_{i+1} \leq \prod_{t=0}^{i+1} \varepsilon_t^{-O(1)}.$$

In particular, after enlarging  $c_{\text{str}}, c_{\text{size}}, c_{\text{sens}}$  if necessary, items (1)–(3) of Lemma 11.13 hold at stage  $i+1$ .

*Proof.* Assume the stated stage- $i$  reduction data. We first build  $\mathcal{J}_{i+1}$  and  $\mathcal{I}_{i+1}$  and record the bounds on local description complexity, degrees, alphabet sizes, and vertex counts needed for the next stage. We then define  $\text{Inst}_{i+1}$  and  $\text{Rec}_{i+1,\Phi}$ , check satisfiability, and prove the soundness statement, verifying items (1) and (2) in the definition. After that we use Lemma 11.2 twice, together with Lemma 11.1, to update the tracking constants through the composition layer and then the fixed alphabet-then-degree reduction, which

verifies items (3)–(5). Finally we pull back the sensitivity lower bound through the new reduction data to obtain item (3) of Lemma 11.13 at stage  $i + 1$ .

**Setup / decoder parameters.** Since the local description complexity, the left degree, and the right alphabet of  $\mathcal{I}_i$  are all at most  $\varepsilon_i^{-c_{\text{str}}}$ , every circuit deciding admissible neighborhood tuples in  $\mathcal{I}_i$  can be padded to decision complexity at most  $M_i$ . Taking  $A$  large enough ensures both  $|\Sigma_V| \leq M_i^\gamma$  and

$$\xi_i = \varepsilon_i^{1/4} \geq M_i^{-\alpha}$$

for all sufficiently small  $\varepsilon_i$ . Let  $\mathcal{D}_i$  be the Dinur–Harsha decoder from Lemma 10.3 at decision complexity  $M_i$  and inner soundness parameter  $\xi_i$ . Its parameters satisfy

$$q_i, m_i, 2^{r_i}, |\sigma_i|, s(M_i) \leq M_i^{O(1)} = \varepsilon_i^{-O(1)}.$$

The trivial proof-degree bound gives

$$d_i \leq \varepsilon_i^{-c_{\text{str}}} q_i 2^{r_i} \leq \varepsilon_i^{-O(1)}.$$

Let  $d_{U,i}$  and  $d_{V,i}$  be the degree bounds for  $\mathcal{I}_i$ , so

$$d_{U,i}, d_{V,i} \leq \varepsilon_i^{-c_{\text{str}}}.$$

**Construction of the next family.** Because the stage family  $\mathcal{I}_i$  already has a common graph and domain, Theorem 10.5 shows that  $\mathcal{J}_{i+1}$  also has a common graph and domain. By Theorem 10.5, the family  $\mathcal{J}_{i+1}$  has local description complexity at most  $\varepsilon_i^{-O(1)}$ , has left degree at most  $\varepsilon_i^{-O(1)}$ , right degree at most  $\varepsilon_i^{-O(1)}$ , right alphabet size at most  $\varepsilon_i^{-O(1)}$ , left alphabet size at most  $\exp(\varepsilon_i^{-O(1)})$ , and its vertex counts are obtained by multiplying those of  $\mathcal{I}_i$  by at most  $\varepsilon_i^{-O(1)}$ . Because every vertex of every instance in  $\mathcal{I}_i$  is non-isolated, the exact degree formulas from Algorithm 2 show that every vertex of every instance in  $\mathcal{J}_{i+1}$  is non-isolated as well. Since

$$\varepsilon_i = \left(\frac{\varepsilon_{i+1}}{20}\right)^4,$$

every polynomial in  $\varepsilon_i^{-1}$  is also a polynomial in  $\varepsilon_{i+1}^{-1}$ . The family  $\mathcal{I}_{i+1}$  is therefore defined by taking the common branch from Lemma 11.6 together with the fixed auxiliary choices from Definition 11.7. Since  $\mathcal{J}_{i+1}$  itself has a common graph and domain, fixing these reduction choices across the family gives the whole image family  $\mathcal{I}_{i+1}$  a common target graph and common alphabets. Hence  $\mathcal{I}_{i+1}$  has local description complexity at most  $\varepsilon_{i+1}^{-O(1)}$  by Lemma 11.8; here the input right alphabet size of  $\mathcal{J}_{i+1}$  is already  $\varepsilon_i^{-O(1)}$ , so the polynomial dependence on  $|\Sigma_V|$  is still polynomial in  $\varepsilon_{i+1}^{-1}$ . The same construction and the bounds from Section 9 imply that its left degree, right degree, right alphabet, and vertex blow-up are all bounded by  $\varepsilon_{i+1}^{-O(1)}$ , while the left alphabet remains at most  $\exp(\varepsilon_{i+1}^{-O(1)})$ . By construction of the alphabet-then-degree reduction, non-isolatedness on both sides is preserved from  $\mathcal{J}_{i+1}$  to  $\mathcal{I}_{i+1}$ . Enlarging  $c_{\text{str}}$  and  $c_{\text{size}}$  if necessary establishes items (1) and (2) of Lemma 11.13 at stage  $i + 1$ .

For each  $\Phi \in \Phi_{N_{\text{base}}}$ , write

$$J_{i+1,\Phi} := \text{Inst}_i(\Phi) \otimes \mathcal{D}_i, \quad R_{i+1,\Phi}^{\text{red}} := R_{\varepsilon_{i+1}, J_{i+1,\Phi}}^{\text{red}}.$$

Define

$$\widehat{\text{Inst}}_{i+1}(\Phi) := J_{i+1,\Phi}, \quad \text{Inst}_{i+1}(\Phi) := \text{red}_{\varepsilon_{i+1}}^{\circ}(J_{i+1,\Phi}).$$

For each  $\Phi \in \Phi_{N_{\text{base}}}$ , let  $\mathcal{R}_{\text{Comp},i,\Phi}$  denote the recovery map from Theorem 10.5 for the outer instance  $\text{Inst}_i(\Phi)$ . Define

$$\widehat{\text{Rec}}_{i+1,\Phi} := \text{Rec}_{i,\Phi} \circ \mathcal{R}_{\text{Comp},i,\Phi}, \quad \text{Rec}_{i+1,\Phi} := \widehat{\text{Rec}}_{i+1,\Phi} \circ R_{i+1,\Phi}^{\text{red}}.$$

Because  $\text{Inst}_i(\Phi)$  is satisfiable, item (1) of Theorem 10.5 shows that  $J_{i+1,\Phi}$  is satisfiable, and the fixed alphabet-then-degree reduction preserves satisfiability. Thus  $\text{Inst}_{i+1}(\Phi)$  is satisfiable for every  $\Phi \in \Phi_{N_{\text{base}}}$ . This defines  $\text{Inst}_{i+1}$  and  $\text{Rec}_{i+1,\Phi}$  and verifies item (1) in the definition.

**Soundness transfer.** We next verify item (2) in the definition. For soundness, fix  $\Phi \in \Phi_{N_{\text{base}}}$  and suppose a labeling of  $\text{Inst}_{i+1}(\Phi)$  has expected value at least  $9\sqrt{\varepsilon_{i+1}}$ . Since  $4\sqrt{\varepsilon_{i+1}} \leq 9\sqrt{\varepsilon_{i+1}}$ , the recovery map  $R_{i+1,\Phi}^{\text{red}}$  yields a labeling of  $J_{i+1,\Phi}$  with expected value at least  $\varepsilon_{i+1}$ . Apply Theorem 10.5 to the outer instance  $\text{Inst}_i(\Phi)$  with

$$\chi_i := 9\sqrt{\varepsilon_i}$$

and

$$\xi_i := \varepsilon_i^{1/4}.$$

Since  $L \leq 2/\xi_i$ , we have

$$\xi_i + \frac{2}{\xi_i} \cdot \chi_i \leq \varepsilon_i^{1/4} + 18\varepsilon_i^{1/4} \leq 20\varepsilon_i^{1/4} = \varepsilon_{i+1}.$$

Therefore  $\mathcal{R}_{\text{Comp},i,\Phi}$  recovers a labeling of  $\text{Inst}_i(\Phi)$  with expected value at least  $9\sqrt{\varepsilon_i}$ . Applying the stage- $i$  soundness guarantee for  $\text{Rec}_{i,\Phi}$  now gives an assignment to  $\Phi$  of expected value at least  $1 - \eta_{\text{FY}}$ .

**Update of  $C_T, C_R, D$ .** We now verify items (3)–(5) in the definition. Let

$$C_{T,\text{Comp}}^{(i)} := d_{U,i} 2^{r_i}, \quad C_{R,\text{Comp}}^{(i)} := d_i (1 + d_{V,i}) 2^{-r_i}, \quad D_{\text{Comp}}^{(i),\text{single}} := d_{U,i} (1 + d_{V,i}) + 1.$$

By the stage- $i$  bounds on degrees, alphabet sizes, and local description complexity, together with the decoder bounds, all three are at most  $\varepsilon_i^{-O(1)}$ . For every one-swap pair  $\Phi, \tilde{\Phi} \in \Phi_{N_{\text{base}}}$ , the stage- $i$  reduction data give

$$\text{SwapDist}(\text{Inst}_i(\Phi), \text{Inst}_i(\tilde{\Phi})) \leq C_T^{(i)}.$$

Because the whole stage family  $\mathcal{I}_i$  has a common target graph and common alphabets, so does its swap closure  $\text{SwapClo}(\mathcal{I}_i)$ , and the same one-outer-swap calculation for how the composition recovery changes under source swaps from item (5) of Theorem 10.5 applies throughout that swap closure. Applying Lemma 11.1 to the swap-closed family  $\text{SwapClo}(\mathcal{I}_i)$  therefore extends that single-swap bound to a uniform composition bound for how the recovery changes under source swaps:

$$D_{\text{Comp}}^{(i),\text{fam}} \leq C_T^{(i)} D_{\text{Comp}}^{(i),\text{single}}.$$

We first treat the composition layer, which produces intermediate reduction data with  $\widehat{\text{Inst}}_{i+1}, \widehat{\text{Rec}}_{i+1,\Phi}, \widehat{C}_T^{(i+1)}, \widehat{C}_R^{(i+1)}$ , and  $\widehat{D}^{(i+1)}$ . Apply Lemma 11.2 to the stage- $i$  reduction data and the composition layer  $I \mapsto I \otimes \mathcal{D}_i$ . This yields the intermediate tracking constants

$$\widehat{C}_T^{(i+1)} = C_T^{(i)} C_{T,\text{Comp}}^{(i)}, \quad \widehat{C}_R^{(i+1)} = C_R^{(i)} C_{R,\text{Comp}}^{(i)},$$

and

$$\widehat{D}^{(i+1)} \leq D_i + C_R^{(i)} D_{\text{Comp}}^{(i),\text{fam}}.$$

Let  $C_{T,\text{red}}^{(i+1)}, C_{R,\text{red}}^{(i+1)} \leq \varepsilon_{i+1}^{-O(1)}$  denote the constants supplied by the existing locality bounds for how the transformation changes under source swaps and how the recovery changes under assignment perturbations for the alphabet-then-degree reduction. Let

$$D_{\text{red}}^{(i+1),\text{single}}$$

denote the bound for how the reduction recovery layer changes under a single pre-reduction swap on the family  $\mathcal{J}_{i+1}$ . By Section 11.2, the reduction  $\text{red}_{\varepsilon_{i+1}}^\circ$  itself is fixed on the stage family  $\mathcal{J}_{i+1}$ , even though the recovery maps  $R_{i+1,\Phi}^{\text{red}}$  still depend on the source instance. For a single pre-reduction swap, the same locality calculation as in the alphabet-reduction proof changes the recovered assignment by at most  $1 + \Delta_U/\delta_V$ , where  $\Delta_U$  and  $\delta_V$  are the maximum left degree and minimum right degree of the relevant pre-reduction instance in  $\mathcal{J}_{i+1}$ . Because  $\Delta_U \leq \varepsilon_i^{-O(1)}$  on  $\mathcal{J}_{i+1}$  and, as noted above, every right vertex there has degree at least 1, this gives

$$D_{\text{red}}^{(i+1),\text{single}} \leq \varepsilon_{i+1}^{-O(1)}.$$

Moreover, for every one-swap pair  $\Phi, \tilde{\Phi} \in \Phi_{N_{\text{base}}}$ ,

$$\text{SwapDist}(J_{i+1,\Phi}, J_{i+1,\tilde{\Phi}}) \leq C_T^{(i)} C_{T,\text{Comp}}^{(i)}.$$

Since  $\mathcal{J}_{i+1}$  has a common target graph and common alphabets, so does its swap closure  $\text{SwapClo}(\mathcal{J}_{i+1})$ , and the same fixed reduction choices apply on that swap-closed family. Therefore Lemma 11.1 extends the corresponding single-swap bound to a uniform reduction bound for how the recovery changes under source swaps:

$$D_{\text{red}}^{(i+1),\text{fam}} \leq C_T^{(i)} C_{T,\text{Comp}}^{(i)} D_{\text{red}}^{(i+1),\text{single}}.$$

We then pass from these intermediate reduction data to the final ones via the alphabet-then-degree reduction. Apply Lemma 11.2 again, now to the intermediate reduction data

$$\widehat{\text{Inst}}_{i+1}(\Phi) = J_{i+1,\Phi}, \quad \widehat{\text{Rec}}_{i+1,\Phi} = \text{Rec}_{i,\Phi} \circ \mathcal{R}_{\text{Comp},i,\Phi},$$

and the reduction layer. This gives

$$C_T^{(i+1)} = \widehat{C}_T^{(i+1)} C_{T,\text{red}}^{(i+1)}, \quad C_R^{(i+1)} = \widehat{C}_R^{(i+1)} C_{R,\text{red}}^{(i+1)},$$

and

$$D_{i+1} \leq \widehat{D}^{(i+1)} + \widehat{C}_R^{(i+1)} D_{\text{red}}^{(i+1),\text{fam}}.$$

Because every factor on the right-hand sides above is polynomial in  $\varepsilon_i^{-1}$  or  $\varepsilon_{i+1}^{-1}$ , and

$$\varepsilon_i = \left( \frac{\varepsilon_{i+1}}{20} \right)^4,$$

we conclude that

$$C_T^{(i+1)}, C_R^{(i+1)}, D_{i+1} \leq \prod_{t=0}^{i+1} \varepsilon_t^{-O(1)}.$$

Thus the constructed data satisfy items (3)–(5) in the definition with the stated uniform bound over the family.

**Pulling back the sensitivity lower bound.** Let  $A$  be any algorithm on  $\text{SwapClo}(\mathcal{I}_{i+1})$  attaining expected value at least  $9\sqrt{\varepsilon_{i+1}}$  on every satisfiable instance, and let

$$S := \text{SwapSens}(A, \text{SwapClo}(\mathcal{I}_{i+1})).$$

Define

$$B(\Phi) := \text{Rec}_{i+1, \Phi}(A(\text{Inst}_{i+1}(\Phi))).$$

Since  $\text{Inst}_{i+1}(\Phi) \in \mathcal{I}_{i+1} \subseteq \text{SwapClo}(\mathcal{I}_{i+1})$  and is satisfiable, the algorithm  $A$  is applicable there, and the soundness transfer just proved shows that  $B$  attains value at least  $1 - \eta_{\text{FY}}$  on  $\Phi_{N_{\text{base}}}$ . The stage- $(i+1)$  data just constructed therefore satisfy the hypotheses of Corollary 11.4 with source family  $\Phi_{N_{\text{base}}}$ , target family  $\text{SwapClo}(\mathcal{I}_{i+1})$ , and constants  $C_T^{(i+1)}$ ,  $C_R^{(i+1)}$ ,  $D_{i+1}$ . The Fleming–Yoshida lower bound on  $\Phi_{N_{\text{base}}}$  gives

$$\text{SwapSens}(B, \Phi_{N_{\text{base}}}) \geq \Omega(N_{\text{base}}).$$

Applying Corollary 11.4 yields

$$S \geq \frac{\Omega(N_{\text{base}}) - D_{i+1}}{C_T^{(i+1)} C_R^{(i+1)}}.$$

Since  $\varepsilon_{t+1} \geq \varepsilon_t^{1/4}$  for every  $t$ , we have

$$\prod_{t=0}^{i+1} \varepsilon_t^{-O(1)} \leq (1/\varepsilon_0)^{O(\sum_{t \geq 0} 4^{-t})} = (\log n_0)^{O(1)} = o(N_{\text{base}}).$$

Hence the additive term  $D_{i+1}$  is absorbed, and we obtain

$$S \geq \Omega\left(\frac{n_0^\delta}{\prod_{t=0}^{i+1} \varepsilon_t^{-O(1)}}\right).$$

Because  $A$  was arbitrary, this proves

$$\text{SwapSens}_{9\sqrt{\varepsilon_{i+1}}}(\text{SwapClo}(\mathcal{I}_{i+1})) \geq \Omega\left(\frac{n_0^\delta}{\prod_{t=0}^{i+1} \varepsilon_t^{-O(1)}}\right).$$

Enlarging  $c_{\text{sens}}$  absorbs the polynomial losses and establishes item (3) of Lemma 11.13 at stage  $i+1$ . This completes the verification of the stage- $(i+1)$  reduction data.  $\square$

## 11.5 Proof of Theorem 1.6

**Lemma 11.13.** Define a sequence  $(\varepsilon_i)_{i \geq 0}$  by

$$\varepsilon_0 := (\log n_0)^{-c_{\text{init}}} \quad \text{and} \quad \varepsilon_{i+1} := 20 \cdot \varepsilon_i^{1/4} \quad \text{for } i \geq 0.$$

There exists an absolute constant  $c_0 > 0$  such that, whenever  $\varepsilon_i < c_0$ , there is a family  $\mathcal{I}_i$  of left-predicate label cover instances with local description complexity at most  $\varepsilon_i^{-c_{\text{str}}}$  and with the following properties:

1. The local description complexity, the left degree, the right degree, and the right alphabet size are all at most  $\varepsilon_i^{-c_{\text{str}}}$ , the left alphabet size is at most  $\exp(\varepsilon_i^{-c_{\text{str}}})$ , and every left and right vertex has degree at least 1.

