# Nearly Linear Sample Bounds for Testing Product-State Satisfiability of Dense Quantum $k$-SAT

Zicheng Han* Xuhong Liu† Shuai Shao‡ Xiande Zhang§

University of Science and Technology of China

**Abstract**

It is a fundamental problem in classical property testing to distinguish satisfiable instances of dense constraint satisfaction problems from those far from satisfiable using only a small random induced subinstance. We study the corresponding quantum problem of testing product-state satisfiability for dense quantum $k$-SAT on $n$ qudits of local dimension $d$. Montanaro, Shao, and Verdon [MSV25] previously obtained a sample bound independent of $n$ for qubits. However, this bound has an iterated-factorial dependence on $1/\varepsilon$, which makes the bound extremely large.

We improve the sample bound to $O_{k,d}(\varepsilon^{-1}\log(2/\varepsilon))$ for every fixed $k$ and $d$. Our method also establishes the same sample bound for classical $k$-ary constraint satisfaction problems over a $d$-element alphabet, improving the bound of Blais and Seth [BS24] by a logarithmic factor for fixed $k$ and $d$. For graph $r$-colorability, our method gives an $O(r\varepsilon^{-1}\log(2/\varepsilon))$ sample bound matching that of Kushnir and Shapira [KS26].

## 1 Introduction

Deciding whether a large system of local constraints is satisfiable can be computationally expensive. A natural relaxation asks whether one can distinguish satisfiable instances from those that are far from satisfiable by using only a small random induced subinstance. This question is central to property testing, introduced by Rubinfeld and Sudan [RS96], and has been studied extensively for dense graph coloring [GGR98] and constraint satisfaction problems [AE02, AS03]. Specifically, consider a constraint satisfaction problem (CSP) on $n$ variables over an alphabet of size $d$, with each constraint involving exactly $k$ variables. In the dense setting, i.e., when the number of constraints is $\Theta(n^k)$, an instance is $\varepsilon$-far from satisfiable if at least $\varepsilon n^k$ constraints must be deleted to make it satisfiable. The natural *canonical tester* samples a set of variables uniformly at random and accepts if and only if the induced subinstance is satisfiable. For fixed $k$ and $d$, Alon and Shapira [AS03] proved that $O_{k,d}(\varepsilon^{-2})$ sampled variables suffice: the tester always accepts satisfiable instances and rejects $\varepsilon$-far instances with probability at least $3/4$, independently of $n$. Subsequent work improved the sample complexity and its dependence on the parameters to $O_{k,d}\Big(\varepsilon^{-1}\log^2(1/\varepsilon)\log\log(1/\varepsilon)\Big)$ [Soh12] and $O_{k,d}(\varepsilon^{-1}\log^2(2/\varepsilon))$ [BS24] respectively. For graph $r$-colorability testing, with improved bounds obtained by Alon and Krivelevich [AK02]. Using graph containers lemma, Blais and Seth [BS25] obtained a $\widetilde{O}(r/\varepsilon)$ sample bound. Kushnir and Shapira [KS26] improved this to $O(r\varepsilon^{-1}\log(2/\varepsilon))$ and proved an $\Omega(\varepsilon^{-1}\log r)$ lower bound.

*hzcqj2020@mail.ustc.edu.cn

†xuhongliu@mail.ustc.edu.cn

‡shao10@ustc.edu.cn

§drzhangx@ustc.edu.cn

In this work, we study testing product-state satisfiability of quantum $k$-SAT instances on qudits using only a small random induced subinstance. Quantum $k$-SAT, introduced by Bravyi for qubits [Bra11], represents local constraints by orthogonal projectors. Formally, let $V = [n]$ index $n$ qudits of local dimension $d$, and let $\mathcal{H}_x = \mathbb{C}^d$ denote the Hilbert space of qudit $x$ for each $x \in V$. For $T \subseteq V$, write $\mathcal{H}_T = \bigotimes_{x \in T} \mathcal{H}_x$. An instance is specified by a family $\Phi = \{\Pi_S : S \subseteq V,\ 1 \le |S| \le k\}$, where each $\Pi_S$ is an orthogonal projector on $\mathcal{H}_V$ that acts nontrivially only on the qudits in $S$; zero projectors are allowed. An instance $\Phi$ is satisfiable if there exists a state $|\psi\rangle \in \mathcal{H}_V$ such that $\Pi_S|\psi\rangle = 0$ for every $S$, and it is *product satisfiable* if $|\psi\rangle = \bigotimes_{x \in V} |\psi_x\rangle$ is a product state. We say that $\Phi$ is *$\varepsilon$-far from product satisfiable* if at least $\varepsilon n^k$ constraints must be removed from $\Phi$ to obtain an instance with a product solution, i.e., for every product state $|\psi\rangle \in \mathcal{H}_V$, we have

$$|\{S \subseteq V : 1 \le |S| \le k,\ \Pi_S|\psi\rangle \neq 0\}| \ge \varepsilon n^k.$$

For every $C \subseteq V$, write $\Phi[C] := \{\Pi_S : S \subseteq C,\ 1 \le |S| \le k\}$ for the instance induced by $C$. Here each projector is viewed as an operator on $\mathcal{H}_C$ by omitting the identity factors outside $C$. Following the standard dense-property-testing convention [AS03, BS24], we use the following definitions.

**Definition 1.1** ($\varepsilon$-test of product-satisfiability of quantum $k$-SAT)**.** Given a quantum $k$-SAT instance $\Phi$ on $n$ qudits of local dimension $d$ that is either product satisfiable or $\varepsilon$-far from product satisfiability, a *canonical $\varepsilon$-tester* with sample size $s \le n$ samples a set $C \subseteq V$ uniformly at random with $|C| = s$. The tester accepts $\Phi$ if and only if $\Phi[C]$ is product satisfiable, and rejects every $\varepsilon$-far instance with probability at least $3/4$. We denote this testing problem by $(\varepsilon, k, d)$-Prod-QSAT, and its sample complexity $s(\varepsilon, k, d)$ is the smallest sample size $s$ for which these guarantees hold for every $n \ge s$.

*Remark* 1.2. The classical case is a special case of $(\varepsilon, k, d)$-Prod-QSAT and is denoted by $(\varepsilon, k, d)$-SAT. The classical version considers assignments in $[d]^V$, whereas the quantum version considers product solutions in the Hilbert space $\mathcal{H}_V$.