2. The numbers of left vertices and right vertices satisfy

$$|U|, |V| \leq n_0 \cdot \prod_{t=0}^i \varepsilon_t^{-c_{\text{size}}}.$$

3. The swap sensitivity of the swap closure satisfies

$$\text{SwapSens}_{9\sqrt{\varepsilon_i}}(\text{SwapClo}(\mathcal{I}_i)) \geq \Omega\left(\frac{n_0^\delta}{\prod_{t=0}^i \varepsilon_t^{-c_{\text{sens}}}}\right).$$

*Proof.* Let  $c_0$  be the constant from Proposition 11.12.

**Base case.** For stage 0, Proposition 11.10 provides the initialized family  $\mathcal{I}_0$  together with the direct initialization map  $\text{Inst}_0$ , the recoveries  $\text{Rec}_{0,\Phi}$ , and constants

$$C_T^{(0)}, C_R^{(0)}, D_0 \leq \varepsilon_0^{-O(1)},$$

satisfying stage-0 reduction data. Its three conclusions are exactly items (1), (2), and (3) at  $i = 0$ .

**Inductive step.** Suppose stage- $i$  reduction data have been constructed on a family  $\mathcal{I}_i$  satisfying items (1)–(3), and that  $\varepsilon_i < c_0$ . Then Proposition 11.12 advances them to stage- $(i + 1)$  reduction data on  $\mathcal{I}_{i+1}$ , with  $\text{Inst}_{i+1}$ , the recoveries  $\text{Rec}_{i+1,\Phi}$ , and constants  $C_T^{(i+1)}, C_R^{(i+1)}, D_{i+1}$ . In particular, items (1), (2), and (3) hold at stage  $i + 1$ . Iterating this update from the initialized family  $\mathcal{I}_0$  proves the lemma.  $\square$

*Proof of Theorem 1.6.* Fix a function  $g : \mathbb{Z}_{>0} \rightarrow (1, \infty)$  with  $g(t) \leq (\log t)^B$ , where  $B > 0$  is the constant from the theorem statement and is chosen so that  $2B < c_{\text{init}}$ . For clarity, write  $n_{\text{LC}}$  for the final left-side size in this proof, and let  $n_{\text{LC}}$  be sufficiently large. From Lemma 11.13 we use only three outputs at stage  $i$ : the threshold  $9\sqrt{\varepsilon_i}$ , the bounds on degrees, alphabet sizes, local description complexity, and vertex counts that are polynomial in  $\varepsilon_i^{-1}$ , and the sensitivity lower bound

$$\Omega\left(\frac{n_0^\delta}{\prod_{t=0}^i \varepsilon_t^{-c_{\text{sens}}}}\right).$$

We therefore choose the last stage whose threshold is at most  $1/g(n_{\text{LC}})$  and then match the resulting size and alphabet bounds to the target parameter  $n_{\text{LC}}$ .

Choose a sufficiently large absolute constant  $K$  and set

$$n_0 := \left\lfloor \frac{n_{\text{LC}}}{(\log n_{\text{LC}})^K} \right\rfloor, \quad \varepsilon_0 := (\log n_0)^{-c_{\text{init}}}.$$

Define

$$\tau(n_{\text{LC}}) := \min\left\{\frac{1}{81g(n_{\text{LC}})^2}, \frac{c_0}{2}\right\}.$$

Then  $\log n_0 = \Theta(\log n_{\text{LC}})$ . Since  $g(n_{\text{LC}}) \leq (\log n_{\text{LC}})^B$  and  $2B < c_{\text{init}}$ , we have

$$\varepsilon_0 = (\log n_0)^{-c_{\text{init}}} \leq \frac{1}{81g(n_{\text{LC}})^2}$$

for all sufficiently large  $n_{\text{LC}}$ . Also  $\varepsilon_0 \rightarrow 0$ , so  $\varepsilon_0 \leq c_0/2$  for all sufficiently large  $n_{\text{LC}}$ . Hence

$$\varepsilon_0 \leq \tau(n_{\text{LC}})$$

once  $n_{\text{LC}}$  is large enough. Define  $(\varepsilon_i)_{i \geq 0}$  as in Lemma 11.13.

Choose the largest integer  $i^* \geq 0$  such that

$$\varepsilon_{i^*} \leq \tau(n_{\text{LC}}).$$

Because  $\varepsilon_0 \leq \tau(n_{\text{LC}})$ , the defining set is non-empty. We therefore have both

$$\varepsilon_{i^*} < c_0$$

and

$$9\sqrt{\varepsilon_{i^*}} \leq 9\sqrt{\tau(n_{\text{LC}})} \leq \frac{1}{g(n_{\text{LC}})}.$$

Maximality and the recursion  $\varepsilon_{i+1} = 20\varepsilon_i^{1/4}$  imply

$$\left(\frac{\tau(n_{\text{LC}})}{20}\right)^4 < \varepsilon_{i^*} \leq \tau(n_{\text{LC}}). \quad (5)$$

Apply Lemma 11.13 with the index  $i^*$  to obtain the family  $\mathcal{I}_{i^*}$  with local description complexity at most  $\varepsilon_{i^*}^{-c_{\text{str}}}$ . Because  $9\sqrt{\varepsilon_{i^*}} \leq 1/g(n_{\text{LC}})$ , every algorithm for  $\text{LabelCover}_{1/g(n_{\text{LC}})}$  on  $\text{SwapClo}(\mathcal{I}_{i^*})$  is also admissible for the threshold in item (3) of Lemma 11.13.

By item (1) of Lemma 11.13 and the lower side of (5),

$$\varepsilon_{i^*}^{-c_{\text{str}}} < \left(\frac{20}{\tau(n_{\text{LC}})}\right)^{4c_{\text{str}}}.$$

If  $\tau(n_{\text{LC}}) = 1/(81g(n_{\text{LC}})^2)$  then the right-hand side is  $g(n_{\text{LC}})^{O(1)}$ , while if  $\tau(n_{\text{LC}}) = c_0/2$  then it is an absolute constant. Hence the left and right degrees and the right alphabet are all bounded by

$$\varepsilon_{i^*}^{-c_{\text{str}}} = g(n_{\text{LC}})^{O(1)},$$

while the left alphabet is bounded by

$$\exp(\varepsilon_{i^*}^{-c_{\text{str}}}) = \exp(g(n_{\text{LC}})^{O(1)}).$$

Item (2) of the lemma gives

$$|U|, |V| \leq n_0 \prod_{t=0}^{i^*} \varepsilon_t^{-c_{\text{size}}}.$$

Since  $\varepsilon_{t+1} = 20\varepsilon_t^{1/4} \geq \varepsilon_t^{1/4}$ , we have  $\varepsilon_t \geq \varepsilon_0^{1/4^t}$ . Therefore, for every absolute constant  $c > 0$ ,

$$\prod_{t=0}^{i^*} \varepsilon_t^{-c} \leq \prod_{t=0}^{i^*} \varepsilon_0^{-c/4^t} = (1/\varepsilon_0)^{c \sum_{t=0}^{i^*} 4^{-t}} \leq (1/\varepsilon_0)^{4c/3} = (\log n_0)^{O(1)}. \quad (6)$$

Hence

$$|U|, |V| \leq n_0 (\log n_0)^{O(1)} \leq n_{\text{LC}}$$

once  $K$  is chosen large enough.

Pad each instance in  $\text{SwapClo}(\mathcal{I}_{i^*})$  by adding  $n_{\text{LC}} - |U|$  isolated left vertices whose left predicates always accept a distinguished symbol. This does not change satisfiability or the value of any assignment on

the original edges, and it preserves swap sensitivity. We therefore obtain a family  $\widehat{\mathcal{I}}_{n_{\text{LC}}}$  with exact left-side size  $|U| = n_{\text{LC}}$ . The right side is unchanged, so

$$|V| \leq n_{\text{LC}}(\log n_{\text{LC}})^{O(1)}.$$

Finally, item (3) of Lemma 11.13 implies

$$\text{SwapSens}_{1/g(n_{\text{LC}})}(\widehat{\mathcal{I}}_{n_{\text{LC}}}) \geq \text{SwapSens}_{9\sqrt{\varepsilon_{i^*}}}(\text{SwapClo}(\mathcal{I}_{i^*})) \geq \Omega\left(\frac{n_0^\delta}{\prod_{t=0}^{i^*} \varepsilon_t^{-c_{\text{sens}}}}\right).$$

Using (6) and  $n_0 = n_{\text{LC}}/(\log n_{\text{LC}})^K$ , we obtain

$$\text{SwapSens}_{1/g(n_{\text{LC}})}(\widehat{\mathcal{I}}_{n_{\text{LC}}}) \geq \Omega\left(\frac{n_{\text{LC}}^\delta}{(\log n_{\text{LC}})^{O(1)}}\right).$$

The right-hand side is at least  $n_{\text{LC}}^{\delta'}$  for some absolute constant  $\delta' \in (0, \delta)$  and all sufficiently large  $n_{\text{LC}}$ . Replacing  $\delta$  by  $\delta'$  proves the theorem.  $\square$

## Part III

# From Label Cover to Covering Problems

## 12 Set Cover

This section proves the two main set-cover sensitivity lower bounds used later in the locality argument, one via the IKW route and one via the Dinur–Harsha route. The former gives a lower bound for logarithmic-approximation algorithms, while the latter yields a polynomial sensitivity lower bound. In both cases, the hard instances have bounded set size and frequency. Both start from the same Feige–Lund–Yannakakis reduction from label cover to set cover, so we first isolate the common transformation and recovery statements.

In the *set cover problem*, an instance is a pair  $(U, \mathcal{F} = (S_1, \dots, S_m))$ , where  $U$  is a finite ground set and  $\mathcal{F}$  is an indexed family of subsets of  $U$ . The goal is to find a subfamily  $\mathcal{S} \subseteq [m]$  of minimum size whose union covers  $U$ , i.e.,  $\bigcup_{i \in \mathcal{S}} S_i = U$ . We call an instance *feasible* if some subfamily covers  $U$ , equivalently if  $\bigcup_{i=1}^m S_i = U$ .

For two instances  $I = (U, \mathcal{F} = (S_1, \dots, S_m))$  and  $\tilde{I} = (U, \tilde{\mathcal{F}} = (\tilde{S}_1, \dots, \tilde{S}_m))$  over the same ground set and indexed family, we define the swap distance by  $\text{SwapDist}(I, \tilde{I}) = \sum_{i=1}^m |S_i \Delta \tilde{S}_i|$ . That is,  $\text{SwapDist}(I, \tilde{I})$  counts the number of *single-element membership changes* across all sets (equivalently, the Hamming distance between the two incidence matrices). We identify a subfamily  $\mathcal{S} \subseteq [m]$  with its indicator vector, and define the Hamming distance by  $\text{Ham}(\mathcal{S}, \tilde{\mathcal{S}}) := |\mathcal{S} \Delta \tilde{\mathcal{S}}|$ . For a (possibly randomized) algorithm  $A$  and an instance  $I$ , define the swap sensitivity by  $\text{SwapSens}(A, I) := \max_{\tilde{I}} \frac{\text{EMD}(A(I), A(\tilde{I}))}{\text{SwapDist}(I, \tilde{I})}$ , where the maximum ranges over instances  $\tilde{I}$  on the same ground set and index set with  $\text{SwapDist}(I, \tilde{I}) > 0$ , and the earth mover's distance uses Ham as the underlying metric on outputs. For a family  $\mathcal{I}$ , set  $\text{SwapSens}(A, \mathcal{I}) := \max_{I \in \mathcal{I}} \text{SwapSens}(A, I)$ . For  $\alpha \geq 1$ , let  $\text{SwapSens}_\alpha(\mathcal{I})$  be the minimum  $\text{SwapSens}(A, \mathcal{I})$  over all algorithms  $A$  defined on all set cover instances and satisfying the  $\alpha$ -approximation guarantee on every feasible  $I \in \mathcal{I}$ .

**Parameter guide.** In the reduction proofs below,  $n_{\text{LC}}$  denotes the size parameter of the source label cover instance,  $N_{\text{SC}}$  the size of the set-cover universe, and  $m_{\text{SC}} := |\mathcal{F}|$  the number of sets. For the IKW route,  $n_{\text{LC}}$  refers to the total number of label-cover vertices; for the Dinur–Harsha route below, we later normalize it to the left-side size. Bare  $n$  is reserved for theorem statements.

## 12.1 Reduction from Label Cover to Set Cover

We follow the basic reduction from label cover to set cover [27] based on the  $(m, l)$ -set system [27, 47], defined below. An  $(m, l)$ -set system consists of a universe  $B$  and collection of subsets  $\mathcal{C} = \{C_1, \dots, C_m\}$  such that, if a collection of at most  $l$  sets in  $\{C_1, \dots, C_m, \bar{C}_1, \dots, \bar{C}_m\}$  covers  $B$ , then the collection must contain both  $C_i$  and  $\bar{C}_i$  for some  $i$ .

**Lemma 12.1** ([47]). An  $(m, l)$ -set system with a universe size  $|B| = O(2^{2l}m^2)$  exists, and can be constructed in  $2^{O(l)}m^{O(1)}$  time.

Fix  $m$  and  $l$ , and let  $(B, \{C_1, \dots, C_m\})$  be an  $(m, l)$ -set system. We define a transformation  $\mathcal{T}_{\text{SC}}$  that maps a label cover instance  $I = (U, V, E, \Sigma_U, \Sigma_V, P, F)$ , where  $P = \{P_u\}_{u \in U}$  and  $|\Sigma_V| = m$ , to a set cover instance  $I' = (E \times B, \mathcal{F})$ . Let

$$\Lambda_I := (V \times \Sigma_V) \sqcup (U \times \Sigma_U).$$

We create the indexed family  $\mathcal{F} = (S_\lambda)_{\lambda \in \Lambda_I}$  of subsets of  $E \times B$ , writing  $S_{v,x}$  for  $S_{(v,x)}$  and  $S_{u,y}$  for  $S_{(u,y)}$ . When the source instance carries left predicates, we encode them directly in the left-label sets below so that inadmissible left labels become useless for covering. After the reduction to set cover, that distinction has been absorbed into the covering instance itself.

- For every vertex  $v \in V$  and label  $x \in \Sigma_V$ , set  $S_{v,x} = \bigcup_{e \ni v} \{e\} \times C_x$ .
- For every vertex  $u \in U$  and label  $y \in \Sigma_U$ , set

$$S_{u,y} = \begin{cases} \bigcup_{e \ni u} \{e\} \times \overline{C_{f_e(y)}} & \text{if } P_u(y) = 1, \\ \emptyset & \text{if } P_u(y) = 0. \end{cases}$$

Repeated subsets with different indices remain distinct members of the indexed family  $\mathcal{F}$ . If  $\Delta(I)$  denotes the maximum degree of the label cover instance, then every set in  $\mathcal{F}$  contains at most  $\Delta(I) \cdot |B|$  elements, and each element of  $E \times B$  is contained in at most  $|\Sigma_U| + |\Sigma_V|$  many indexed sets.

Next, we describe the recovery map  $\mathcal{R}_{\text{SC}, I}$ , which depends on the source instance  $I$ . Given a subfamily  $\mathcal{S} \subseteq \Lambda_I$ , define label sets  $L_u = \{y \in \Sigma_U : (u, y) \in \mathcal{S} \text{ and } P_u(y) = 1\}$  for each  $u \in U$ , and  $L_v = \{x \in \Sigma_V : (v, x) \in \mathcal{S}\}$  for each  $v \in V$ . Intuitively,  $L_u$  (resp.  $L_v$ ) records the labels that the set cover solution assigns to  $u$  (resp.  $v$ ). To obtain a randomized assignment  $\pi$ , we independently pick  $\pi(u)$  uniformly from  $L_u$  and  $\pi(v)$  uniformly from  $L_v$  for every  $u \in U$  and  $v \in V$  (if some  $L_u$  or  $L_v$  is empty we choose an arbitrary fixed label). When the source instance is fixed, we may suppress the subscript  $I$ .

**Lemma 12.2.** Let  $I = (U, V, E, \Sigma_U, \Sigma_V, P, F)$  be a label cover instance, and let  $I' = \mathcal{T}_{\text{SC}}(I)$ . Then the following hold.

- If  $I$  is satisfiable, then  $I'$  has a cover of size at most  $|U| + |V|$ .
- Let  $\mathcal{S} \subseteq \Lambda_I$  be any cover of  $I'$ . After deleting from  $\mathcal{S}$  all selected indices whose associated sets are empty, the cover property is unchanged and the size does not increase. For a fixed edge  $e = (u, v)$ , the nonempty selected sets on the slice  $\{e\} \times B \subset E \times B$  are exactly the sets

$$C_x \quad \text{with } x \in L_v \quad \text{and} \quad \overline{C_{f_e(y)}} \quad \text{with } y \in L_u.$$

- For a fixed source instance  $I$  and any two subfamilies  $\mathcal{S}, \tilde{\mathcal{S}} \subseteq \Lambda_I$ ,  $\text{EMD}(\mathcal{R}_{\text{SC},I}(\mathcal{S}), \mathcal{R}_{\text{SC},I}(\tilde{\mathcal{S}})) \leq \text{Ham}(\mathcal{S}, \tilde{\mathcal{S}})$ .
- If  $I$  and  $\tilde{I}$  differ only by one predicate swap at a single left vertex  $u_0$ , then for any fixed subfamily  $\mathcal{S}$  on the common index set,  $\text{EMD}(\mathcal{R}_{\text{SC},I}(\mathcal{S}), \mathcal{R}_{\text{SC},\tilde{I}}(\mathcal{S})) \leq 1$ . If  $I$  and  $\tilde{I}$  differ only by a projection swap, then  $\mathcal{R}_{\text{SC},I} = \mathcal{R}_{\text{SC},\tilde{I}}$  on the common index set.

*Proof.* If  $I$  is satisfiable, choose a satisfying label for each  $u \in U$  and  $v \in V$ . The corresponding family of  $|U| + |V|$  sets covers every slice  $\{e\} \times B$  because for each edge  $e = (u, v)$  the chosen labels satisfy  $f_e(y) = x$ , so  $C_x \cup \overline{C_{f_e(y)}} = B$ .

Let  $\mathcal{S} \subseteq \Lambda_I$  be any cover of  $I'$ , and delete from it all selected indices whose associated sets are empty to obtain  $\mathcal{S}'$ . The set covered by  $\mathcal{S}'$  is unchanged and  $|\mathcal{S}'| \leq |\mathcal{S}|$ . Fix an edge  $e = (u, v)$ . On the slice  $\{e\} \times B$ , a selected right-label set contributes exactly  $C_x$  for some  $x \in L_v$ , while a selected left-label set contributes either  $\overline{C_{f_e(y)}}$  for some admissible  $y \in L_u$  or nothing at all. After the empty indices are removed, the nonempty selected sets on  $\{e\} \times B$  are therefore exactly the sets

$$C_x \quad \text{with } x \in L_v \quad \text{and} \quad \overline{C_{f_e(y)}} \quad \text{with } y \in L_u.$$

For fixed  $I$ , toggling one selected index changes only one label set  $L_u$  or  $L_v$ . Hence it affects at most one output coordinate. Coupling the uniform choices identically on all unchanged coordinates gives

$$\text{EMD}(\mathcal{R}_{\text{SC},I}(\mathcal{S}), \mathcal{R}_{\text{SC},I}(\tilde{\mathcal{S}})) \leq \text{Ham}(\mathcal{S}, \tilde{\mathcal{S}}).$$

If  $I$  and  $\tilde{I}$  differ only by a predicate swap at  $u_0$ , then the admissibility filter changes only the label set at  $u_0$ : all  $L_v$  are identical and all  $L_u$  with  $u \neq u_0$  are identical. Hence the recovered assignments can differ only at coordinate  $u_0$ , which gives

$$\text{EMD}(\mathcal{R}_{\text{SC},I}(\mathcal{S}), \mathcal{R}_{\text{SC},\tilde{I}}(\mathcal{S})) \leq 1.$$

For a projection swap the predicates  $P_u$  do not change, so the recovery map itself is unchanged.  $\square$

**Lemma 12.3** (Slice soundness for  $\mathcal{T}_{\text{SC}}$ ). Let  $I = (U, V, E, \Sigma_U, \Sigma_V, P, F)$  be a label cover instance, let  $I' = \mathcal{T}_{\text{SC}}(I)$ , and let  $\mathcal{S} \subseteq \Lambda_I$  be a cover of  $I'$ . Fix an edge  $e = (u, v) \in E$ , and write

$$t_e := |L_u| + |L_v|$$

for the recovered label-set sizes associated with  $\mathcal{S}$ . If  $t_e \leq l$ , then for  $\pi \sim \mathcal{R}_{\text{SC},I}(\mathcal{S})$ ,

$$\Pr[\pi \text{ satisfies } e] \geq \frac{4}{l^2}.$$

*Proof.* By the second bullet of Lemma 12.2, after deleting from  $\mathcal{S}$  all selected indices whose associated sets are empty, the nonempty selected sets on the slice  $\{e\} \times B$  are exactly

$$\{C_x : x \in L_v\} \cup \{\overline{C_{f_e(y)}} : y \in L_u\}.$$

These sets still cover  $B$ , and there are at most  $t_e \leq l$  of them. Hence the defining property of the  $(m, l)$ -set system implies that some index  $i \in [m]$  appears on both sides: both  $C_i$  and  $\overline{C_i}$  occur among the selected sets on this slice. Consequently there exist  $x \in L_v$  and  $y \in L_u$  with

$$x = i = f_e(y).$$

Since  $y \in L_u$  also implies  $P_u(y) = 1$ , the edge  $e$  is satisfied whenever  $\pi(v) = x$  and  $\pi(u) = y$ . Under the independent uniform choices in  $\mathcal{R}_{\text{SC},I}$ , this happens with probability at least

$$\frac{1}{|L_u||L_v|} \geq \frac{4}{(|L_u| + |L_v|)^2} \geq \frac{4}{l^2},$$

where the middle inequality is the arithmetic–geometric mean bound.  $\square$

## 12.2 IKW-Based Bound

**Theorem 12.4.** There exist absolute constants  $C_{\text{SC}} > 0$ ,  $c_0 > 0$ , and an integer  $\Gamma_0 \geq 2$  such that the following holds. For all sufficiently large target ground-set sizes  $n$  and every integer  $\Gamma \geq \Gamma_0$  satisfying

$$\log \Gamma \cdot (\log \log \Gamma)^3 \leq c_0 \log n,$$

any algorithm  $A_{\text{SC}}$  for set cover on ground-set size  $n$  that returns a  $C_{\text{SC}} \log \Gamma$ -approximate cover on every feasible instance must have swap sensitivity at least

$$2^{\Omega\left(\frac{\log n}{(\log \log \Gamma)^3}\right)}.$$

Moreover, the lower bound is already witnessed by neighboring instances  $J_0, J_1$  on the same ground set and the same indexed family such that

$$|\mathcal{F}(J_0)| = |\mathcal{F}(J_1)| = O(n),$$

every set in  $J_0$  and  $J_1$  has size at most  $\Gamma$ , every element has frequency at most  $\Gamma$ , and

$$\text{EMD}(A_{\text{SC}}(J_0), A_{\text{SC}}(J_1)) \geq 2^{\Omega\left(\frac{\log n}{(\log \log \Gamma)^3}\right)}.$$

**Lemma 12.5** (IKW-friendly soundness for  $\mathcal{T}_{\text{SC}}$ ). Let  $I = (U, V, E, \Sigma_U, \Sigma_V, \mathcal{P}, F)$  be a left-predicate bi-regular label cover instance satisfying

$$\max\{|U|, |V|\} \leq 2 \min\{|U|, |V|\},$$

and let

$$I' = (E \times B, \mathcal{F}) := \mathcal{T}_{\text{SC}}(I)$$

be built from an  $(m, l)$ -set system with  $m = |\Sigma_V|$ . If  $S$  is a cover of  $I'$  with

$$|S| \leq \frac{l}{8}(|U| + |V|),$$

then for  $\pi \sim \mathcal{R}_{\text{SC},I}(S)$ ,

$$\mathbf{E}[\text{val}_I(\pi)] \geq \frac{2}{l^2}.$$

*Proof.* Delete from  $S$  all selected indices whose associated sets are empty. By Lemma 12.2, the cover property is unchanged and the size does not increase, so we may assume every selected index corresponds to a nonempty set. For each  $u \in U$  and  $v \in V$ , let  $L_u \subseteq \Sigma_U$  and  $L_v \subseteq \Sigma_V$  be the label sets recovered from  $S$ . If we write

$$A := \sum_{u \in U} |L_u|, \quad B := \sum_{v \in V} |L_v|,$$

then  $A + B = |S|$ , and hence

$$A + B \leq \frac{l}{8}(|U| + |V|)$$

For a uniformly random edge  $e = (u, v) \in E$ , bi-regularity implies that  $u$  is uniform in  $U$  and  $v$  is uniform in  $V$ . Writing

$$t_e := |L_u| + |L_v|,$$

and setting

$$m := \min\{|U|, |V|\},$$

we obtain

$$\mathbf{E}[t_e] = \frac{A}{|U|} + \frac{B}{|V|} \leq \frac{A+B}{m} \leq \frac{l}{8} \cdot \frac{|U|+|V|}{m} \leq \frac{3l}{8}.$$

Hence Markov's inequality gives

$$\Pr[t_e \leq l] \geq \frac{5}{8}.$$

For every edge  $e = (u, v)$  with  $t_e \leq l$ , Lemma 12.3 gives

$$\Pr_{\pi \sim \mathcal{R}_{\text{SC}, I(S)}}[\pi \text{ satisfies } e] \geq \frac{4}{l^2}.$$

Averaging over the good  $5/8$  fraction of edges gives

$$\mathbf{E}[\text{val}_I(\pi)] \geq \frac{5}{8} \cdot \frac{4}{l^2} = \frac{5}{2l^2} \geq \frac{2}{l^2}. \quad \square$$

*Proof of Theorem 12.4.* Let  $C_{\text{IKW}} > 0$  be the soundness constant from Theorem 5.1. Fix a sufficiently large target ground-set size  $n$  and an integer  $\Gamma$  satisfying

$$\log \Gamma \cdot (\log \log \Gamma)^3 \leq c_0 \log n.$$

Set

$$N_{\text{SC}} := n, \quad \alpha_{\Gamma} := C_{\text{SC}} \log \Gamma.$$

We choose absolute constants  $C_k$  sufficiently large and  $c_l > 0$  sufficiently small, and then shrink  $C_{\text{SC}}, c_0 > 0$  and enlarge  $\Gamma_0$  finitely many times so that every displayed estimate below holds for all  $\Gamma \geq \Gamma_0$ .

**Step 1: Choose  $k$  and  $l$  from  $\Gamma$ .** Define

$$k := \lceil C_k (\log \log \Gamma)^3 \rceil, \quad l := \lceil c_l \log \Gamma \rceil.$$

For sufficiently large  $\Gamma$ , we have  $c_l \log \Gamma \geq 1$  and hence

$$l \leq 2c_l \log \Gamma.$$

Therefore

$$2 \log l \leq 2 \log(2c_l \log \Gamma) = O(\log \log \Gamma).$$

On the other hand,

$$\sqrt{k} \geq \sqrt{C_k} (\log \log \Gamma)^{3/2}.$$

Since  $(\log \log \Gamma)^{3/2}$  dominates  $\log \log \Gamma$ , choosing  $C_k$  sufficiently large ensures that for all sufficiently large  $\Gamma$ ,

$$C_{\text{IKW}} \sqrt{k} \geq 2 \log l.$$

Equivalently,

$$\exp\left(-C_{\text{IKW}} \sqrt{k}\right) \leq \frac{1}{l^2}.$$

**Step 2: Choose the IKW source size.** By Theorem 5.1, there exist absolute constants  $a, b, \eta > 0$  such that for all sufficiently large  $N_{\text{IKW}}$  and all  $k \leq o(\log N_{\text{IKW}})$  there is a hard family of satisfiable left-predicate biregular label-cover instances with

$$n_{\text{LC}} \in [N_{\text{IKW}}^{ak}, N_{\text{IKW}}^{bk}],$$

sensitivity lower bound at least  $n_{\text{LC}}^{\eta/k}$ , soundness threshold  $\exp(-C_{\text{IKW}}\sqrt{k})$ , and alphabet and degree parameters at most  $2^{O(k)}$ .

Set

$$N_{\text{IKW}} := \left\lfloor \left( \frac{N_{\text{SC}}}{100\Gamma} \right)^{1/(bk)} \right\rfloor.$$

The regime hypothesis implies

$$\frac{\log \Gamma}{\log N_{\text{SC}}} \leq \frac{c_0}{(\log \log \Gamma)^3} = o(1),$$

so after shrinking  $c_0$  and enlarging  $\Gamma_0$  if needed we have

$$\log \left( \frac{N_{\text{SC}}}{100\Gamma} \right) \geq \frac{1}{2} \log N_{\text{SC}}.$$

For all sufficiently large  $N_{\text{SC}}$ , the quantity inside the floor is at least 2, and hence

$$N_{\text{IKW}} \geq \frac{1}{2} \left( \frac{N_{\text{SC}}}{100\Gamma} \right)^{1/(bk)}.$$

Therefore

$$\log N_{\text{IKW}} \geq \frac{1}{bk} \log \left( \frac{N_{\text{SC}}}{100\Gamma} \right) - \log 2 = \Omega \left( \frac{\log N_{\text{SC}}}{k} \right).$$

Moreover,

$$\frac{\log N_{\text{SC}}}{k^2} = \Omega \left( \frac{\log N_{\text{SC}}}{(\log \log \Gamma)^6} \right) = \Omega \left( \frac{\log \Gamma}{(\log \log \Gamma)^3} \right) \rightarrow \infty,$$

where the second bound again uses the regime assumption. Combining the last two displays gives

$$\frac{k}{\log N_{\text{IKW}}} \rightarrow 0.$$

Hence  $k \leq o(\log N_{\text{IKW}})$  for all sufficiently large  $N_{\text{SC}}$ , so Theorem 5.1 applies.

The resulting hard family satisfies

$$n_{\text{LC}} \leq N_{\text{IKW}}^{bk} \leq \frac{N_{\text{SC}}}{100\Gamma}.$$

For the lower bound, the estimate on  $N_{\text{IKW}}$  gives

$$\log n_{\text{LC}} \geq ak \log N_{\text{IKW}} \geq \frac{a}{b} \log \left( \frac{N_{\text{SC}}}{100\Gamma} \right) - ak \log 2.$$

Since  $k = O((\log \log \Gamma)^3) = o(\log N_{\text{SC}})$ , after shrinking  $c_0$  once more we obtain

$$\log n_{\text{LC}} = \Omega(\log N_{\text{SC}}).$$

**Step 3: Balance the two sides.** Let

$$I = (U, V, E, \Sigma_U, \Sigma_V, \mathcal{P}, F)$$

be an instance from this hard family. Since  $I$  is biregular, if  $d_U$  and  $d_V$  denote its common left and right degrees, then

$$|U|d_U = |V|d_V.$$

Every vertex has positive degree, and both  $d_U$  and  $d_V$  are at most  $2^{O(k)}$ , so

$$\frac{|V|}{|U|} = \frac{d_U}{d_V} \leq 2^{O(k)} \quad \text{and} \quad \frac{|U|}{|V|} = \frac{d_V}{d_U} \leq 2^{O(k)}.$$

If  $|U| \leq |V|$ , let

$$r := \left\lceil \frac{|V|}{|U|} \right\rceil$$

and duplicate each left vertex  $u \in U$  into  $r$  copies carrying the same predicate and the same incident projections. If  $|V| < |U|$ , do the symmetric operation on the right side. Let  $I^{\text{bal}}$  denote the resulting instance. Then  $I^{\text{bal}}$  remains left-predicate and biregular, its two sides differ by at most a factor 2, its alphabets are unchanged, and its maximum degree grows by at most a factor  $r \leq 2^{O(k)}$ . Hence there is an absolute quantity

$$D_\Gamma := 2^{O(k)}$$

that simultaneously bounds the maximum degree of  $I^{\text{bal}}$  and the quantity  $|\Sigma_U| + |\Sigma_V|$ . Also,

$$|U(I^{\text{bal}})| + |V(I^{\text{bal}})| = \Theta(n_{\text{LC}}).$$

Given an algorithm on the balanced family, recover an assignment for the original family by choosing one copy uniformly and independently for each duplicated original vertex on the copied side and leaving the untouched side unchanged. This recovery preserves expected value, because each original edge is represented by the same number of copied edges. It is 1-Lipschitz in Hamming distance. A single source swap in the original family becomes at most  $r \leq 2^{O(k)}$  swaps after balancing. Therefore Corollary 11.4 yields a balanced hard family with sensitivity lower bound

$$L_{\text{LC}} \geq \frac{n_{\text{LC}}^{\eta/k} - 1}{2^{O(k)}} = 2^{\Omega(\log n_{\text{LC}}/k)} = 2^{\Omega(\log N_{\text{SC}}/k)}.$$

The last equality uses  $\log n_{\text{LC}} = \Omega(\log N_{\text{SC}})$  and  $k = o(\log n_{\text{LC}})$ .