Montanaro, Shao, and Verdon [MSV25] first showed that, for sufficiently large $n$, a random sample of qubits of size independent of $n$ suffices to distinguish dense quantum $k$-SAT instances that are satisfiable, possibly by an entangled state, from those that are $\varepsilon$-far from satisfiable by a product state. Their proof combines tools from algebraic geometry with backtracking to handle the continuous space of quantum states. Although the sample bound is independent of $n$, its dependence on $\varepsilon$ and $k$ involves iterated factorials, making the bound extremely large.

## Our contribution

We give an explicit sample complexity that is nearly linear on $1/\varepsilon$ for $(\varepsilon, k, d)$-Prod-QSAT. This also implies an improved sample complexity for classical $(\varepsilon, k, d)$-SAT.

**Theorem 1.3.** *The sample complexity of* $(\varepsilon, k, d)$-Prod-QSAT *is* $O_{k,d}(\varepsilon^{-1} \log(2/\varepsilon))$.

We study testing product-state satisfiability, whereas [MSV25] allow YES instances to be satisfied by arbitrary quantum states. Nevertheless, our result also improves the sample complexity in their original setting. In both settings, the tester must reject instances that are $\varepsilon$-far from product-state satisfiable with probability at least $3/4$. Our main result improves the sample size needed to guarantee this rejection. Combining this result with [MSV25, Theorem 1.6], we obtain a sample bound of $O_k(\varepsilon^{-1} \log(2/\varepsilon))$ in their original setting for qubit instances and sufficiently large $n$.

Our extension to qudits of local dimension $d$ also covers classical CSPs over $d$-element alphabets. These correspond to instances in which all constraint projectors are diagonal in a fixed local product basis. Thus, Theorem 1.3 yields the following corollary.

**Corollary 1.4.** *The sample complexity of* $(\varepsilon, k, d)$-SAT *is* $O_{k,d}(\varepsilon^{-1}\log(2/\varepsilon))$.

For fixed $k$ and $d$, this improves the previous sample bound for general CSPs [BS24] by a logarithmic factor. Viewing graph $r$-colorability testing as $(\varepsilon, 2, r)$-SAT and tracking the dependence on $r$ in our classical sampling argument yields the following corollary.

**Corollary 1.5.** *The sample complexity of* $r$*-colorability of graphs is* $O(r\varepsilon^{-1}\log(2/\varepsilon))$.

These bounds suggest that, in terms of sample complexity, quantum testing is no harder than classical testing in our setting. The following table compares our results with previous work.

| **Problem** | **Previous work** | **This work** |
|---|---|---|
| $(\varepsilon, k, d)$**-Prod-QSAT** | Iterated factorials [MSV25] $(d = 2)$ | $O_{k,d}(\varepsilon^{-1}\log(2/\varepsilon))$ |
| $(\varepsilon, k, d)$**-SAT** | $O_{k,d}(\varepsilon^{-2})$ [AS03]<br>$O_{k,d}\Big(\varepsilon^{-1}\log^2(1/\varepsilon)\log\log(1/\varepsilon)\Big)$ [Soh12]<br>$O_{k,d}(\varepsilon^{-1}\log^2(2/\varepsilon))$ [BS24] | $O_{k,d}(\varepsilon^{-1}\log(2/\varepsilon))$ |
| **Graph** $r$**-colorability** | $O(r\varepsilon^{-2}\log r)$ [AK02]<br>$O(r\varepsilon^{-1}\log^2(2/\varepsilon))$ [BS25]<br>$O(r\varepsilon^{-1}\log(2/\varepsilon))$ [KS26] | $O(r\varepsilon^{-1}\log(2/\varepsilon))$ |

Table 1: Comparison of sample bounds. The notation $O_{k,d}$ hides the constant factor depending on $k, d$.

More broadly, our setting can be viewed as testing a structured CSP over a continuous domain. In this formulation, variables range over single-qudit pure states, and local projector constraints impose homogeneous multilinear equations on the state amplitudes [LLM$^+$10, AGR26]. We hope that this perspective, together with our techniques, will motivate further work on testing CSPs over continuous domains with similar algebraic structure.

## Proof Overview and Technical Innovations

We call a product-state assignment to a subset of qudits a *local assignment* if it satisfies all constraints supported entirely on that subset. Following the idea of [MSV25] a natural next step is to extend a local assignment to more qudits. As more qudits are assigned, additional constraints come into play, making further extensions more difficult. We therefore introduce a potential $W_{\mathcal{A}}$ to quantify the remaining freedom to extend a local assignment $\mathcal{A}$. Intuitively, we view $W_{\mathcal{A}}$ as the remaining *budget* for extensions, with each extension spending part of this budget. If adding a qudit consumes a large part of the budget, further extensions become more difficult. Following Sohler's historical minimum-cost approach [Soh12], we consider extending a given local assignment $\mathcal{A}$ to an unassigned qudit. We look back over the entire assignment process to find when this qudit could have been added at minimum cost. If even this minimum cost is at least $\tau_i n^{k-1}$ with $\tau_i = 2^{-i}$, we call it a *bad qudit* at scale $i$. All such qudits form the *bad set* $R_i(\mathcal{A})$ at scale $i$.

We prove two key lemmas to prove the main theorem, the first guarantees a sufficient supply of bad qudits, while the second shows that a small random sample can find them.

**Lemma 1.6** (Informal statement for Lemma 3.1)**.** *If an instance is* $\varepsilon$*-far from product-state satisfiable, then every local assignment* $\mathcal{A}$ *has sufficiently many bad qudits.*

**Lemma 1.7** (Informal statement for Lemma 3.2)**.** *A small random sample contains a bad qudit from every sufficiently large bad set with high probability.*

**Proof sketch of Theorem 3.3.** Choose a random sample and suppose it has a product-state solution. We start with the empty local assignment and repeatedly update it as follows. Whenever the current assignment has sufficiently many bad qudits, Lemma 3.2 allows us to update it using a bad qudit in the sample and its state in the product-state solution. Each update consumes part of the budget. Since the budget is bounded, the process terminates after finitely many steps. Then we get a local assignment that has too few bad qudits to allow further extension, contradicting Lemma 3.1. Thus, the sample has no product-state solution with high probability. □

The proof of Lemma 3.1 builds on [MSV25, Lemma 3.10], where constraints are removed after each extension to preserve the spaces of possible extensions, and the number of removed constraints is bounded by a weighted sum of dimension reductions. We combine this argument with Sohler's approach [Soh12] of taking the minimum weight attained by each variable during the assignment process and assigning variables in increasing order of their time stamps.