**Step 4: Build the set-cover core.** We now choose the set-system parameters so that the resulting core instance has maximum set size and maximum element frequency at most  $\Gamma$ . Since

$$\log D_\Gamma = O(k) = O((\log \log \Gamma)^3) = o(\log \Gamma),$$

we have  $D_\Gamma = \Gamma^{o(1)}$ . Let

$$m := |\Sigma_V|.$$

By Lemma 12.1, there is an  $(m, l)$ -set system  $(B_0, \{C_1^{(0)}, \dots, C_m^{(0)}\})$  with  $|B_0| \leq 2^{O(l)} m^{O(1)} = \Gamma^{O(c_l) + o(1)}$ . Choosing  $c_l$  sufficiently small and then enlarging  $\Gamma_0$  if needed, we may assume

$$D_\Gamma |B_0| \leq \frac{\Gamma}{32}.$$

Set

$$q := \left\lceil \frac{\Gamma}{32D_\Gamma|B_0|} \right\rceil, \quad B := [q] \times B_0.$$

For each  $i \in [m]$ , define

$$C_i := [q] \times C_i^{(0)} \subseteq B.$$

This is still an  $(m, l)$ -set system: any collection of at most  $l$  sets from  $\{C_1, \dots, C_m, \bar{C}_1, \dots, \bar{C}_m\}$  that covers  $B$  also covers each copy  $\{j\} \times B_0$ , so already on one copy it must contain both  $C_i^{(0)}$  and  $\bar{C}_i^{(0)}$  for some  $i$ . Since

$$\frac{\Gamma}{32D_\Gamma|B_0|} \geq 1,$$

the choice of  $q$  gives

$$\frac{\Gamma}{32D_\Gamma} \leq |B| \leq \left( \frac{\Gamma}{32D_\Gamma|B_0|} + 1 \right) |B_0| \leq \frac{\Gamma}{16D_\Gamma}.$$

Let

$$I_{\text{core}} = (E \times B, \mathcal{F}_{\text{core}}) := \mathcal{T}_{\text{SC}}(I^{\text{bal}}).$$

By the reduction facts before Lemma 12.2,

$$s_{\text{core}} := \max_{S \in \mathcal{F}_{\text{core}}} |S| \leq \Delta(I^{\text{bal}}) \cdot |B| \leq D_\Gamma |B| \leq \Gamma$$

and

$$f_{\text{core}} := \max_{x \in E \times B} \#\{S \in \mathcal{F}_{\text{core}} : x \in S\} \leq |\Sigma_U| + |\Sigma_V| \leq D_\Gamma \leq \Gamma$$

for all sufficiently large  $\Gamma$ .

Writing

$$N_{\text{core}} := |E| \cdot |B|, \quad m_{\text{core}} := |\mathcal{F}_{\text{core}}|,$$

positivity of the biregular degrees gives

$$|E| \geq \max\{|U|, |V|\} \geq \frac{n_{\text{LC}}}{2},$$

while the maximum-degree bound gives

$$|E| \leq D_\Gamma n_{\text{LC}}.$$

Hence

$$\frac{n_{\text{LC}}\Gamma}{64D_\Gamma} \leq N_{\text{core}} \leq \frac{n_{\text{LC}}\Gamma}{16} \leq \frac{N_{\text{SC}}}{16}.$$

Also,

$$m_{\text{core}} = |V||\Sigma_V| + |U||\Sigma_U| = O(n_{\text{LC}}D_\Gamma) = O(N_{\text{core}}),$$

since  $D_\Gamma = \Gamma^{o(1)}$ .

**Step 5: Exact-size padding.** Let

$$\rho := \left\lceil \frac{N_{\text{SC}}}{N_{\text{core}}} \right\rceil, \quad t := N_{\text{SC}} - \rho N_{\text{core}}.$$

Since  $N_{\text{core}} \leq N_{\text{SC}}/16$ , we have  $\rho \geq 1$ . Form  $I_{\text{rep}}$  as the disjoint union of  $\rho$  indexed copies of  $I_{\text{core}}$ . Partition the  $t$  fresh elements into blocks of size at most  $s_{\text{core}}$  and add one dedicated set for each block. Let

$$r := \left\lceil \frac{t}{s_{\text{core}}} \right\rceil$$

be the number of padding sets, and write the resulting exact-size instance as

$$I_{\text{pad}} = (U_{\text{pad}}, \mathcal{F}_{\text{pad}}).$$

Then

$$|U_{\text{pad}}| = N_{\text{SC}}.$$

Moreover,

$$r \leq \frac{N_{\text{core}}}{s_{\text{core}}} + 1 \leq \frac{|E||B|}{\Delta(I^{\text{bal}})|B|} + 1 \leq \min\{|U|, |V|\} + 1 \leq |U| + |V|.$$

Therefore

$$|\mathcal{F}_{\text{pad}}| = \rho m_{\text{core}} + r = O(N_{\text{SC}}).$$

Every core set still has size at most  $\Gamma$ . Every padding set has size at most  $s_{\text{core}} \leq \Gamma$ , and each padding element has frequency exactly 1. Hence every set in  $I_{\text{pad}}$  has size at most  $\Gamma$ , and every element has frequency at most  $\Gamma$ .

**Step 6: Recovery and approximation lift.** Define an algorithm  $A_{\text{LC}}$  on the balanced IKW family as follows. On input  $I^{\text{bal}}$ , construct  $I_{\text{pad}}$ , run  $A_{\text{SC}}$  on  $I_{\text{pad}}$ , delete the dedicated padding sets, choose a uniformly random core copy, restrict to that copy to obtain a cover  $\mathcal{S}_{\text{core}}$ , and then apply  $\mathcal{R}_{\text{SC}, I^{\text{bal}}}$ .

If  $I^{\text{bal}}$  is satisfiable, then Lemma 12.2 gives

$$\text{opt}(I_{\text{core}}) \leq |U| + |V|.$$

Hence

$$\text{opt}(I_{\text{pad}}) \leq \rho(|U| + |V|) + r.$$

Since  $I_{\text{pad}}$  is feasible and  $A_{\text{SC}}$  satisfies the  $\alpha_{\Gamma}$ -approximation guarantee, the output  $A_{\text{SC}}(I_{\text{pad}})$  is a cover of  $I_{\text{pad}}$  almost surely and

$$\mathbf{E}[|A_{\text{SC}}(I_{\text{pad}})|] \leq \alpha_{\Gamma} \text{opt}(I_{\text{pad}}) \leq \alpha_{\Gamma} (\rho(|U| + |V|) + r).$$

Conditioning on the output of  $A_{\text{SC}}$ , the random core copy satisfies

$$\mathbf{E}[|\mathcal{S}_{\text{core}}| \mid A_{\text{SC}}(I_{\text{pad}})] \leq \frac{|A_{\text{SC}}(I_{\text{pad}})|}{\rho}.$$

Therefore

$$\mathbf{E}[|\mathcal{S}_{\text{core}}|] \leq \frac{\mathbf{E}[|A_{\text{SC}}(I_{\text{pad}})|]}{\rho} \leq \alpha_{\Gamma} \left( |U| + |V| + \frac{r}{\rho} \right) \leq 2\alpha_{\Gamma}(|U| + |V|),$$

since  $\rho \geq 1$  and  $r \leq |U| + |V|$ .

After shrinking  $C_{\text{SC}}$  if necessary, we may assume

$$2\alpha_{\Gamma}(|U| + |V|) \leq \frac{l}{32}(|U| + |V|).$$

Indeed,  $l = \lceil c_l \log \Gamma \rceil$  and  $\alpha_{\Gamma} = C_{\text{SC}} \log \Gamma$ , so it suffices to choose  $C_{\text{SC}}$  small enough compared with  $c_l$ . Markov's inequality then yields

$$\Pr \left[ |\mathcal{S}_{\text{core}}| \leq \frac{l}{8}(|U| + |V|) \right] \geq \frac{3}{4}.$$

On this event, Lemma 12.5 applies because the balanced family has its two sides within a factor 2, and gives

$$\mathbf{E}[\text{val}_{I^{\text{bal}}}(\boldsymbol{\pi}) \mid |\mathcal{S}_{\text{core}}| \leq \frac{l}{8}(|U| + |V|)] \geq \frac{2}{l^2}.$$

Consequently,

$$\mathbf{E}[\text{val}_{I^{\text{bal}}}(\boldsymbol{\pi})] \geq \frac{1}{l^2} \geq \exp\left(-C_{\text{IKW}}\sqrt{k}\right),$$

where the last inequality is exactly the estimate from Step 1. Thus the approximation scale  $C_{\text{SC}} \log \Gamma$  is strong enough for the slice-soundness decoding, and  $A_{\text{LC}}$  is a valid algorithm on the balanced IKW family.

**Step 7: Sensitivity transfer and neighboring witnesses.** Since  $A_{\text{LC}}$  is valid on the balanced hard family, its sensitivity is at least  $L_{\text{LC}}$ . Let  $\mathcal{R}_{\text{SC}, I^{\text{bal}}}^{\oplus}$  denote the repeated-copy recovery that deletes the padding sets, chooses a uniform random copy, and then applies  $\mathcal{R}_{\text{SC}, I^{\text{bal}}}$  on that copy.

For a neighboring balanced source pair  $I^{\text{bal}}, \tilde{I}^{\text{bal}}$ , the corresponding core instances satisfy

$$\text{SwapDist}(I_{\text{core}}, \tilde{I}_{\text{core}}) \leq |\Sigma_U| \Gamma.$$

Indeed, for a projection swap only the sets  $S_{u,y}$  on one  $\{e\} \times B$  slice can change, which contributes at most  $|\Sigma_U| |B| \leq |\Sigma_U| \Gamma$ . For a predicate swap at one left vertex  $u_0$ , each of the  $|\Sigma_U|$  left-label sets can change on at most

$$\deg(u_0) |B| \leq \Delta(I^{\text{bal}}) |B| \leq \Gamma$$

elements, giving the same bound. Repeating the core  $\rho$  times multiplies the swap distance by  $\rho$ , and the padding gadget is fixed, so

$$\text{SwapDist}(I_{\text{pad}}, \tilde{I}_{\text{pad}}) \leq \rho |\Sigma_U| \Gamma.$$

Coupling the random copy choice identically and using Lemma 12.2, we obtain the fixed-source Lipschitz bound

$$\text{EMD}\left(\mathcal{R}_{\text{SC}, I^{\text{bal}}}^{\oplus}(S), \mathcal{R}_{\text{SC}, I^{\text{bal}}}^{\oplus}(\tilde{S})\right) \leq \frac{1}{\rho} \text{Ham}(S, \tilde{S}).$$

The same lemma gives one-swap drift at most 1 for predicate swaps and 0 for projection swaps, so taking  $D = 1$  uniformly, Corollary 11.4 yields

$$\text{SwapSens}(A_{\text{SC}}) \geq \frac{L_{\text{LC}} - 1}{|\Sigma_U| \Gamma}.$$

**Step 8: Final asymptotic simplification.** Since  $|\Sigma_U| \leq D_\Gamma = \Gamma^{o(1)}$ , we have

$$|\Sigma_U|\Gamma = \Gamma^{1+o(1)}.$$

Hence, for some absolute constant  $c_1 > 0$ ,

$$\text{SwapSens}(A_{\text{SC}}) \geq \frac{2^{c_1 \log N_{\text{SC}}/k}}{\Gamma^{1+o(1)}} = 2^{c_1 \log N_{\text{SC}}/k - (1+o(1)) \log \Gamma}.$$

Because

$$k = \Theta((\log \log \Gamma)^3),$$

the regime assumption implies

$$\log \Gamma \leq O(c_0) \cdot \frac{\log N_{\text{SC}}}{k}.$$

After shrinking  $c_0$  if necessary, the subtraction from  $|\Sigma_U|\Gamma$  is absorbed into the main exponent, and therefore

$$\text{SwapSens}(A_{\text{SC}}) \geq 2^{\Omega(\log N_{\text{SC}}/k)} = 2^{\Omega(\log N_{\text{SC}}/(\log \log \Gamma)^3)}.$$

Finally, Lemma 2.3 applied to the exact-size padded family produces neighboring instances  $J_0, J_1$  on the same ground set and the same indexed family such that

$$\text{EMD}(A_{\text{SC}}(J_0), A_{\text{SC}}(J_1)) \geq 2^{\Omega(\log N_{\text{SC}}/(\log \log \Gamma)^3)}.$$

By construction,

$$|\mathcal{F}(J_0)| = |\mathcal{F}(J_1)| = O(N_{\text{SC}}),$$

every set in  $J_0$  and  $J_1$  has size at most  $\Gamma$ , and every element has frequency at most  $\Gamma$ .  $\square$

### 12.3 Dinur–Harsha–Based Bound

**Theorem 12.6.** There exist absolute constants  $B_{\log}^{\text{SC}}, c_{\text{blowup}}^{\text{SC}}, C_{\text{approx}}^{\text{SC}}, \delta_{\text{sens}}^{\text{SC}} > 0$  such that the following holds. For every function  $q = q(n)$  with

$$2 \leq q(n) \leq (\log n)^{B_{\log}^{\text{SC}}}$$

and all sufficiently large  $n$ , any algorithm  $A$  defined on all set cover instances of ground-set size  $n$  and returning a  $C_{\text{approx}}^{\text{SC}} \sqrt{q(n)}$ -approximate cover on every feasible such instance must have sensitivity  $\Omega(n^{\delta_{\text{sens}}^{\text{SC}}})$ . Moreover, the lower bound already holds on instances in which every set has size at most  $\exp\left(q(n)^{c_{\text{blowup}}^{\text{SC}}}\right)$ , every element belongs to at most  $\exp\left(q(n)^{c_{\text{blowup}}^{\text{SC}}}\right)$  sets, and the indexed family size satisfies  $|\mathcal{F}| = O(n)$ .

We prove Theorem 12.6 by starting from the hard family of Theorem 1.6. That family is not stated to be bi-regular, and after trimming isolated left vertices we therefore first balance the remaining instance by the following regularization lemma before applying the common set-cover reduction.

**Lemma 12.7.** Let  $I = (U, V, E, \Sigma_U, \Sigma_V, P, F)$  be a label cover instance with no isolated vertices. For  $u \in U$  and  $v \in V$ , write  $d(u) := \deg_I(u)$  and  $d(v) := \deg_I(v)$ , and let

$$\Delta := \max\left\{\max_{u \in U} d(u), \max_{v \in V} d(v)\right\}, \quad M_\Delta := \text{lcm}(1, \dots, \Delta), \quad K := M_\Delta^2.$$

Define the balanced regularization  $\widehat{I} = \text{Bal}(I)$  by

$$\widehat{U} := \{(u, i) : u \in U, i \in [d(u)]\}, \quad \widehat{V} := \{(v, j) : v \in V, j \in [d(v)]\},$$

and for every edge  $e = (u, v) \in E$ , every  $i \in [d(u)]$ , every  $j \in [d(v)]$ , and every  $t \in [K/(d(u)d(v))]$ , add the copied edge

$$\widehat{e} = (e, i, j, t)$$

between  $(u, i)$  and  $(v, j)$  with projection  $f_{\widehat{e}} := f_e$ . For each copied left vertex  $(u, i) \in \widehat{U}$ , set  $\widehat{P}_{(u, i)} := P_u$ . Then  $|\widehat{U}| = |\widehat{V}| = |E|$ , every left and right vertex of  $\widehat{I}$  has degree exactly  $K$ , and  $|\widehat{E}| = K|E|$ . Moreover, the following hold.

- If  $I$  is satisfiable, then  $\widehat{I}$  is satisfiable.
- For any deterministic assignment  $\widehat{\pi}$  to  $\widehat{I}$ , let  $\text{Proj}_I(\widehat{\pi})$  be the distribution on assignments to  $I$  obtained by independently choosing a uniform copy in  $[d(u)]$  for each  $u \in U$  and a uniform copy in  $[d(v)]$  for each  $v \in V$ , and then reading off the labels of the chosen copies. Then

$$\mathbf{E}_{\pi \sim \text{Proj}_I(\widehat{\pi})} [\text{val}_I(\pi)] = \text{val}_{\widehat{I}}(\widehat{\pi}).$$

By linearity, the same identity holds for randomized assignments  $\widehat{\pi}$ .

- For any two assignments  $\widehat{\pi}, \widetilde{\pi}$  to  $\widehat{I}$ ,

$$\text{EMD}(\text{Proj}_I(\widehat{\pi}), \text{Proj}_I(\widetilde{\pi})) \leq \text{Ham}(\widehat{\pi}, \widetilde{\pi}).$$

- Viewing  $\text{Proj}_I$  as a map on distributions over assignments to  $\widehat{I}$ , for any two such distributions  $\mu, \nu$ ,

$$\text{EMD}(\text{Proj}_I(\mu), \text{Proj}_I(\nu)) \leq \text{EMD}(\mu, \nu).$$

*Proof.* Since every degree lies in  $[\Delta]$ , the number  $K/(d(u)d(v))$  is an integer for every edge  $(u, v) \in E$ . The size identity is immediate from

$$|\widehat{U}| = \sum_{u \in U} d(u) = |E| = \sum_{v \in V} d(v) = |\widehat{V}|.$$

Fix a copied left vertex  $(u, i) \in \widehat{U}$ . For each neighbor  $v \in N(u)$ , the block above the edge  $(u, v)$  contributes  $d(v) \cdot K/(d(u)d(v)) = K/d(u)$  incident copied edges to  $(u, i)$ . Summing over the  $d(u)$  neighbors of  $u$  shows that  $(u, i)$  has degree exactly  $K$ . The same calculation on the right proves that every copied right vertex also has degree  $K$ . Therefore  $|\widehat{E}| = K|\widehat{U}| = K|E|$ . If  $I$  is satisfiable, assign every copy of a vertex the satisfying label of the original vertex, which satisfies every copied edge.

Fix a deterministic assignment  $\widehat{\pi}$  to  $\widehat{I}$ . For an original edge  $e = (u, v)$ , the projection  $\text{Proj}_I$  chooses  $(u, i)$  and  $(v, j)$  uniformly and independently from the corresponding copy fibers, so

$$\Pr_{\pi \sim \text{Proj}_I(\widehat{\pi})} [\pi \text{ satisfies } e] = \frac{1}{d(u)d(v)} \sum_{i=1}^{d(u)} \sum_{j=1}^{d(v)} \mathbf{1}[P_u(\widehat{\pi}(u, i)) = 1 \text{ and } f_e(\widehat{\pi}(u, i)) = \widehat{\pi}(v, j)].$$

Because  $\widehat{P}_{(u, i)} = P_u$ , a copied edge  $\widehat{e} = (e, i, j, t)$  is satisfied by  $\widehat{\pi}$  if and only if

$$P_u(\widehat{\pi}(u, i)) = 1 \quad \text{and} \quad f_e(\widehat{\pi}(u, i)) = \widehat{\pi}(v, j).$$

Every pair  $(i, j)$  appears in exactly  $K/(d(u)d(v))$  copied edges above  $e$ , so the same quantity is the fraction of satisfied copied edges lying above  $e$ . Since each original edge contributes exactly  $K$  copied edges, averaging over  $e \in E$  gives

$$\begin{aligned} \mathbf{E}_{\pi \sim \text{Proj}_I(\hat{\pi})} [\text{val}_I(\pi)] &= \frac{1}{|E|} \sum_{e=(u,v) \in E} \frac{1}{d(u)d(v)} \sum_{i=1}^{d(u)} \sum_{j=1}^{d(v)} \\ &\quad \mathbf{1}[P_u(\hat{\pi}(u, i)) = 1 \text{ and } f_e(\hat{\pi}(u, i)) = \hat{\pi}(v, j)] \\ &= \text{val}_{\hat{I}}(\hat{\pi}). \end{aligned}$$

The randomized case follows by linearity.

For the deterministic Lipschitz claim, couple the random representative choices identically in the two projections. If  $\hat{\pi}$  and  $\tilde{\pi}$  differ on  $k_u$  copies of an original left vertex  $u$ , then the projected labels at  $u$  differ with probability  $k_u/d(u)$ ; similarly, if they differ on  $\ell_v$  copies of an original right vertex  $v$ , then the projected labels at  $v$  differ with probability  $\ell_v/d(v)$ . Hence under this coupling the expected Hamming distance of the projected assignments is

$$\sum_{u \in U} \frac{k_u}{d(u)} + \sum_{v \in V} \frac{\ell_v}{d(v)} \leq \sum_{u \in U} k_u + \sum_{v \in V} \ell_v = \text{Ham}(\hat{\pi}, \tilde{\pi}),$$

which proves the earth mover bound.

For the distributional statement, take an optimal coupling  $(\hat{\pi}, \tilde{\pi})$  of  $\mu$  and  $\nu$ . Conditional on a sampled pair of deterministic assignments, apply the coupling from the previous paragraph to their projections. The resulting pair has marginals  $\text{Proj}_I(\mu)$  and  $\text{Proj}_I(\nu)$ , and its expected Hamming distance is at most

$$\mathbf{E}[\text{Ham}(\hat{\pi}, \tilde{\pi})] = \text{EMD}(\mu, \nu).$$

Taking the infimum over couplings on the projected side proves the claim.  $\square$

**Lemma 12.8** (DH-friendly soundness for  $\mathcal{T}_{\text{SC}}$ ). Let  $I = (U, V, E, \Sigma_U, \Sigma_V, P, F)$  be a bi-regular label cover instance with  $|U| = |V|$ , and let  $I' = \mathcal{T}_{\text{SC}}(I)$ . If  $\mathcal{S} \subseteq \Lambda_I$  is a cover of  $I'$  with

$$|\mathcal{S}| \leq \frac{l}{8}(|U| + |V|),$$

then for  $\pi \sim \mathcal{R}_{\text{SC}, I}(\mathcal{S})$ ,

$$\mathbf{E}[\text{val}_I(\pi)] \geq \frac{2}{l^2}.$$

*Proof.* Let  $\mathcal{S} \subseteq \Lambda_I$  be any cover of  $I'$ , and delete from it all selected indices whose associated sets are empty to obtain  $\mathcal{S}'$ . Write

$$A := \sum_{u \in U} |L_u|, \quad B := \sum_{v \in V} |L_v|.$$

Because  $\mathcal{S}'$  contains no selected index whose associated set is empty, each selected index contributes to exactly one term of  $A$  or  $B$ , and hence

$$A + B = |\mathcal{S}'| \leq |\mathcal{S}| \leq \frac{l}{8}(|U| + |V|) = \frac{l}{4}|U|.$$

For a uniformly random edge  $e = (u, v) \in E$ , bi-regularity implies that  $u$  is uniform in  $U$  and  $v$  is uniform in  $V$ , so if we write  $t_e := |L_u| + |L_v|$ , then

$$\mathbf{E}_e[t_e] = \frac{1}{|U|} \sum_{u \in U} |L_u| + \frac{1}{|V|} \sum_{v \in V} |L_v| = \frac{A+B}{|U|} \leq \frac{l}{4}.$$

Therefore at least a  $3/4$  fraction of the edges satisfy  $t_e \leq l$ . For every edge  $e = (u, v)$  with  $t_e \leq l$ , Lemma 12.3 gives

$$\Pr_{\pi \sim \mathcal{R}_{\text{SC}, I}(\mathcal{S})} [\pi \text{ satisfies } e] \geq \frac{4}{l^2}.$$

Averaging over the good  $3/4$  fraction of edges gives

$$\mathbf{E}_{\pi \sim \mathcal{R}_{\text{SC}, I}(\mathcal{S})} [\text{val}_I(\pi)] \geq \frac{3}{4} \cdot \frac{4}{l^2} = \frac{3}{l^2} \geq \frac{2}{l^2}. \quad \square$$

*Proof of Theorem 12.6.* Let  $B_0, \delta_{\text{LC}} > 0$  be the constants from Theorem 1.6. Choose an absolute constant  $c_{\text{blowup}}^{\text{SC}} \geq 1$  large enough that every factor of the form  $q^{O(1)}$  or  $\exp(q^{O(1)})$  arising below is bounded by  $\exp(q^{c_{\text{blowup}}^{\text{SC}}})$  once  $q \geq 2$ . Next choose

$$B_{\log}^{\text{SC}} \in (0, B_0/4) \quad \text{with} \quad 2B_{\log}^{\text{SC}} c_{\text{blowup}}^{\text{SC}} < 1, \quad \delta_{\text{sens}}^{\text{SC}} := \delta_{\text{LC}}/2,$$

and finally fix  $C_{\text{approx}}^{\text{SC}} \geq 1$  large enough for the approximation-transfer estimates below.

Fix any function  $q(\cdot)$  with  $2 \leq q(t) \leq (\log t)^{B_{\log}^{\text{SC}}}$ , fix a sufficiently large ground-set size  $n$ , and suppose there exists an algorithm  $A'$  defined on all set cover instances of ground-set size  $n$  whose output is a  $C_{\text{approx}}^{\text{SC}} \sqrt{q(n)}$ -approximate cover on every feasible input of that size. Write

$$q_* := q(n), \quad l := \lceil 64C_{\text{approx}}^{\text{SC}} \sqrt{q_*} \rceil, \quad g_* := l^2, \quad \Lambda_* := \left\lceil \exp\left(q_*^{c_{\text{blowup}}^{\text{SC}}}\right) \right\rceil.$$

Choose

$$n_{\text{LC}} := \left\lfloor \frac{n}{8\Lambda_*^2} \right\rfloor.$$

Because  $q_* \leq (\log n)^{B_{\log}^{\text{SC}}}$  and  $2B_{\log}^{\text{SC}} c_{\text{blowup}}^{\text{SC}} < 1$ , we have  $\Lambda_*^2 = n^{o(1)}$ , hence  $n_{\text{LC}} \rightarrow \infty$  and

$$q_* \leq (\log n_{\text{LC}})^{B_0}$$

for all sufficiently large  $n$ . Since  $g_* = O(q_*)$ , Theorem 1.6 therefore applies to the constant function  $g \equiv g_*$  at size  $n_{\text{LC}}$ .

Given a satisfiable left-predicate label cover instance  $I = (U, V, E, \Sigma_U, \Sigma_V, P, F)$  from the hard family supplied by Theorem 1.6 at size  $n_{\text{LC}}$ , fix default labels  $\sigma_U^* \in \Sigma_U$  and  $\sigma_V^* \in \Sigma_V$  once and for all, delete all isolated vertices to obtain the trimmed instance  $I^\circ = (U^\circ, V^\circ, E^\circ, \Sigma_U, \Sigma_V, P^\circ, F^\circ)$ , and form the balanced regularization

$$\widehat{I}_0 := \text{Bal}(I^\circ).$$

The hard family of Theorem 1.6 is nontrivial, so the trimmed instance is nonempty and  $|E^\circ| \geq 1$ .

**Core parameters and the exact-size target instance.** Let

$$\Delta := \max \left\{ \max_{u \in U^\circ} \deg_{I^\circ}(u), \max_{v \in V^\circ} \deg_{I^\circ}(v) \right\}, \quad M_\Delta := \text{lcm}(1, \dots, \Delta), \quad K := M_\Delta^2.$$

Since  $I^\circ$  is a subinstance of  $I$ , Theorem 1.6 gives

$$\Delta = q_*^{O(1)}, \quad |E^\circ| \leq |U| \Delta = n_{\text{LC}} q_*^{O(1)}, \quad |\Sigma_V| = q_*^{O(1)}, \quad |\Sigma_U| \leq \exp(q_*^{O(1)}).$$

Let

$$m_{\text{sys}} := |\Sigma_V|,$$

and let  $(B_{\text{sys}}, \{C_1, \dots, C_{m_{\text{sys}}}\})$  be the  $(m_{\text{sys}}, l)$ -set system from Lemma 12.1. Then

$$K = \exp(q_*^{O(1)}), \quad |B_{\text{sys}}| = 2^{O(l)} m_{\text{sys}}^{O(1)} = \exp(q_*^{O(1)}).$$

Define

$$R_* := \max \left\{ 1, \left\lceil \frac{|\Sigma_U| + |\Sigma_V|}{K |B_{\text{sys}}|} \right\rceil \right\}, \quad K_* := R_* K.$$

Repeat every copied edge of  $\widehat{I}_0$  exactly  $R_*$  times, and denote the resulting bi-regular instance again by  $\widehat{I}$ . All conclusions of Lemma 12.7 remain valid for  $\widehat{I}$ , because every block above each original edge is scaled uniformly. In particular,  $\widehat{I}$  is  $K_*$ -regular,  $|\widehat{U}| = |\widehat{V}| = |E^\circ|$ , and  $|\widehat{E}| = K_* |E^\circ|$ . By the choice of  $c_{\text{blowup}}^{\text{SC}}$ , all of

$$\Delta, \quad K_* |B_{\text{sys}}|, \quad |\Sigma_U| + |\Sigma_V|, \quad C_T := |\Sigma_U| \Delta K_* |B_{\text{sys}}|$$

are at most  $\Lambda_*$ . Therefore the core set cover instance

$$I_{\text{core}} = (\widehat{E} \times B_{\text{sys}}, \mathcal{F}_{\text{core}}) := \mathcal{T}_{\text{SC}}(\widehat{I})$$

has universe size

$$N_{\text{core}} := |\widehat{E}| |B_{\text{sys}}| = K_* |E^\circ| |B_{\text{sys}}| \leq \Lambda_* |E^\circ| \leq \Lambda_*^2 n_{\text{LC}} \leq \frac{n}{8}.$$

Every set in  $I_{\text{core}}$  has size at most  $K_* |B_{\text{sys}}| \leq \Lambda_*$ , every element belongs to at most  $|\Sigma_U| + |\Sigma_V| \leq \Lambda_*$  sets, and

$$|\mathcal{F}_{\text{core}}| = |\widehat{V}| |\Sigma_V| + |\widehat{U}| |\Sigma_U| = |E^\circ| (|\Sigma_U| + |\Sigma_V|) \leq |E^\circ| K_* |B_{\text{sys}}| = N_{\text{core}}.$$

Let

$$\rho := \left\lfloor \frac{n}{N_{\text{core}}} \right\rfloor, \quad t := n - \rho N_{\text{core}}, \quad r := \left\lceil \frac{t}{\Lambda_*} \right\rceil.$$

Form  $I_{\text{rep}}$  as the disjoint union of  $\rho$  indexed copies of  $I_{\text{core}}$ , and then pad it to exact ground-set size  $n$  by adjoining  $t$  fresh elements partitioned into blocks of size at most  $\Lambda_*$ , together with one dedicated set covering each block and no other element. Denote the final instance by  $I_{\text{pad}} = (U_{\text{pad}}, \mathcal{F}_{\text{pad}})$ . Since  $t < N_{\text{core}} \leq \Lambda_* |E^\circ|$ , we have

$$r \leq \frac{N_{\text{core}}}{\Lambda_*} + 1 \leq |E^\circ| + 1 \leq 2|E^\circ|.$$

The disjoint-union part contributes at most  $\rho |\mathcal{F}_{\text{core}}| \leq \rho N_{\text{core}} \leq n$  indexed sets, and the padding gadget adds  $r = O(n/\Lambda_*) = O(n)$  further sets, so  $|\mathcal{F}_{\text{pad}}| = O(n)$ .

**Approximation transfer.** Algorithm  $A$  on input  $I$  performs the exact-size transformation above, runs  $A'$  on  $I_{\text{pad}}$  to obtain a random subfamily  $\mathcal{S}_{\text{pad}}$ , deletes the dedicated padding sets to obtain  $\mathcal{S}_{\text{rep}}$ , chooses a uniform random copy index  $\kappa \in [\rho]$ , restricts to the  $\kappa$ -th copy to obtain  $\mathcal{S}_{\text{core}}$ , applies  $\mathcal{R}_{\text{SC}, \hat{I}}$  to  $\mathcal{S}_{\text{core}}$ , projects via  $\text{Proj}_{I^\circ}$ , and finally inserts the default labels on the deleted isolated vertices.