The proof of Lemma 3.2 takes a different approach from [MSV25]. Before we compare them, we first recall why the classical proof does not extend directly to the quantum setting. The classical approach builds a finite tree of assignments and applies a union bound over all branches. In the quantum setting, each qubit has infinitely many possible states, so such a finite tree cannot be constructed directly. [MSV25] use algebraic geometry to reduce the candidate states considered during backtracking to a finite set. Repeatedly applying the factorial bound on the number of candidates in the recursion leads to a sample bound involving iterated factorials.

In this paper, we use uniform approximation and the signs of polynomials to control the complexity arising from the continuous space of quantum states. Our approach features the following two key innovations. First, we approximate the remaining budget using a small number of sampled terms, yielding approximate bad sets that are easier to analyze while preserving the required properties. This step makes the subsequent analysis independent of the instance size $n$. Second, we describe these approximate bad sets using the signs of polynomials, thereby addressing the challenge posed by continuous assignments. Thus, a single random sample hits all relevant large approximate bad sets with high probability. This completes the proof of Lemma 3.2.

### Organization

Section 2 establishes the notation and introduces local assignments, extension spaces, the potential, and historical bad sets. In Section 3, we combine a weighted density lemma with a simultaneous hitting lemma to prove the main sample-complexity bound and derive its quantum and classical corollaries. Section 4 proves these lemmas: the first uses a constraint-deletion argument, while the second uses uniform approximation of the potential and growth-function bounds for approximate bad sets. Finally, Appendix A proves the graph-colorability corollary.

## 2 Preliminaries

We write $[n] = \{1, \dots, n\}$, $|S|$ for the cardinality of a set $S$, and $\binom{S}{j}$ for the family of its $j$-element subsets. We use standard linear algebra notation, with $M^*$ denoting the conjugate transpose of $M$. All logarithms are natural unless a base is specified. The notation $O_{k,d}(\cdot)$ hides constants depending only on $k$ and $d$.

We retain the notation introduced in Section 1. For a product state $|\psi\rangle = \bigotimes_{x \in V} |\psi_x\rangle$ and $S \subseteq V$, write $|\psi_S\rangle := \bigotimes_{x \in S} |\psi_x\rangle$. Whenever $S \subseteq T \subseteq V$ and $1 \le |S| \le k$, we also use $\Pi_S$ for its extension to

$\mathcal{H}_T$, acting as the identity on the qudits in $T \setminus S$.

## 2.1 Local assignments and bad sets

In this subsection, we recall the framework of Montanaro, Shao, and Verdon [MSV25]. Combining this framework with Sohler's historical minimum-cost idea [Soh12], we introduce the basic definitions of local assignments, extension spaces, potentials, and bad sets used in our analysis.

**Definition 2.1** (Local assignment)**.** A *local assignment* is a pair $\mathcal{A} = (F, |\psi\rangle)$, where $F = (x_1, \ldots, x_T)$ is an ordered tuple of distinct indices in $V$, called its *support*. The state $|\psi\rangle = \bigotimes_{j=1}^{T} |\psi_{x_j}\rangle \in \mathcal{H}_F$ is a product state satisfying every constraint supported entirely on $F$. We write $\mathcal{P}(F)$ for the set of all such product states and $|\mathcal{A}| := |F| = T$ for the length of $\mathcal{A}$.

The ordered support determines a sequence of prefix assignments. For $1 \le t \le T$, set $\mathcal{A}_t := ((x_1, \ldots, x_t), |\psi^{(t)}\rangle)$, where $|\psi^{(t)}\rangle = \bigotimes_{j=1}^{t} |\psi_{x_j}\rangle$. Set $\mathcal{A}_0 := (\varnothing, -)$, where the dash denotes the empty state assignment. Then $\mathcal{A}_T = \mathcal{A}$. We call $\mathcal{A}_t$ the *history* of $\mathcal{A}$ at time $t$, for $0 \le t \le T$.

Given a local assignment $\mathcal{A} = (F, |\psi\rangle)$, we call the qudits in $V \setminus F$ *unassigned*.[1] A natural way to extend a local assignment is to choose a state for an unassigned qudit and take its tensor product with the state of the local assignment. This motivates the following definition of the extension space of $\mathcal{A}$ on $X$, where $X \subseteq V \setminus F$ and $1 \le |X| \le k-1$:

$$L_{\mathcal{A}}(X) = \{|v\rangle \in \mathcal{H}_X : \Pi_S(|v\rangle \otimes |\psi\rangle) = 0 \text{ for every } S \text{ with } X \subseteq S \subseteq F \cup X \text{ and } |S| \le k\}.$$

*Remark* 2.2. The vectors in $L_{\mathcal{A}}(X)$ are not required to be normalized. Since the defining constraints are homogeneous linear equations, $L_{\mathcal{A}}(X)$ is a linear subspace of $\mathcal{H}_X$.

For an unassigned qudit $x \in V \setminus F$, we say that a state $|v\rangle$ is *compatible* at $x$ if $|v\rangle \in L_{\mathcal{A}}(\{x\})$. The extended local assignment $\mathcal{A}_{(x,|v\rangle)}$ is obtained by appending $x$ to the tuple and assigning it the compatible state $|v\rangle$:

$$\mathcal{A}_{(x,|v\rangle)} = ((x_1, \ldots, x_T, x), |\psi\rangle \otimes |v\rangle),$$

and $(\mathcal{A}_0)_{(x,|v\rangle)} = ((x), |v\rangle)$.

We treat the weighted dimensions of the extension spaces as a budget $W_{\mathcal{A}}$ for future extensions. Assigning $x$ the state $|\eta\rangle$ spends $\Delta_{\mathcal{A}}(x, |\eta\rangle)$ of this budget.

**Definition 2.3** (Extension potential and exact drop)**.** The extension potential of a local assignment $\mathcal{A}$ is

$$W_{\mathcal{A}} = \sum_{\substack{X \subseteq V \setminus F \\ 1 \le |X| \le k-1}} n^{k-|X|-1} \dim L_{\mathcal{A}}(X).$$

Its exact potential drop at $(x, |v\rangle)$ is

$$\Delta_{\mathcal{A}}(x, |v\rangle) = W_{\mathcal{A}} - W_{\mathcal{A}_{(x,|v\rangle)}}.$$

*Remark* 2.4. As mentioned in Section 1, we may think of it as a budget: each extension spends part of this budget, and the potential drop records the amount spent. Extension constraints are monotone, so $\Delta_{\mathcal{A}}(x, |v\rangle) \ge 0$. The complete drop includes both the contraction of surviving extension spaces and the removal of summands indexed by sets containing $x$. Since $\dim L_{\mathcal{A}}(X) \le d^{|X|}$, we have the capacity bound

$$0 \le W_{\mathcal{A}} \le n^{k-1} \sum_{j=1}^{k-1} \frac{d^j}{j!} \le e^d n^{k-1}. \tag{1}$$

[1] Although $F$ is an ordered tuple, we identify it with the set of its entries whenever it occurs in set notation or as a subsystem label.

For $0 < \varepsilon \leq 1$, set $J = \lceil \log_2(16/\varepsilon) \rceil$ and $\tau_i = 2^{-i}$ for $1 \leq i \leq J+1$. We call an unassigned qudit "bad" at scale $i$ if even its cheapest compatible insertion over the assignment history would spend at least $\tau_i n^{k-1}$ of the potential budget.

**Definition 2.5** (Bad set)**.** Fix a local assignment $\mathcal{A}_T$ with support $F_T$, and let $(\mathcal{A}_t)_{t=0}^T$ be its history. For $1 \leq i \leq J+1$, define the bad set as follows:

$$R_i(\mathcal{A}_T) = \left\{ y \in V \setminus F_T : \begin{array}{c} \Delta_{\mathcal{A}_t}(y, |\eta\rangle) \geq \tau_i n^{k-1} \\ \text{for every } t \in \{0, \ldots, T\} \\ \text{and every unit vector } |\eta\rangle \in L_{\mathcal{A}_t}(\{y\}) \end{array} \right\}.$$

*Remark* 2.6. Intuitively, $R_i(\mathcal{A}_T)$ consists of the unassigned qudits whose insertion into the local assignment at any time with any compatible state incurs a potential drop of at least $\tau_i n^{k-1}$. Moreover, given a local assignment $\mathcal{A}_T$, the definition also immediately implies the following monotonicity:

$$R_i(\mathcal{A}_{t'}) \subseteq R_i(\mathcal{A}_t), \qquad \text{for } 0 \leq t \leq t' \leq T$$

# 3 Proof of the Main Theorem

We prove the main theorem and derive its classical corollaries using two key lemmas, whose proofs are deferred to Section 4. We first state these lemmas.

**Lemma 3.1.** *Suppose that $\Phi$ is $\varepsilon$-far from product-state satisfiability. Then every local assignment $\mathcal{A}$ satisfies*

$$\sum_{i=1}^{J} \tau_i |R_i(\mathcal{A})| \geq \frac{15\varepsilon n}{16}.$$

**Lemma 3.2.** *Fix $k, d \geq 2$, and let $A_{k,d} > 0$ be sufficiently large. For $0 < \varepsilon \leq 1$, choose a uniform random $c$-subset $C \subseteq V$, where $c = \lceil A_{k,d} \varepsilon^{-1} \log(2/\varepsilon) \rceil \leq n$. With probability at least $3/4$,*

$$|R_i(\mathcal{A})| > \frac{\varepsilon n(|\mathcal{A}|+1)}{128 e^d} \quad \Longrightarrow \quad C \cap R_{i+1}(\mathcal{A}) \neq \varnothing \tag{2}$$

*holds simultaneously for every $i \in [J]$ and every local assignment $\mathcal{A}$ supported on $C$ with $|\mathcal{A}| \leq 4e^d/\tau_i$.*

Lemma 3.1 guarantees a sufficient weighted supply of bad qudits, while Lemma 3.2 ensures that a single random sample hits the relevant large bad sets. Their proofs appear in Sections 4.1 and 4.2. We now combine the two lemmas to prove the main theorem.

**Theorem 3.3.** *The sample complexity of $(\varepsilon, k, d)$-*Prod-QSAT *is $O_{k,d}(\varepsilon^{-1} \log(2/\varepsilon))$.*

*Proof.* Since the restriction of a product-state solution to any induced subinstance is also a product-state solution, it suffices to consider instances that are $\varepsilon$-far from product-state satisfiability.

Let $A_{k,d} > 0$ be sufficiently large for Lemma 3.2, and set

$$c = \left\lceil A_{k,d} \varepsilon^{-1} \log \frac{2}{\varepsilon} \right\rceil. \tag{3}$$

For $n < c$, the tester inspects the entire instance and decides product satisfiability exactly.

For $n \geq c$, choose $C$ uniformly from $\binom{V}{c}$. Fix a sample $C$ satisfying (2) and suppose that $\Phi[C]$ has a product-state solution $|\psi_C\rangle = \bigotimes_{x \in C} |\psi_x\rangle$ for a contradiction.

Starting from the empty local assignment $\mathcal{A}_0$, we use the following algorithm to construct a local assignment.

**Algorithm 1** Extending along a product-state solution

**Input:** A sample $C$ satisfying (2), and a product-state solution $|\psi_C\rangle = \bigotimes_{x\in C}|\psi_x\rangle$ of $\Phi[C]$.
**Output:** A local assignment $\mathcal{A}_{\text{fin}}$.

$\mathcal{A} \leftarrow \mathcal{A}_0$.
$T \leftarrow \left\{i \in [J] : |R_i(\mathcal{A})| > \frac{\varepsilon n(|\mathcal{A}|+1)}{128e^d}\right\}$.
**while** $T \neq \varnothing$ **do**
    $i \leftarrow \min T$.
    Let $F$ be the support of $\mathcal{A}$.
    Choose $x \in C \cap R_{i+1}(\mathcal{A})$.
    $\mathcal{A} \leftarrow \mathcal{A}_{(x,|\psi_x\rangle)}$.
    $T \leftarrow \left\{i \in [J] : |R_i(\mathcal{A})| > \frac{\varepsilon n(|\mathcal{A}|+1)}{128e^d}\right\}$.
$\mathcal{A}_{\text{fin}} \leftarrow \mathcal{A}$.
**return** $\mathcal{A}_{\text{fin}}$.