If  $I$  is satisfiable, then  $I^\circ$  is satisfiable, hence  $\hat{I}_0$  and  $\hat{I}$  are satisfiable by Lemma 12.7. By the completeness part of Lemma 12.2, each copy of  $I_{\text{core}}$  has an optimal cover of size at most

$$|\hat{U}| + |\hat{V}| = 2|E^\circ|.$$

Therefore

$$\text{opt}(I_{\text{pad}}) \leq 2\rho|E^\circ| + r.$$

Since  $I_{\text{pad}}$  is feasible and  $A'$  satisfies the  $C_{\text{approx}}^{\text{SC}}\sqrt{q_*}$ -approximation guarantee, the random subfamily  $\mathcal{S}_{\text{pad}}$  is a cover of  $I_{\text{pad}}$  almost surely and

$$\mathbf{E}[|\mathcal{S}_{\text{pad}}|] \leq C_{\text{approx}}^{\text{SC}}\sqrt{q_*} \text{opt}(I_{\text{pad}}) \leq C_{\text{approx}}^{\text{SC}}\sqrt{q_*} (2\rho|E^\circ| + r).$$

For each copy index  $a \in [\rho]$ , let  $\mathcal{S}^{(a)}$  be the restriction of  $\mathcal{S}_{\text{rep}}$  to the  $a$ -th copy. Because the copies are disjoint, every  $\mathcal{S}^{(a)}$  is a cover of  $I_{\text{core}}$  almost surely. Conditional on  $\mathcal{S}_{\text{pad}}$ , the random index  $\kappa$  is uniform, so

$$\mathbf{E}[|\mathcal{S}_{\text{core}}| \mid \mathcal{S}_{\text{pad}}] = \frac{|\mathcal{S}_{\text{rep}}|}{\rho} \leq \frac{|\mathcal{S}_{\text{pad}}|}{\rho}.$$

Taking expectations gives

$$\mathbf{E}[|\mathcal{S}_{\text{core}}|] \leq \frac{\mathbf{E}[|\mathcal{S}_{\text{pad}}|]}{\rho} \leq C_{\text{approx}}^{\text{SC}}\sqrt{q_*} \left( 2|E^\circ| + \frac{r}{\rho} \right) \leq 4C_{\text{approx}}^{\text{SC}}\sqrt{q_*} |E^\circ|,$$

where the last inequality uses  $r/\rho \leq r \leq 2|E^\circ|$ . Since

$$\frac{l}{8} (|\hat{U}| + |\hat{V}|) = \frac{l}{4} |E^\circ|,$$

Markov's inequality gives

$$\Pr \left[ |\mathcal{S}_{\text{core}}| > \frac{l}{8} (|\hat{U}| + |\hat{V}|) \right] \leq \frac{4C_{\text{approx}}^{\text{SC}}\sqrt{q_*}|E^\circ|}{(l/4)|E^\circ|} \leq \frac{1}{4}.$$

On the complementary event, Lemma 12.8 yields

$$\mathbf{E}_{\hat{\pi} \sim \mathcal{R}_{\text{SC}, \hat{I}}(\mathcal{S}_{\text{core}})} [\text{val}_{\hat{I}}(\hat{\pi}) \mid |\mathcal{S}_{\text{core}}| \leq \frac{l}{8} (|\hat{U}| + |\hat{V}|)] \geq \frac{2}{l^2}.$$

Averaging over the good event therefore gives

$$\mathbf{E}_{\hat{\pi} \sim \mathcal{R}_{\text{SC}, \hat{I}}(\mathcal{S}_{\text{core}})} [\text{val}_{\hat{I}}(\hat{\pi})] \geq \frac{3}{4} \cdot \frac{2}{l^2} \geq \frac{1}{g_*}.$$

Applying the projection identity from Lemma 12.7 and then extending back to the isolated vertices shows that  $A$  is a valid algorithm for  $\text{LabelCover}_{1/g_*}$  on the hard label cover family.

**Sensitivity transfer.** By Theorem 1.6, the induced label cover algorithm  $A$  has swap sensitivity at least  $\Omega(n_{\text{LC}}^{\delta_{\text{LC}}})$ . Fix a neighboring witness pair  $I, \tilde{I}$  from Lemma 2.3. Because a source swap changes only one predicate or one projection and never the underlying graph, the instances  $I^\circ$  and  $\tilde{I}^\circ$  have the same core graph, hence the same parameters  $\Delta, K, R_*, K_*, N_{\text{core}}, \rho$ , and the same padding gadget. For a single core copy, the transformation distance under one source swap is bounded by

$$C_T = |\Sigma_U| \Delta K_* |B_{\text{sys}}| \leq \Lambda_*.$$

Repeating the core  $\rho$  times multiplies this by  $\rho$ , so

$$\text{SwapDist}(I_{\text{pad}}, \tilde{I}_{\text{pad}}) \leq \rho C_T.$$

Let  $\text{Rec}_I^\oplus$  denote the recovery that deletes the dedicated padding sets, chooses a uniform copy, applies  $\mathcal{R}_{\text{SC}, \hat{I}}$  on that copy, projects via  $\text{Proj}_{I^\circ}$ , and then extends by the fixed default labels on the isolated vertices. For any two subfamilies on the common index set of  $I_{\text{pad}}$ , couple the random copy choice identically. The fixed-target Lipschitz bounds from Lemmas 12.2 and 12.7 then give

$$\text{EMD}(\text{Rec}_I^\oplus(\mathcal{S}), \text{Rec}_I^\oplus(\tilde{\mathcal{S}})) \leq \frac{1}{\rho} \text{Ham}(\mathcal{S}, \tilde{\mathcal{S}}),$$

so the repeated-copy recovery has Lipschitz constant  $C_R = 1/\rho$  for the fixed source instance. Under one source swap, the same coupling of the random copy choice together with the predicate-swap / projection-swap analysis for one core copy shows that the recovery map still changes by at most  $D = 1$ . Hence Corollary 11.4 gives

$$\text{SwapSens}(A') \geq \Omega\left(\frac{n_{\text{LC}}^{\delta_{\text{LC}}}}{(\rho C_T)(1/\rho)}\right) = \Omega\left(\frac{n_{\text{LC}}^{\delta_{\text{LC}}}}{C_T}\right) \geq \Omega\left(\frac{n_{\text{LC}}^{\delta_{\text{LC}}}}{\Lambda_*}\right).$$

Because  $n_{\text{LC}} = \Theta(n/\Lambda_*^2)$  and  $\Lambda_* = n^{o(1)}$ , the right-hand side is at least  $\Omega(n^\delta)$  after shrinking  $\delta$  once more if necessary. This proves the claimed sensitivity lower bound at ground-set size  $n$ .  $\square$

## 13 Dominating Set

This section proves the two main dominating-set sensitivity lower bounds: the IKW-based bound Theorem 13.3 and the Dinur–Harsha-based bound Theorem 13.4. We first isolate the common reduction from set cover to dominating set and then specialize it in the two approximation regimes.

**Parameter guide.** In the lower-bound proofs below,  $N_{\text{SC}}$  denotes the size of the source set-cover universe,  $m_{\text{SC}} := |\mathcal{F}|$  the number of sets,  $n_{\text{DS}}$  the target number of vertices, and  $\Gamma$  the parameter bounding both set size and element frequency in the dominating-set reduction. Bare  $n$  is reserved for theorem statements.

### 13.1 Reduction from Set Cover to Dominating Set

Fix a parameter  $\Gamma \geq 1$ . We begin with the classical reduction from set cover to dominating set due to [11]; we refer to it as  $\mathcal{T}_{\text{DS}}$ . Given a set cover instance  $I = (U = \{u_1, \dots, u_{N_{\text{SC}}}\}, \mathcal{F} = (S_1, \dots, S_{m_{\text{SC}}}))$  whose maximum set size and maximum element frequency are at most  $\Gamma$ , form the bipartite incidence graph  $G_0 = (U \cup \mathcal{F}, E_0)$  with  $E_0 = \{(u, S) \mid u \in S\}$ . Introduce an additional set

$$W = \{w_1, \dots, w_{\lceil m_{\text{SC}}/\Gamma \rceil}\}$$

of helper vertices and connect each set vertex  $S_i$  to the helper vertex  $w_{\lceil i/\Gamma \rceil}$ . Then every helper vertex has degree at most  $\Gamma$ , and the choice of helper edges is deterministic once the index set  $[m_{\text{SC}}]$  and the parameter  $\Gamma$  are fixed. Let  $G = \mathcal{T}_{\text{DS}}(I)$  denote the resulting graph. It is bipartite and satisfies

$$\Delta(G) \leq \Gamma + 1, \quad |V(G)| = N_{\text{SC}} + m_{\text{SC}} + \left\lceil \frac{m_{\text{SC}}}{\Gamma} \right\rceil.$$

Let  $\text{sc}(I)$  be the optimum set cover size of  $I$ , and let  $\text{ds}(G)$  be the optimum dominating set size of  $G$ . For approximation guarantees we only apply this reduction to feasible set-cover instances, and the hard instances from Section 12 are feasible. For the later sensitivity pullback, however, it is convenient to keep the recovery map defined on arbitrary neighboring instances as well. For every set-cover instance  $I = (U, \mathcal{F})$  and every element  $u \in U$ , define

$$\sigma_I(u) \in [m_{\text{SC}}] \cup \{\perp\}$$

by a fixed deterministic rule: if  $u$  belongs to at least one set, let  $\sigma_I(u)$  be the smallest index  $i$  with  $u \in S_i$ ; otherwise set  $\sigma_I(u) := \perp$ . When  $I$  is feasible,  $\sigma_I(u) \in [m_{\text{SC}}]$  for every  $u \in U$ . Define the recovery map  $\mathcal{R}_{\text{DS},I}$  on vertex subsets  $D \subseteq V(G)$  by

$$\mathcal{R}_{\text{DS},I}(D) := \{i \in [m_{\text{SC}}] : S_i \in D\} \cup \{\sigma_I(u) : u \in D \cap U, \sigma_I(u) \neq \perp\}.$$

In words, we keep all selected set-vertices, add one fixed incident set for each selected element-vertex when such a set exists, and ignore helper vertices.

**Lemma 13.1.** Let  $I = (U, \mathcal{F})$  be a set-cover instance and let  $G = \mathcal{T}_{\text{DS}}(I)$ . Then the following hold.

- If  $I$  is feasible and  $D$  is a dominating set of  $G$ , then  $\mathcal{R}_{\text{DS},I}(D)$  is a set cover of  $I$ .
- For every vertex subset  $D \subseteq V(G)$ ,  $|\mathcal{R}_{\text{DS},I}(D)| \leq |D|$ .
- Define  $\phi_I(S_i) = i$ ,  $\phi_I(u) = \sigma_I(u)$ , and  $\phi_I(w) = \perp$ . Then  $\mathcal{R}_{\text{DS},I}(D) = \phi_I(D) \setminus \{\perp\}$ . Consequently, for any two vertex subsets  $D, \tilde{D} \subseteq V(G)$ ,

$$\mathcal{R}_{\text{DS},I}(D) \Delta \mathcal{R}_{\text{DS},I}(\tilde{D}) \subseteq \phi_I(D \Delta \tilde{D}),$$

and therefore  $|\mathcal{R}_{\text{DS},I}(D) \Delta \mathcal{R}_{\text{DS},I}(\tilde{D})| \leq |D \Delta \tilde{D}|$ .

- If  $I$  and  $\tilde{I}$  differ by toggling exactly one membership relation  $(u_0, S_j)$ , and the designated choices  $\sigma_I, \sigma_{\tilde{I}}$  are defined by the same deterministic rule, then for any fixed vertex subset  $D$  we have

$$|\mathcal{R}_{\text{DS},I}(D) \Delta \mathcal{R}_{\text{DS},\tilde{I}}(D)| \leq 2.$$

*Proof.* Assume first that  $I$  is feasible, let  $D$  be a dominating set of  $G$ , and fix any element  $u \in U$ . If  $u \in D$ , then the set index  $\sigma_I(u)$  belongs to  $\mathcal{R}_{\text{DS},I}(D)$  and covers  $u$ . If  $u \notin D$ , then because  $D$  dominates  $u$  and only set-vertices are adjacent to  $u$ , some selected set-vertex  $S_i \in D$  contains  $u$ , so  $i \in \mathcal{R}_{\text{DS},I}(D)$  and  $u$  is covered. Thus  $\mathcal{R}_{\text{DS},I}(D)$  is a set cover.

Every selected set-vertex contributes at most one set index, every selected element-vertex contributes at most one set index, and helper vertices contribute nothing. Duplicates only decrease the image size, so  $|\mathcal{R}_{\text{DS},I}(D)| \leq |D|$ .

The identity  $\mathcal{R}_{\text{DS},I}(D) = \phi_I(D) \setminus \{\perp\}$  is immediate from the definition of  $\phi_I$ . Now fix another vertex subset  $\tilde{D} \subseteq V(G)$ . Therefore

$$\mathcal{R}_{\text{DS},I}(D) \Delta \mathcal{R}_{\text{DS},I}(\tilde{D}) \subseteq \phi_I(D \Delta \tilde{D}),$$

which implies

$$|\mathcal{R}_{\text{DS},I}(D) \Delta \mathcal{R}_{\text{DS},I}(\tilde{D})| \leq |D \Delta \tilde{D}|.$$

If  $I$  and  $\tilde{I}$  differ by toggling one membership relation  $(u_0, S_j)$ , then the deterministic rule can change only the designated choice for  $u_0$ ; all other values of  $\phi_I$  and  $\phi_{\tilde{I}}$  agree. Hence, for a fixed  $D$ , the image of  $D$  can change only through the contribution of  $u_0$ , and at worst one set index is inserted, deleted, or replaced by another. Therefore  $|\mathcal{R}_{\text{DS},I}(D) \Delta \mathcal{R}_{\text{DS},\tilde{I}}(D)| \leq 2$ .  $\square$

In particular, for every feasible set-cover instance  $I$ ,  $\text{sc}(I) \leq \text{ds}(G) \leq \text{sc}(I) + |W| \leq \text{sc}(I) + \frac{|\mathcal{F}|}{\Gamma} + 1$ . The lower bound follows by applying  $\mathcal{R}_{\text{DS},I}$  to an optimal dominating set, while the upper bound is obtained from an optimal set cover by selecting the corresponding set-vertices and all helper vertices.

For later exact-size padding, we also fix the following gadget. For integers  $\Delta \geq 2$  and  $t \geq 0$ , let  $\mathcal{P}_{\Delta,t}$  denote any fixed star forest on  $t$  vertices whose component sizes are all at most  $\Delta$  and whose dominating number is exactly  $\lceil t/\Delta \rceil$ ; for example, write  $t = a\Delta + b$  with  $0 \leq b < \Delta$ , take  $a$  copies of  $K_{1,\Delta-1}$ , and if  $b > 0$  add one copy of  $K_{1,b-1}$  (interpreting  $K_{1,0}$  as an isolated vertex).

Whenever a target size  $n_{\text{DS}}$  is fixed and  $n_{\text{DS}} \geq |V(G)|$ , write

$$t_I := n_{\text{DS}} - |V(G)|, \quad H_I := G \sqcup \mathcal{P}_{\Gamma+1,t_I}.$$

Then  $H_I$  has exactly  $n_{\text{DS}}$  vertices and maximum degree at most  $\Gamma + 1$ . Define the padded recovery map on vertex subsets  $D \subseteq V(H_I)$  by

$$\mathcal{R}_{\text{DS},I}^{\text{pad}}(D) := \mathcal{R}_{\text{DS},I}(D \cap V(G)).$$

**Lemma 13.2.** Fix  $\Gamma \geq 1$  and a target size  $n_{\text{DS}}$ . Let  $I = (U, \mathcal{F})$  be a set-cover instance with maximum set size and maximum element frequency at most  $\Gamma$ , let  $G = \mathcal{T}_{\text{DS}}(I)$ , and let  $H_I$  and  $\mathcal{R}_{\text{DS},I}^{\text{pad}}$  be defined as above. Then the following hold.

- If  $I$  is feasible and  $D$  is a dominating set of  $H_I$ , then  $\mathcal{R}_{\text{DS},I}^{\text{pad}}(D)$  is a set cover of  $I$  and

$$|\mathcal{R}_{\text{DS},I}^{\text{pad}}(D)| \leq |D|.$$

- For any two vertex subsets  $D, \tilde{D} \subseteq V(H_I)$ ,

$$|\mathcal{R}_{\text{DS},I}^{\text{pad}}(D) \Delta \mathcal{R}_{\text{DS},I}^{\text{pad}}(\tilde{D})| \leq |D \Delta \tilde{D}|.$$

- If  $I = (U, \mathcal{F})$  and  $\tilde{I} = (U, \tilde{\mathcal{F}})$  differ by toggling exactly one membership relation, and the designated choices  $\sigma_I, \sigma_{\tilde{I}}$  are defined by the same deterministic rule, then for any fixed vertex subset  $D$  on the common vertex set,

$$|\mathcal{R}_{\text{DS},I}^{\text{pad}}(D) \Delta \mathcal{R}_{\text{DS},\tilde{I}}^{\text{pad}}(D)| \leq 2.$$

- If  $I = (U, \mathcal{F})$  and  $\tilde{I} = (U, \tilde{\mathcal{F}})$  differ by toggling exactly one membership relation, then  $H_I$  and  $H_{\tilde{I}}$  live on a common vertex set and differ by exactly one edge.

*Proof.* Write  $P := \mathcal{P}_{\Gamma+1,t_I}$  for the padding forest, so  $H_I = G \sqcup P$ . If  $D$  dominates  $H_I$ , then  $D \cap V(G)$  dominates  $G$  because  $G$  is a connected component of the disjoint union. Hence Lemma 13.1 implies that  $\mathcal{R}_{\text{DS},I}^{\text{pad}}(D)$  is a set cover of  $I$ . Moreover,

$$|\mathcal{R}_{\text{DS},I}^{\text{pad}}(D)| = |\mathcal{R}_{\text{DS},I}(D \cap V(G))| \leq |D \cap V(G)| \leq |D|.$$

For any two vertex subsets  $D, \tilde{D} \subseteq V(H_I)$ , deleting the padding component cannot increase symmetric difference, so

$$|(D \cap V(G)) \Delta (\tilde{D} \cap V(G))| \leq |D \Delta \tilde{D}|.$$

Applying Lemma 13.1 to the core graph  $G$  gives

$$|\mathcal{R}_{\text{DS},I}^{\text{pad}}(D) \Delta \mathcal{R}_{\text{DS},\tilde{I}}^{\text{pad}}(\tilde{D})| \leq |(D \cap V(G)) \Delta (\tilde{D} \cap V(G))| \leq |D \Delta \tilde{D}|.$$

If  $I$  and  $\tilde{I}$  differ by toggling one membership relation on the same ground set and indexed family, then the two core graphs share the same vertex set  $U \cup \mathcal{F} \cup W$ . Since the indexed family size is unchanged, the same padding forest is attached in both padded graphs. Hence for any fixed  $D$  on the common vertex set,

$$|\mathcal{R}_{\text{DS},I}^{\text{pad}}(D) \Delta \mathcal{R}_{\text{DS},\tilde{I}}^{\text{pad}}(D)| = |\mathcal{R}_{\text{DS},I}(D \cap V(G)) \Delta \mathcal{R}_{\text{DS},\tilde{I}}(D \cap V(G))| \leq 2$$

by Lemma 13.1. Under the same hypothesis, the core graphs differ by exactly one incidence edge and the padding forest is identical, so  $H_I$  and  $H_{\tilde{I}}$  differ by exactly one edge.  $\square$

### 13.2 IKW-Based $O(\log \Delta)$ Bound

**Theorem 13.3.** There exist an absolute constant  $C > 0$ , an integer  $\Delta_0 \geq 2$ , and an absolute constant  $c_0 > 0$  such that for all sufficiently large  $n$  and every integer  $\Delta \geq \Delta_0$  satisfying

$$\log \Delta \cdot (\log \log \Delta)^3 \leq c_0 \log n,$$

any algorithm for the dominating set problem on graphs with  $n$  vertices and maximum degree at most  $\Delta$  with approximation ratio

$$C \log \Delta$$

requires sensitivity at least

$$2^{\Omega\left(\frac{\log n}{(\log \log \Delta)^3}\right)}.$$

Set  $\Gamma := \Delta - 1$ . Apply Theorem 12.4 with this value of  $\Gamma$ , and then compose it with the common set-cover  $\rightarrow$  dominating-set reduction above. This is exactly the IKW label-cover  $\rightarrow$  set-cover  $\rightarrow$  dominating-set route, where  $\Gamma = \Delta - 1$  is the bound on maximum set size and element frequency in the set-cover step. Combined with the sensitivity-to-locality transfer theorem, it also yields Theorem 1.1. For larger  $\Delta$ , the locality lower bound in Theorem 1.1 is trivial because its right-hand side is already below 1.

*Proof of Theorem 13.3.* Let  $C_{\text{SC}}, c_0^{\text{SC}} > 0$  and  $\Gamma_0 \geq 2$  be the constants from Theorem 12.4. Fix a sufficiently large integer  $n$  and an integer  $\Delta \geq \Delta_0$ , where  $\Delta_0 := \Gamma_0 + 1$ , satisfying

$$\log \Delta \cdot (\log \log \Delta)^3 \leq c_0 \log n$$

for a sufficiently small absolute constant  $c_0 \leq c_0^{\text{SC}}/2$ . Set

$$\Gamma := \Delta - 1.$$

Assume there exists an absolute constant  $C_{\text{approx}} > 0$  and an algorithm  $A'$  for dominating set on  $n$ -vertex graphs of maximum degree at most  $\Delta$  such that

$$r(\Delta) := C_{\text{approx}} \log \Delta.$$

Moreover, for every such graph  $G$ , the output  $A'(G)$  is a dominating set almost surely and

$$\mathbf{E}[|A'(G)|] \leq r(\Delta) \text{ds}(G).$$

**Step 1: The bounded set-cover family on which we run the reduction.** Fix an absolute constant  $C_{|\mathcal{F}|} \geq 1$  that witnesses the  $|\mathcal{F}| = O(N_{\text{SC}})$  bound from Theorem 12.4, and define

$$C_{\text{pad}} := 8(C_{|\mathcal{F}|} + 1), \quad C_{\text{lift}}^{\text{DS}} := C_{|\mathcal{F}|} + 2C_{\text{pad}} + 3, \quad N_{\text{SC}} := \left\lfloor \frac{n}{C_{\text{pad}}} \right\rfloor.$$

Since  $N_{\text{SC}} = \Theta(n)$ , after decreasing  $c_0$  further if necessary we also have

$$\log \Gamma \cdot (\log \log \Gamma)^3 \leq c_0^{\text{SC}} \log N_{\text{SC}}.$$

Let  $\mathcal{Y}_{N_{\text{SC}}, \Gamma}^{\text{SC}}$  denote the family of set-cover instances

$$J = (U, \mathcal{F})$$

satisfying

$$|U| = N_{\text{SC}}, \quad |\mathcal{F}| \leq C_{|\mathcal{F}|} N_{\text{SC}}, \quad \text{maximum set size} \leq \Gamma, \quad \text{maximum frequency} \leq \Gamma.$$

**Step 2: Exact graph size and degree.** Fix the target size  $n_{\text{DS}} := n$ . For any instance  $J = (U, \mathcal{F}) \in \mathcal{Y}_{N_{\text{SC}}, \Gamma}^{\text{SC}}$ , write  $m_{\text{SC}} := |\mathcal{F}|$  and form the core graph

$$G_J := \mathcal{T}_{\text{DS}}(J)$$

with parameter  $\Gamma$ . The common reduction facts above give

$$\Delta(G_J) \leq \Gamma + 1 = \Delta, \quad n_{\text{core}} := |V(G_J)| = N_{\text{SC}} + m_{\text{SC}} + \left\lceil \frac{m_{\text{SC}}}{\Gamma} \right\rceil.$$

Since  $m_{\text{SC}} \leq C_{|\mathcal{F}|} N_{\text{SC}}$ , we get

$$n_{\text{core}} \leq (1 + 2C_{|\mathcal{F}|}) N_{\text{SC}} + 1.$$

By the choice of  $C_{\text{pad}}$  and  $N_{\text{SC}} = \lfloor n/C_{\text{pad}} \rfloor$ , this implies

$$n_{\text{core}} \leq \frac{n}{2}$$

for all sufficiently large  $n$ . Let

$$t := n_{\text{DS}} - n_{\text{core}}$$

and let

$$H_J := G_J \sqcup \mathcal{P}_{\Delta, t}.$$

Then every such  $H_J$  has exactly  $n_{\text{DS}} = n$  vertices and maximum degree at most  $\Delta$ .

**Step 3: Approximation transfer.** We define a set-cover algorithm  $A$  on all instances of ground-set size  $N_{\text{SC}}$  as follows. If

$$J \in \mathcal{Y}_{N_{\text{SC}}, \Gamma}^{\text{SC}},$$

then construct  $H_J$ , run  $A'$  on  $H_J$ , and apply the padded recovery map  $\mathcal{R}_{\text{DS}, J}^{\text{pad}}$  to the output. If  $J \notin \mathcal{Y}_{N_{\text{SC}}, \Gamma}^{\text{SC}}$  and is feasible, let  $A(J)$  be an optimal set cover; if  $J$  is infeasible, define  $A(J)$  arbitrarily.

We now verify that  $A$  is a valid approximation algorithm. Let  $J$  be a feasible input of ground-set size  $N_{\text{SC}}$ . If  $J \notin \mathcal{Y}_{N_{\text{SC}}, \Gamma}^{\text{SC}}$ , then  $A(J)$  is optimal by definition. Assume henceforth that

$$J \in \mathcal{Y}_{N_{\text{SC}}, \Gamma}^{\text{SC}}.$$

Since every set in  $J$  has size at most  $\Gamma$ ,

$$\text{sc}(J) \geq \frac{N_{\text{SC}}}{\Gamma}.$$

Therefore

$$\frac{m_{\text{SC}}}{\Gamma} + 1 \leq \frac{C_{|\mathcal{F}|} N_{\text{SC}}}{\Gamma} + 1 \leq (C_{|\mathcal{F}|} + 1) \text{sc}(J).$$

Also,

$$\text{ds}(\mathcal{P}_{\Delta, t}) = \left\lceil \frac{t}{\Delta} \right\rceil \leq \frac{n}{\Delta} + 1.$$

Since  $n \leq 2C_{\text{pad}} N_{\text{SC}}$  for all sufficiently large  $n$  and  $\Delta = \Gamma + 1$ , this gives

$$\text{ds}(\mathcal{P}_{\Delta, t}) \leq \frac{2C_{\text{pad}} N_{\text{SC}}}{\Gamma + 1} + 1 \leq (2C_{\text{pad}} + 1) \text{sc}(J).$$

By Lemma 13.1,

$$\text{ds}(G_J) \leq \text{sc}(J) + \frac{m_{\text{SC}}}{\Gamma} + 1.$$

Combining the last three displays yields

$$\text{ds}(H_J) = \text{ds}(G_J) + \text{ds}(\mathcal{P}_{\Delta, t}) \leq C_{\text{lift}}^{\text{DS}} \text{sc}(J).$$

Since  $A'(H_J)$  is a dominating set of  $H_J$  almost surely, Lemma 13.2 implies that  $A(J)$  is a set cover of  $J$  almost surely and

$$\mathbf{E}[|A(J)|] \leq \mathbf{E}[|A'(H_J)|] \leq r(\Delta) \text{ds}(H_J) \leq C_{\text{lift}}^{\text{DS}} r(\Delta) \text{sc}(J).$$

Choosing  $C_{\text{approx}}$  sufficiently small (as an absolute constant depending only on  $C_{\text{SC}}$  and  $C_{\text{lift}}^{\text{DS}}$ ) ensures

$$C_{\text{lift}}^{\text{DS}} r(\Delta) \leq C_{\text{SC}} \log \Gamma$$

for all sufficiently large  $n$ . Therefore  $A$  is a valid set-cover algorithm on ground-set size  $N_{\text{SC}}$ .

Applying Theorem 12.4 to  $A$ , we obtain neighboring instances

$$J_0, J_1$$

of ground-set size  $N_{\text{SC}}$  such that

$$|\mathcal{F}(J_0)| = |\mathcal{F}(J_1)| = O(N_{\text{SC}}),$$

every set in  $J_0$  and  $J_1$  has size at most  $\Gamma$ , every element has frequency at most  $\Gamma$ , and

$$\text{EMD}(A(J_0), A(J_1)) \geq L := 2^{\Omega(\log N_{\text{SC}} / (\log \log \Gamma)^3)}.$$

Hence, for all sufficiently large  $n$ , we have

$$J_0, J_1 \in \mathcal{Y}_{N_{\text{SC}}, \Gamma}^{\text{SC}}.$$

**Step 4: Sensitivity transfer back to dominating set.** Because the witness path keeps the ground set and indexed family fixed, the neighboring instances  $J_0$  and  $J_1$  satisfy the hypotheses of Lemma 13.2. That lemma yields transformation constant  $C_T = 1$ , fixed-source Lipschitz constant  $C_R = 1$ , and one-toggle drift at most  $D = 2$  for the padded reduction. Hence Corollary 11.4 gives

$$\text{SwapSens}(A') \geq L - 2.$$

Finally, since  $N_{\text{SC}} = \Theta(n)$ ,

$$L = 2^{\Omega(\log n / (\log \log \Delta)^3)}.$$

Because every padded graph  $H_J$  has exactly  $n$  vertices and maximum degree at most  $\Delta$ , this proves Theorem 13.3.  $\square$

### 13.3 Dinur–Harsha–Based Polylogarithmic-in-Degree Bound

**Theorem 13.4.** There exist  $\beta, \delta, \kappa \in (0, 1)$ , an absolute constant  $C > 0$ , and an integer  $\Delta_0 \geq 2$  such that for all sufficiently large  $n$  and every integer  $\Delta$  satisfying

$$\Delta_0 \leq \Delta \leq \exp(C(\log n)^\kappa),$$

any  $O((\log \Delta)^\beta)$ -approximation algorithm for the dominating set problem on  $n$ -vertex graphs with maximum degree at most  $\Delta$  requires sensitivity  $\Omega(n^\delta)$ .

Combined with the sensitivity-to-locality transfer theorem [28], Theorem 13.4 immediately implies Theorem 1.2, with the only wrinkle being that the upper bound  $\Delta \leq \exp(C(\log n)^\kappa)$  can be removed at the cost of replacing the constant  $\beta$  in the approximation factor  $\log^\beta(\Delta)$  with the worse constant  $\beta\kappa$ .

To prove Theorem 13.4, we start from the exact-size set-cover hard family supplied by Theorem 12.6, choose  $q_\Delta$  from the target degree  $\Delta$ , set  $\Gamma := \Delta - 1$ , and then apply the common reduction and exact-size padding above. The parameter choice is made so that the lifted dominating-set approximation matches the  $O(\sqrt{q})$  threshold in Theorem 12.6.

*Proof of Theorem 13.4.* Let  $B_{\log}^{\text{SC}}, c_{\text{blowup}}^{\text{SC}}, C_{\text{approx}}^{\text{SC}}, \delta_{\text{sens}}^{\text{SC}} > 0$  be the constants from Theorem 12.6. Since Theorem 12.6 remains valid after decreasing the exponent  $B_{\log}^{\text{SC}}$ , we may assume

$$0 < B_{\log}^{\text{SC}} < 1.$$

Set

$$\delta := \min \left\{ \delta_{\text{sens}}^{\text{SC}}, \frac{1}{2} \right\},$$

and choose

$$\beta \in \left( 0, \min \left\{ \frac{1}{2}, \frac{1}{4c_{\text{blowup}}^{\text{SC}}}, \frac{B_{\log}^{\text{SC}}}{4} \right\} \right), \quad \kappa := \frac{B_{\log}^{\text{SC}}}{2},$$

so in particular  $\beta, \delta, \kappa \in (0, 1)$ . Let  $C_{\text{approx}} > 0$  absorb the hidden constant in the assumed  $O((\log \Delta)^\beta)$ -approximation guarantee for dominating set. Let  $C_{|\mathcal{F}|} \geq 1$  be an absolute constant witnessing the linear-size bound  $|\mathcal{F}| = O(N_{\text{SC}})$  in Theorem 12.6, and define

$$C_{\text{pad}} := 8(C_{|\mathcal{F}|} + 1), \quad C_{\text{lift}}^{\text{DS}} := C_{|\mathcal{F}|} + 2C_{\text{pad}} + 3.$$

Fix a sufficiently large  $n$  and an integer  $\Delta$  with

$$\Delta_0 \leq \Delta \leq \exp(C(\log n)^\kappa),$$

where  $\Delta_0$  is chosen large enough, and  $C > 0$  small enough, so that the estimates below all hold. Set

$$\Gamma := \Delta - 1, \quad q_\Delta := \max \left\{ 2, \left\lceil \left( \frac{64C_{\text{approx}}C_{\text{lift}}^{\text{DS}}}{C_{\text{approx}}^{\text{SC}}} \right)^2 (\log \Delta)^{2\beta} \right\rceil \right\}, \quad M_\Delta^{\text{SC}} := \left\lceil \exp \left( q_\Delta^{c_{\text{blowup}}^{\text{SC}}} \right) \right\rceil.$$

Because  $2\beta c_{\text{blowup}}^{\text{SC}} < 1$ , enlarging  $\Delta_0$  if necessary ensures

$$M_\Delta^{\text{SC}} + 1 \leq \Gamma.$$

The upper bound  $\Delta \leq \exp(C(\log n)^\kappa)$ , the choice  $\kappa = B_{\log}^{\text{SC}}/2$ , and the inequality  $\beta < 1/2$  imply

$$q_\Delta \leq (\log n)^{B_{\log}^{\text{SC}}/2}$$

after possibly decreasing  $C$  and increasing  $\Delta_0$ , because  $2\beta\kappa = \beta B_{\log}^{\text{SC}} < B_{\log}^{\text{SC}}/2$ . Therefore, once we set

$$N_{\text{SC}} := \left\lfloor \frac{n}{C_{\text{pad}}} \right\rfloor,$$

we also have

$$q_\Delta \leq (\log N_{\text{SC}})^{B_{\log}^{\text{SC}}}$$

after increasing the lower threshold on  $n$  if necessary.

Apply Theorem 12.6 to the constant function  $q \equiv q_\Delta$  at ground-set size  $N_{\text{SC}}$ . Unpacking the proof of that theorem, we may choose the resulting family of hard instances  $\mathcal{H}_{N_{\text{SC}}}$  so that all instances have universe size  $N_{\text{SC}}$ , share the same indexed family size

$$m_{\text{SC}} := |\mathcal{F}| \leq C_{|\mathcal{F}|} N_{\text{SC}},$$

differ only in their membership relations, every set has size at most  $M_\Delta^{\text{SC}}$ , every element belongs to at most  $M_\Delta^{\text{SC}}$  sets, and every  $C_{\text{approx}}^{\text{SC}} \sqrt{q_\Delta}$ -approximation algorithm on those instances has sensitivity  $\Omega(N_{\text{SC}}^\delta)$ .

Let

$$\partial\mathcal{H}_{N_{\text{SC}}} := \{\tilde{I} : \exists I \in \mathcal{H}_{N_{\text{SC}}} \text{ with } \text{SwapDist}(I, \tilde{I}) = 1\}$$

be the distance-1 neighborhood of this family under one membership toggle. Since a single membership toggle changes the size of at most one set and the frequency of at most one element by at most 1, and since  $M_\Delta^{\text{SC}} + 1 \leq \Gamma$ , every instance in  $\mathcal{H}_{N_{\text{SC}}} \cup \partial\mathcal{H}_{N_{\text{SC}}}$  still has maximum set size and maximum frequency at most  $\Gamma$ .