We first verify that every iteration is well defined. Throughout the algorithm, $\mathcal{A}$ agrees with $|\psi_C\rangle$ on its ordered support $F \subseteq C$. Thus $\mathcal{A}$ is a local assignment, and $|\psi_x\rangle$ is compatible with $\mathcal{A}$ for every unassigned $x \in C$. Moreover, by the monotonicity in Remark 2.6, $R_i(\mathcal{A}') \subseteq R_i(\mathcal{A})$ for every $i \in [J]$ whenever $\mathcal{A}'$ extends $\mathcal{A}$. Thus $T$ can only shrink throughout the procedure.

To justify each application of Lemma 3.2, consider an iteration at scale $i$, assuming inductively that all previous iterations were well defined. Every completed extension was made at a scale $j \le i$ and decreased the potential by at least $\tau_j n^{k-1}/2 \ge \tau_i n^{k-1}/2$. Since exactly $|\mathcal{A}|$ extensions have been completed, (1) gives

$$\frac{\tau_i n^{k-1}}{2}|\mathcal{A}| \le W_{\mathcal{A}_0} - W_{\mathcal{A}} \le e^d n^{k-1}.$$

Consequently, $|\mathcal{A}| \le 2e^d/\tau_i < 4e^d/\tau_i$. Thus, the size condition in Lemma 3.2 also holds, so

$$C \cap R_{i+1}(\mathcal{A}) \neq \varnothing.$$

The algorithm can therefore select $x$ from this intersection and extend $\mathcal{A}$ using $|\psi_x\rangle$. By the definition of $R_{i+1}(\mathcal{A})$, this extension decreases the potential by at least $\tau_i n^{k-1}/2$. Since $0 \le W_{\mathcal{A}} \le e^d n^{k-1}$, at most $2e^d/\tau_i$ extensions are possible at scale $i$.

Let $m_i$ be the number of extensions at scale $i$, and let $|\mathcal{A}| = t_i = \sum_{j=1}^{i} m_j$ be the assignment length after processing scales $1, \dots, i$. Processing scale $i$ must therefore end with an assignment $\mathcal{A}$ of length $t_i$ satisfying

$$|R_i(\mathcal{A})| \le \frac{\varepsilon n(t_i+1)}{128e^d};$$

otherwise, Lemma 3.2 would permit further extensions, contradicting the potential bound.

Summing the potential drops gives $\sum_{i=1}^{J} m_i\tau_i \le 2e^d$. The monotonicity of the bad sets implies

$$|R_i(\mathcal{A}_{\text{fin}})| \le \frac{\varepsilon n(t_i+1)}{128e^d} \qquad \text{for every } i \in [J].$$

Using $\sum_{i=j}^{J} \tau_i \le 2\tau_j$, we obtain

$$\begin{aligned}\sum_{i=1}^{J} \tau_i |R_i(\mathcal{A}_{\text{fin}})| &\le \frac{\varepsilon n}{128 e^d} \sum_{i=1}^{J} \tau_i (t_i + 1) \\ &\le \frac{\varepsilon n}{128 e^d} \left(1 + 2 \sum_{i=1}^{J} m_i \tau_i\right) \\ &\le \frac{5\varepsilon n}{128} < \frac{15\varepsilon n}{16},\end{aligned}$$

contradicting Lemma 3.1. Thus every sample satisfying (2) induces an instance with no product-state solution. By Lemma 3.2, this occurs with probability at least 3/4. □

A classical constraint with allowed assignments $P_S \subseteq [d]^S$ can be represented by the diagonal projector

$$\Pi_S = \sum_{a \in [d]^S \setminus P_S} |a\rangle\langle a|.$$

Under this encoding, classical satisfiability is equivalent to product-state satisfiability. Since deletion distance and induced subinstances are also preserved, Theorem 3.3 yields the following corollary.

**Corollary 3.4.** *The sample complexity of* $(\varepsilon, k, d)$-SAT *is* $O_{k,d}(\varepsilon^{-1} \log(2/\varepsilon))$.

Let $G = (V, E)$ be a graph with $|V| = n$. The graph $r$-colorability problem asks whether its vertices can be assigned colors from $[r]$ so that adjacent vertices receive different colors. Testing this property is a special case of $(\varepsilon, 2, r)$-SAT: each vertex $v$ corresponds to a variable $x_v \in [r]$, and each edge $uv \in E$ imposes the binary constraint $x_u \neq x_v$. Nonedges impose no restriction. We say that $G$ is $\varepsilon$-far from $r$-colorability if at least $\varepsilon n^2$ edges must be deleted to make it $r$-colorable. A proper partial coloring $\mathcal{A}$ on $F \subseteq V$ is a proper $r$-coloring of the induced subgraph $G[F]$. Our proof method yields the following upper bound, matching the bound obtained by Kushnir and Shapira [KS26].

**Corollary 3.5.** *The sample complexity of* $\varepsilon$*-testing* $r$*-colorability of graphs is* $O(r\varepsilon^{-1} \log(2/\varepsilon))$.

The proof of Corollary 3.5 requires a few refinements of our analysis to match the bound of [KS26]. Our approach differs from theirs, and we defer the proof to Appendix A.

# 4 Proofs of the two Lemmas

This section proves the two lemmas used in Section 3. We prove the weighted density bound of Lemma 3.1 in Section 4.1, and the simultaneous hitting guarantee of Lemma 3.2 in Section 4.2.

## 4.1 Proof of the weighted density lemma

The proof constructs a product solution after deleting constraints and bounds the deletion cost.

*Proof of Lemma 3.1.* **Step 1: Partition of the unassigned qudits.** Fix a local assignment $\mathcal{A}_T = (F_T, |\psi^{(T)}\rangle)$, where $F_T = (x_1, \ldots, x_T)$. For $0 \le h \le T$, write $\mathcal{A}_h = (F_h, |\psi^{(h)}\rangle)$ for its history at time $h$, with $F_h = (x_1, \ldots, x_h)$. For convenience, write $R_i = R_i(\mathcal{A}_T)$ when the local assignment is clear for $1 \le i \le J$. By the definition of $R_i$, we have the inclusion chain

$$\emptyset \subseteq R_1 \subseteq R_2 \subseteq \cdots \subseteq R_J \subseteq V \setminus F_T.$$

For this proof only, set $\tau_0 = 1$, $R_0 = \varnothing$, and $R_{J+1} = V \setminus F_T$. This gives the decomposition

$$V \setminus F_T = \bigsqcup_{i=0}^{J} (R_{i+1} \setminus R_i).$$

If $y \in R_{i+1} \setminus R_i$ for $1 \le i \le J$, then $y \notin R_i$. By Definition 2.5, we can choose a time $t_y \in \{0, \dots, T\}$ and a compatible state $|v_y\rangle \in L_{\mathcal{A}_{t_y}}(\{y\})$ such that $\Delta_{\mathcal{A}_{t_y}}(y, |v_y\rangle) < \tau_i n^{k-1}$. If $y \in R_1 = R_1 \setminus R_0$, set $t_y = -1$ and choose an arbitrary single-qudit state $|v_y\rangle$.