Fix the target size  $n_{\text{DS}} := n$ . For each  $I \in \mathcal{H}_{N_{\text{SC}}} \cup \partial\mathcal{H}_{N_{\text{SC}}}$ , let

$$G_I := \mathcal{T}_{\text{DS}}(I)$$

with parameter  $\Gamma$ . The common reduction facts above give

$$\Delta_{\text{fin}} := \Gamma + 1 = \Delta, \quad n_{\text{core}} := |V(G_I)| = N_{\text{SC}} + m_{\text{SC}} + \left\lceil \frac{m_{\text{SC}}}{\Gamma} \right\rceil \leq (1 + 2C_{|\mathcal{F}|})N_{\text{SC}} + 1.$$

Since  $N_{\text{SC}} = \lfloor n/C_{\text{pad}} \rfloor$  and  $C_{\text{pad}} = 8(C_{|\mathcal{F}|} + 1)$ , we may assume  $n_{\text{core}} \leq n/2$  for every such  $I$  after increasing the lower threshold on  $n$  if necessary. We may also assume

$$n \leq 2C_{\text{pad}}N_{\text{SC}}.$$

Let

$$t := n_{\text{DS}} - n_{\text{core}}$$

and

$$H_I := G_I \sqcup \mathcal{P}_{\Delta,t}.$$

All padded graphs  $H_I$  therefore have exactly  $n_{\text{DS}} = n$  vertices and maximum degree at most  $\Delta$ .

Suppose there exists an algorithm  $A'$  that, on every  $n$ -vertex graph of maximum degree at most  $\Delta$ , outputs a dominating set almost surely and satisfies

$$\mathbf{E}[|A'(G)|] \leq C_{\text{approx}}(\log \Delta)^\beta \text{ds}(G)$$

for every such graph  $G$ . We convert  $A'$  into a set-cover algorithm  $A$  on all ground sets of size  $N_{\text{SC}}$  as follows. If  $J \in \mathcal{H}_{N_{\text{SC}}} \cup \partial\mathcal{H}_{N_{\text{SC}}}$ , construct  $H_J$ , run  $A'$  on  $H_J$ , and then apply the padded recovery map  $\mathcal{R}_{\text{DS},J}^{\text{pad}}$ . If  $J$  lies outside  $\mathcal{H}_{N_{\text{SC}}} \cup \partial\mathcal{H}_{N_{\text{SC}}}$  and is feasible, let  $A(J)$  be an optimal set cover; if  $J$  is infeasible, define  $A(J)$  arbitrarily.

We now verify that  $A$  is a  $C_{\text{approx}}^{\text{SC}}\sqrt{q_\Delta}$ -approximation algorithm on every feasible input of ground-set size  $N_{\text{SC}}$ . Let  $J$  be such an input. If  $J \notin \mathcal{H}_{N_{\text{SC}}} \cup \partial\mathcal{H}_{N_{\text{SC}}}$ , then  $A(J)$  is optimal by definition.

Assume henceforth that  $J \in \mathcal{H}_{N_{\text{SC}}} \cup \partial\mathcal{H}_{N_{\text{SC}}}$ . Every set of  $J$  has size at most  $\Gamma$ , so

$$\text{sc}(J) \geq \frac{N_{\text{SC}}}{\Gamma}.$$

Since  $m_{\text{SC}} \leq C_{|\mathcal{F}|}N_{\text{SC}}$  and  $\text{sc}(J) \geq 1$ , it follows that

$$\frac{m_{\text{SC}}}{\Gamma} + 1 \leq C_{|\mathcal{F}|} \frac{N_{\text{SC}}}{\Gamma} + 1 \leq (C_{|\mathcal{F}|} + 1)\text{sc}(J).$$

Moreover,

$$\text{ds}(\mathcal{P}_{\Delta,t}) = \left\lceil \frac{t}{\Delta} \right\rceil \leq \frac{n}{\Delta} + 1 \leq \frac{2C_{\text{pad}}N_{\text{SC}}}{\Gamma + 1} + 1 \leq (2C_{\text{pad}} + 1)\text{sc}(J),$$

because  $\Delta = \Gamma + 1$  and  $n \leq 2C_{\text{pad}}N_{\text{SC}}$ . Combining these estimates with

$$\text{ds}(G_J) \leq \text{sc}(J) + \frac{m_{\text{SC}}}{\Gamma} + 1$$

gives

$$\text{ds}(H_J) = \text{ds}(G_J) + \text{ds}(\mathcal{P}_{\Delta,t}) \leq C_{\text{lift}}^{\text{DS}} \text{sc}(J).$$

Since  $A'(H_J)$  is a dominating set of  $H_J$  almost surely, Lemma 13.2 implies that  $A(J)$  is a set cover of  $J$  almost surely and

$$\mathbf{E}[|A(J)|] \leq \mathbf{E}[|A'(H_J)|] \leq C_{\text{approx}}(\log \Delta)^\beta \text{ds}(H_J) \leq C_{\text{approx}}C_{\text{lift}}^{\text{DS}}(\log \Delta)^\beta \cdot \text{sc}(J).$$

By the definition of  $q_\Delta$ ,

$$C_{\text{approx}}C_{\text{lift}}^{\text{DS}}(\log \Delta)^\beta \leq C_{\text{approx}}^{\text{SC}}\sqrt{q_\Delta}.$$

Thus  $A$  is a  $C_{\text{approx}}^{\text{SC}}\sqrt{q\Delta}$ -approximation algorithm on every feasible set-cover instance of ground-set size  $N_{\text{SC}}$ .

Applying Theorem 12.6 to  $A$ , we obtain a feasible instance  $I \in \mathcal{H}_{N_{\text{SC}}}$  with

$$\text{SwapSens}(A, I) \geq \Omega(N_{\text{SC}}^\delta).$$

By Lemma 2.3, there exists a neighboring instance  $\tilde{I} \in \partial\mathcal{H}_{N_{\text{SC}}}$ , obtained from  $I$  by toggling one membership relation, such that

$$\text{EMD}(A(I), A(\tilde{I})) \geq \Omega(N_{\text{SC}}^\delta).$$

Since  $I$  and  $\tilde{I}$  live on the same ground set and indexed family, Lemma 13.2 applies to their padded graphs. It yields transformation constant  $C_T = 1$ , fixed-source Lipschitz constant  $C_R = 1$ , and one-toggle drift at most  $D = 2$ . Therefore Corollary 11.4 yields

$$\text{SwapSens}(A') \geq \Omega(N_{\text{SC}}^\delta)$$

on  $n$ -vertex graphs of maximum degree at most  $\Delta$ . Since  $N_{\text{SC}} = \Theta(n)$ , this implies

$$\text{SwapSens}(A') \geq \Omega(n^\delta),$$

which proves the theorem. □

## References

- [1] A. Aamand, M. Aliakbarpour, J. Y. Chen, S. Narayanan, and S. Silwal. On the structure of replicable hypothesis testers. In *Proceedings of the 2026 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 5771–5823, 2026.
- [2] A. Akbari, X. Coiteux-Roy, F. d’Amore, F. Le Gall, H. Lievonon, D. Melnyk, A. Modanese, S. Pai, M.-O. Renou, V. Rozhoň, et al. Online locality meets distributed quantum computing. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing (STOC)*, pages 1295–1306, 2025.
- [3] N. Alon. Explicit expanders of every degree and size. *Combinatorica*, 41(4):447–463, 2021.
- [4] H. Arfaoui and P. Fraigniaud. What can be computed without communications? *ACM SIGACT News*, 45(3):82–104, 2014.
- [5] M. Bafna, N. Lifshitz, and D. Minzer. Constant degree direct product testers with small soundness. In *Proceedings of the 65th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 862–869, 2024.
- [6] M. Bafna and D. Minzer. Characterizing direct product testing via coboundary expansion. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing (STOC)*, pages 1978–1989, 2024.
- [7] M. Bafna, D. Minzer, and N. Vyas. Quasi-linear size PCPs with small soundness from HDX. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing (STOC)*, pages 45–53, 2025.

- [8] A. Balliu, C. Coupette, A. Cruciani, F. d’Amore, M. Equi, H. Lievonen, A. Modanese, D. Olivetti, and J. Suomela. New limits on distributed quantum advantage: Dequantizing linear programs. In *International Symposium on Distributed Computing*, pages 1–22. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2025.
- [9] M. Bun, M. Gaboardi, M. Hopkins, R. Impagliazzo, R. Lei, T. Pitassi, J. Sorrell, and S. Sivakumar. Stability is stable: Connections between replicability, privacy, and adaptive generalization. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing (STOC)*, pages 520–527, 2023.
- [10] Y.-J. Chang and Z. Li. The complexity of distributed approximation of packing and covering integer linear programs. In *Proceedings of the 2023 ACM Symposium on Principles of Distributed Computing (PODC)*, pages 32–43, 2023.
- [11] M. Chlebík and J. Chlebíková. Approximation hardness of dominating set problems in bounded degree graphs. *Information and Computation*, 206(11):1264–1275, 2008.
- [12] X. Coiteux-Roy, F. d’Amore, R. Gajjala, F. Kuhn, F. Le Gall, H. Lievonen, A. Modanese, M.-O. Renou, G. Schmid, and J. Suomela. No distributed quantum advantage for approximate graph coloring. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing (STOC)*, pages 1901–1910, 2024.
- [13] C. Coupette and C. Lenzen. A breezing proof of the KMW bound. In *Proceedings of the Symposium on Simplicity in Algorithms (SOSA)*, pages 184–195, 2021.
- [14] I. Diakonikolas, J. Gao, D. Kane, S. Liu, and C. Ye. Replicable distribution testing. *Advances in Neural Information Processing Systems*, 38, 2025.
- [15] Y. Dikstein and I. Dinur. Agreement theorems for high dimensional expanders in the low acceptance regime: The role of covers. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing (STOC)*, pages 1967–1977, 2024.
- [16] Y. Dikstein and I. Dinur. Swap cosystolic expansion. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing (STOC)*, pages 1956–1966, 2024.
- [17] Y. Dikstein, I. Dinur, and A. Lubotzky. Low acceptance agreement tests via bounded-degree symplectic HDXs. In *Proceedings of the 65th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 826–861, 2024.
- [18] Y. Dikstein and M. Hopkins. Chernoff bounds and reverse hypercontractivity on HDX. In *Proceedings of the 65th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 870–919, 2024.
- [19] Y. Dikstein, M. Hopkins, R. Impagliazzo, and T. Pitassi. High rate efficient local list decoding from hdx. *arXiv preprint arXiv:2601.22535*, 2026.
- [20] I. Dinur and P. Harsha. Composition of low-error 2-query PCPs using decodable PCPs. *SIAM Journal on Computing*, 42(6):2452–2486, 2013.
- [21] I. Dinur and I. Livni Navon. Exponentially small soundness for the direct product z-test. *Theory of Computing*, 19(1):1–56, 2023.

- [22] I. Dinur and O. Meir. Derandomized parallel repetition via structured pcps. *computational complexity*, 20(2):207–327, 2011.
- [23] M. Dory, M. Ghaffari, and S. Ilchi. Near-optimal distributed dominating set in bounded arboricity graphs. In *Proceedings of the 2022 ACM Symposium on Principles of Distributed Computing*, pages 292–300, 2022.
- [24] C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference*, pages 265–284. Springer, 2006.
- [25] E. Eaton, M. Hussing, M. Kearns, and J. Sorrell. Replicable reinforcement learning. *Advances in Neural Information Processing Systems*, 36, 2024.
- [26] M. Ebbens and Y. Yoshida. Average sensitivity of geometric algorithms. In *17th Innovations in Theoretical Computer Science Conference (ITCS 2026)*, pages 53–1, 2026.
- [27] U. Feige. A threshold of  $\ln n$  for approximating set cover. *Journal of the ACM*, 45(4):634–652, 1998.
- [28] N. Fleming and Y. Yoshida. Sensitivity lower bounds for approximation algorithms. In K. G. Larsen and B. Saha, editors, *Proceedings of the 2026 Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2026, Vancouver, BC, Canada, January 11-14, 2026*, pages 4043–4076. SIAM, 2026.
- [29] C. Gavoille, A. Kosowski, and M. Markiewicz. What can be observed locally? round-based models for quantum distributed computing. In *Proceedings of the International Symposium on Distributed Computing (DISC)*, pages 243–257, 2009.
- [30] M. Ghaffari, F. Kuhn, and Y. Maus. On the complexity of local distributed graph problems. In *Proceedings of the 49th Annual ACM Symposium on Theory of Computing (STOC)*, pages 784–797, 2017.
- [31] S. Hara and Y. Yoshida. Average sensitivity of decision tree learning. In *Proceedings of the 11th International Conference on Learning Representations (ICLR)*, 2023.
- [32] P. Harsha. Limits of approximation algorithms: PCPs and Unique Games. Technical report, Tata Institute of Fundamental Research, 2010.
- [33] J. Håstad. Some optimal inapproximability results. *Journal of the ACM*, 48(4):798–859, 2001.
- [34] A. E. Holroyd, O. Schramm, and D. B. Wilson. Finitary coloring. *The Annals of Probability*, pages 2867–2898, 2017.
- [35] M. Hopkins, R. Impagliazzo, D. Kane, S. Liu, and C. Ye. Replicability in high dimensional statistics. In *Proceedings of the 65th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 1–8, 2024.
- [36] M. Hopkins, S. Liu, C. Ye, and Y. Yoshida. From generative to episodic: Sample-efficient replicable reinforcement learning. *arXiv preprint arXiv:2507.11926*, 2025.
- [37] M. Hopkins and S. Moran. The role of randomness in stability. In *Proceedings of the 42nd International Conference on Machine Learning (ICML)*, 2025.
- [38] R. Impagliazzo, V. Kabanets, and A. Wigderson. New direct-product testers and 2-query pcps. *SIAM Journal on Computing*, 41(6):1722–1768, 2012.

- [39] R. Impagliazzo, R. Lei, T. Pitassi, and J. Sorrell. Reproducibility in learning. In *Proceedings of the 54th Annual ACM Symposium on Theory of Computing (STOC)*, pages 818–831, 2022.
- [40] L. Jia, R. Rajaraman, and T. Suel. An efficient distributed algorithm for constructing small dominating sets. *Distributed Computing*, 15(4):193–205, 2002.
- [41] A. Karbasi, G. Velegkas, L. Yang, and F. Zhou. Replicability in reinforcement learning. *Advances in Neural Information Processing Systems*, 36:74702–74735, 2023.
- [42] F. Kuhn, T. Moscibroda, and R. Wattenhofer. Local computation: Lower and upper bounds. *Journal of the ACM*, 63(2):1–44, 2016.
- [43] F. Kuhn and R. Wattenhofer. Constant-time distributed dominating set approximation. In *Proceedings of the twenty-second annual symposium on Principles of distributed computing*, pages 25–32, 2003.
- [44] S. Kumabe and Y. Yoshida. Average sensitivity of dynamic programming. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1925–1961, 2022.
- [45] N. Linial. Locality in distributed graph algorithms. *SIAM Journal on Computing*, 21(1):193–201, 1992.
- [46] S. Liu and C. Ye. Replicable uniformity testing. *Advances in Neural Information Processing Systems*, 37:32039–32075, 2024.
- [47] C. Lund and M. Yannakakis. On the hardness of approximating minimization problems. *Journal of the ACM*, 41(5):960–981, 1994.
- [48] D. Moshkovitz. Sliding scale conjectures in pcp. *ACM SIGACT News*, 50(3):25–33, 2019.
- [49] D. Moshkovitz and R. Raz. Two-query PCP with subconstant error. *Journal of the ACM*, 57(5):1–29, 2008.
- [50] S. Murai and Y. Yoshida. Sensitivity analysis of centralities on unweighted networks. In *Proceedings of the Web Conference (WWW)*, pages 1332–1342, 2019.
- [51] R. O’Donnell and N. G. Singer. Low-soundness direct-product testers and pcps from kaufman–oppenheim complexes. *arXiv preprint arXiv:2511.10514*, 2025.
- [52] R. Raz. A parallel repetition theorem. In *Proceedings of the twenty-seventh annual ACM symposium on Theory of computing*, pages 447–456, 1995.
- [53] N. Varma and Y. Yoshida. Average sensitivity of graph algorithms. *SIAM Journal on Computing*, 52(4):1039–1081, 2023.
- [54] Y. Yoshida and S. Ito. Average sensitivity of euclidean k-clustering. *Advances in Neural Information Processing Systems*, 35:32487–32498, 2022.
- [55] Y. Yoshida and Z. Zhang. Low-sensitivity matching via sampling from gibbs distributions. In *Proceedings of the 2026 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 120–149, 2026.
- [56] Y. Yoshida and S. Zhou. Sensitivity analysis of the maximum matching problem. In *Proceedings of the 12th Innovations in Theoretical Computer Science Conference (ITCS)*, pages 58–1, 2021.