At this point, we have assigned a time $t_y$ and a state $|v_y\rangle$ to each unassigned qudit $y \in V \setminus F_T$. We then group the qudit–state pairs according to their assigned times by defining

$$U_i := \{(y, |v_y\rangle) : y \in V \setminus F_T,\ t_y = i\}, \qquad i = -1, 0, \dots, T.$$

In particular, $U_{-1} = \{(y, |v_y\rangle) : y \in R_1\}$. Each qudit in $V \setminus F_T$ occurs as the first component of exactly one pair in $U_{-1}, U_0, \dots, U_T$.

**Step 2: Construct a global product-state solution.** In this step, we design an algorithm that takes a local assignment $\mathcal{A}_T = (F_T, |\psi^{(T)}\rangle)$ as input. The algorithm starts from the empty local assignment and iteratively extends it, removing selected constraints along the way, until it yields a product-state solution of the remaining instance. To make the algorithm easier to follow, we first recall the relevant notation and introduce a few new symbols.

Let $\Sigma \subseteq \{S \subseteq V : 1 \le |S| \le k\}$ be the set of deleted constraint supports. During the construction, local assignments and compatibility are understood relative to the remaining instance. We use the extension notation of Section 2.1, appending each newly assigned qudit to the ordered support. For a local assignment $\mathcal{A} = (F, |\psi\rangle)$ of the remaining instance and $X \subseteq V \setminus F$, $1 \le |X| \le k-1$, define its extension space after deletion by

$$L_{\mathcal{A}}^{\Sigma}(X) := \left\{ |u\rangle \in \mathcal{H}_X : \begin{array}{l} \Pi_S(|u\rangle \otimes |\psi\rangle) = 0 \text{ for every} \\ X \subseteq S \subseteq F \cup X, \quad |S| \le k, \quad S \notin \Sigma \end{array} \right\}.$$

Thus $L_{\mathcal{A}}^{\varnothing}(X) = L_{\mathcal{A}}(X)$ whenever $\mathcal{A}$ is a local assignment of the original instance.

Throughout the construction, we extend the local assignment in the order shown below. When processing each $U_i$, we delete selected constraints to ensure that further extensions remain possible.

$$U_{-1} \longrightarrow U_0 \longrightarrow x_1 \longrightarrow U_1 \longrightarrow x_2 \longrightarrow U_2 \longrightarrow \cdots \longrightarrow x_T \longrightarrow U_T$$

Figure 1: Assignment order in the construction. Each $U_h$ is processed in an arbitrary order, and each $x_h$ is assigned its corresponding state.

For the current assignment $\mathcal{B} = (G, |\phi\rangle)$, an unassigned qudit $y \in V \setminus G$, and $X \subseteq V \setminus (G \cup \{y\})$ with $1 \le |X| \le k-1$, write

$$D_y(X) := \left\{ S \subseteq V : \begin{array}{l} X \cup \{y\} \subseteq S \subseteq G \cup \{y\} \cup X, \\ |S| \le k, \quad S \notin \Sigma \end{array} \right\}.$$

We now give a formal description of the algorithm:

**Algorithm 2** Constructing a product-state solution by deleting constraints

**Input:** A local assignment $\mathcal{A}_T = (F_T, |\psi^{(T)}\rangle)$, with $F_T = (x_1, \ldots, x_T)$.

**Output:** A global product state $|\phi\rangle$ satisfying all constraints outside $\Sigma_{\text{fin}}$.

**Initialization:** Calculate the stage sets $U_i$
$\Sigma \leftarrow \{S \subseteq V : 1 \le |S| \le k,\ S \cap R_1 \neq \varnothing\}$, $\mathcal{B} \leftarrow \mathcal{A}_0$. ▷ Delete all constraints whose supports intersect $R_1$.
**for** each $(y, |v_y\rangle) \in U_{-1}$, in an arbitrary order **do**
    $\mathcal{B} \leftarrow \mathcal{B}_{(y,|v_y\rangle)}$.
**for** $h = 0, 1, \ldots, T$ **do**
    **for** each $(y, |v_y\rangle) \in U_h$, in an arbitrary order **do**
        Write $\mathcal{B} = (G, |\phi\rangle)$.
        $\mathcal{B}' \leftarrow \mathcal{B}_{(y,|v_y\rangle)}$, $D \leftarrow \varnothing$.
        **for** each $X \subseteq V \setminus (G \cup \{y\})$ with $1 \le |X| \le k-1$ **do**
            Restore each extension space that shrinks after assigning $y$.
            **if** $L^{\Sigma}_{\mathcal{B}'}(X) \subsetneq L^{\Sigma}_{\mathcal{B}}(X)$ **then**
                $D \leftarrow D \cup D_y(X)$.
        $\Sigma \leftarrow \Sigma \cup D$, $\mathcal{B} \leftarrow \mathcal{B}'$.
    **if** $h < T$ **then**
        $x \leftarrow x_{h+1}$.
        $\mathcal{B} \leftarrow \mathcal{B}_{(x,|\psi_x\rangle)}$.
$\Sigma_{\text{fin}} \leftarrow \Sigma$.
Write $\mathcal{B} = (G, |\phi\rangle)$.
**return** $(|\phi\rangle, \Sigma_{\text{fin}})$.

*Correctness.* At the beginning of stage $h$ and after processing each pair in $U_h$, we maintain the following invariant for the current local assignment $\mathcal{B} = (G, |\phi\rangle)$ and deletion set $\Sigma$:

$$L^{\Sigma}_{\mathcal{B}}(X) = L_{\mathcal{A}_h}(X) \quad (X \subseteq V \setminus G,\ 1 \le |X| \le k-1). \tag{4}$$

At initialization, we delete all constraints whose supports intersect $R_1$. The invariant (4) then holds trivially because, for every $X$ disjoint from the assigned qudits in $R_1$, the constraints supported entirely on $X$ are unchanged.

At stage $h$, consider a pair $(y, |v_y\rangle) \in U_h$. If extending $\mathcal{B}$ by assigning $|v_y\rangle$ to $y$ yields $L^{\Sigma}_{\mathcal{B}'}(X) \subsetneq L^{\Sigma}_{\mathcal{B}}(X)$, we delete all constraints in $D_y(X)$, namely the surviving constraints containing $y$ whose remaining unassigned support is exactly $X$. Therefore, writing $\Sigma'$ for the updated deletion set, we obtain

$$L^{\Sigma'}_{\mathcal{B}'}(X) = L^{\Sigma}_{\mathcal{B}}(X) = L_{\mathcal{A}_h}(X).$$

If $L^{\Sigma}_{\mathcal{B}'}(X) = L^{\Sigma}_{\mathcal{B}}(X) = L_{\mathcal{A}_h}(X)$ after the extension, no deletions are needed for this $X$.

It remains to verify that the equality holds when we extend $\mathcal{B}$ by assigning $x_{h+1} \in F_T$ without any deletion. Write $\mathcal{B} = (G, |\phi\rangle)$, set $x = x_{h+1}$, and let $\mathcal{B}' = \mathcal{B}_{(x,|\psi_x\rangle)}$. Fix $X \subseteq V \setminus (G \cup \{x\})$ with $1 \le |X| \le k-1$.

Suppose first that $1 \le |X| \le k-2$. Partitioning the extension equations according to whether their supports contain $x$, we obtain, for every $|v\rangle \in \mathcal{H}_X$,

$$\begin{aligned} |v\rangle \in L^{\Sigma}_{\mathcal{B}'}(X) &\iff \begin{cases} |v\rangle \in L^{\Sigma}_{\mathcal{B}}(X), \\ |v\rangle \otimes |\psi_x\rangle \in L^{\Sigma}_{\mathcal{B}}(X \cup \{x\}) \end{cases} \\ &\iff \begin{cases} |v\rangle \in L_{\mathcal{A}_h}(X), \\ |v\rangle \otimes |\psi_x\rangle \in L_{\mathcal{A}_h}(X \cup \{x\}) \end{cases} \\ &\iff |v\rangle \in L_{\mathcal{A}_{h+1}}(X). \end{aligned}$$

Now suppose that $|X| = k-1$. Since every constraint has support of size at most $k$, the only newly activated support is $X \cup \{x\}$. This support has not been deleted: every previously deleted support meets $G$, whereas $X \cup \{x\}$ is disjoint from $G$. Thus the same additional constraint is imposed on the two extension spaces. By the invariant,

$$\begin{aligned} L^{\Sigma}_{\mathcal{B}'}(X) &= \{|v\rangle \in L_{\mathcal{A}_h}(X) : \Pi_{X\cup\{x\}}(|v\rangle \otimes |\psi_x\rangle) = 0\} \\ &= L_{\mathcal{A}_{h+1}}(X). \end{aligned}$$

Hence the equality is preserved after assigning $x_{h+1}$.

Thus the invariant (4) holds at the beginning of each stage and after processing each pair in $U_h$. Its singleton case ensures that every selected state, whether for a pair in $U_h$ or for the next qudit in $F_T$, is compatible with the current assignment. Consequently, each insertion satisfies all newly fully assigned surviving constraints, while previously satisfied constraints remain satisfied. After processing $U_T$, every qudit in $V$ has been assigned. The algorithm therefore returns a global product-state solution of the remaining instance together with the deletion set.

**Step 3: Deletion cost.** In this step, we complete the proof of the lemma by bounding the size of the deletion set.

For each pair $(y, |v_y\rangle) \in U_{-1}$, the number of constraints whose supports contain $y$ is at most $\sum_{r=0}^{k-1} \binom{n-1}{r} \le n^{k-1}$. Here and below, the number of subsets of size at most $p$ of an $a$-element set is at most $(a+1)^p$: encode a subset in increasing order and pad its list to length $p$ with one additional symbol. Thus, processing $U_{-1}$ requires deleting at most $|U_{-1}|n^{k-1}$ constraints. Since $|U_{-1}| = |R_1| - |R_0|$ and $\tau_0 = 1$, we express this bound in the form $(|R_1| - |R_0|)\tau_0 n^{k-1}$.

At stage $h$, when assigning an unassigned qudit $y$, the algorithm deletes constraints with remaining unassigned support $X$ only if $L^{\Sigma}_{\mathcal{B}'}(X) \subsetneq L^{\Sigma}_{\mathcal{B}}(X)$. Here $\mathcal{B}'$ denotes the assignment obtained by extending the current assignment $\mathcal{B}$ to $y$ with the chosen state $|v_y\rangle$. For $|X| = j$, the number of constraints in $D_y(X)$ is at most $\sum_{r=0}^{k-j-1} \binom{n-j-1}{r} \le n^{k-j-1}$. Whenever a strict inclusion occurs, the invariant gives

$$L_{(\mathcal{A}_h)_{(y,|v_y\rangle)}}(X) \subseteq L^{\Sigma}_{\mathcal{B}'}(X) \subsetneq L^{\Sigma}_{\mathcal{B}}(X) = L_{\mathcal{A}_h}(X).$$

Therefore, the number of constraints deleted when assigning $y$ is at most

$$\begin{aligned} &\sum_{j=1}^{k-1} n^{k-j-1} \sum_{X \in \binom{V\setminus(G\cup\{y\})}{j}} \mathbf{1}\{L^{\Sigma}_{\mathcal{B}'}(X) \subsetneq L^{\Sigma}_{\mathcal{B}}(X)\} \\ &\quad \le \sum_{j=1}^{k-1} n^{k-j-1} \sum_{X \in \binom{V\setminus(F_h\cup\{y\})}{j}} \left(\dim L_{\mathcal{A}_h}(X) - \dim L_{(\mathcal{A}_h)_{(y,|v_y\rangle)}}(X)\right) \\ &\quad \le \Delta_{\mathcal{A}_h}(y, |v_y\rangle). \end{aligned} \tag{5}$$

Charge each constraint deleted at initialization to one qudit in its support belonging to $R_1$. Charge every other deleted constraint to the qudit whose insertion causes its deletion, and let $\Sigma_y$ be the set of constraint supports charged to $y$. No constraint is charged to a qudit in $F_T$. The preceding bounds therefore give

$$\begin{aligned}
|\Sigma_{\mathrm{fin}}| &\le \sum_{y \in V \setminus F_T} |\Sigma_y| \\
&\le n^{k-1} \sum_{i=0}^{J} \tau_i \big( |R_{i+1}| - |R_i| \big) \\
&= n^{k-1} \left( \tau_J |R_{J+1}| + \sum_{i=1}^{J} (\tau_{i-1} - \tau_i) |R_i| \right) \\
&= \tau_J (n - T) n^{k-1} + n^{k-1} \sum_{i=1}^{J} \tau_i |R_i| \\
&\le \frac{\varepsilon}{16} n^k + n^{k-1} \sum_{i=1}^{J} \tau_i |R_i|. \qquad (6)
\end{aligned}$$

The second inequality follows by grouping the qudits in $V \setminus F_T$ into the disjoint sets $R_{i+1} \setminus R_i$, $0 \le i \le J$. Each set contains $|R_{i+1}| - |R_i|$ qudits. For $i = 0$, the bound of $\tau_0 n^{k-1}$ deletions per qudit follows from the initialization estimate above. For $1 \le i \le J$, the time $h = t_y$ and state $|v_y\rangle$ chosen in Step 1 satisfy $\Delta_{\mathcal{A}_h}(y, |v_y\rangle) < \tau_i n^{k-1}$, so (5) gives the same bound of $\tau_i n^{k-1}$ deletions per qudit. Summing over all these sets gives the inequality.

Since the instance is $\varepsilon$-far from product-state satisfiability, we have $|\Sigma_{\mathrm{fin}}| \ge \varepsilon n^k$. Combining this with (6) and using $\tau_J \le \varepsilon/16$ gives

$$\sum_{i=1}^{J} \tau_i |R_i(\mathcal{A}_T)| \ge \varepsilon n - \tau_J (n - T) \ge \frac{15 \varepsilon n}{16}.$$

This completes the proof. □

## 4.2 Proof of the simultaneous hitting lemma

We now prove Lemma 3.2, which guarantees that one random sample hits all relevant large bad sets simultaneously. We collect the tools and definitions in Section 4.2.1, obtain a uniform approximation of the potential in Section 4.2.2, and bound the growth functions of the approximate bad sets in Section 4.2.3. We combine these results in Section 4.2.4.

### 4.2.1 Tools and definitions

We begin by introducing some definitions and tools that will be used in the proof.

**Definition 4.1.** For a nonempty finite set $U$ and a range family $\mathcal{R} \subseteq 2^U$, define its growth function by

$$\pi_{\mathcal{R}}(m) := \max_{z_1, \ldots, z_m \in U} \left| \{ (\mathbf{1}_R(z_1), \ldots, \mathbf{1}_R(z_m)) : R \in \mathcal{R} \} \right|.$$

**Lemma 4.2** ([BEHW89, Theorem A2.1])**.** *Let $U$ be a nonempty finite set and let $\mathcal{R} \subseteq 2^U$ be a range family with growth function $\pi_{\mathcal{R}}$. For $0 < \alpha < 1$, let $S$ be a uniformly random $m$-element subset of $U$, where $1 \le m \le |U|$. If $\alpha m \ge 8$, then*

$$\Pr[\exists R \in \mathcal{R} : |R| \ge \alpha |U| \text{ and } S \cap R = \varnothing] \le 2 \pi_{\mathcal{R}}(2m) \exp(-\alpha m / 4).$$

**Definition 4.3** (Pseudo-dimension)**.** For a class $\mathcal{G}$ of real-valued functions on $U$, its pseudo-dimension is the largest integer $m$ for which there are points $z_1, \dots, z_m \in U$ and thresholds $a_1, \dots, a_m \in \mathbb{R}$ such that

$$\{(\mathbf{1}\{g(z_\ell) > a_\ell\})_{\ell=1}^m : g \in \mathcal{G}\} = \{0,1\}^m.$$

**Theorem 4.4** ([GR17, Theorem 3.2])**.** *Let $\mathcal{G}$ have range in $[0, H]$ and pseudo-dimension at most $p$. For any probability distribution $P$ on $U$, independent samples $Z_1, \dots, Z_m$ from $P$ satisfy*

$$\Pr\left[\sup_{g\in\mathcal{G}} \left| \frac{1}{m}\sum_{\ell=1}^m g(Z_\ell) - \mathbb{E}_{Z\sim P} g(Z) \right| \le \eta \right] \ge 1 - \delta$$

*whenever*

$$m \ge A\left(\frac{H}{\eta}\right)^2 \left(p \log \frac{2H}{\eta} + \log \frac{2}{\delta}\right), \qquad 0 < \eta \le H, \quad 0 < \delta < 1,$$

*for a universal constant $A$.*

**Theorem 4.5** ([GJ95, Corollary 2.1])**.** *Let $p_1, \dots, p_s$ be real polynomials of degree at most $D \ge 1$ in $r$ variables, where $s \ge r \ge 1$. Then*

$$\left|\left\{(\operatorname{sgn}(p_1(x)), \dots, \operatorname{sgn}(p_s(x))) : x \in \mathbb{R}^r\right\}\right| \le \left(\frac{8eDs}{r}\right)^r.$$

To obtain a sample bound independent of $n$, we introduce approximate versions of the potential and its drop.

**Definition 4.6** (Approximate potential)**.** For each support $F$, fix a sequence $(X_\ell)_{\ell=1}^L$ of subsets of $V \setminus F$ with $|X_\ell| \le k-1$, depending only on $F$. For a local assignment $\mathcal{A} = (F, |\phi\rangle)$, its approximate potential is

$$\widehat{W}_{\mathcal{A}} := \frac{e^d n^{k-1}}{L} \sum_{\ell : X_\ell \neq \varnothing} d^{-|X_\ell|} \dim L_{\mathcal{A}}(X_\ell).$$

Its approximate drop at $(x, |v\rangle)$ is

$$\widehat{\Delta}_{\mathcal{A}}(x, |v\rangle) := \widehat{W}_{\mathcal{A}} - \widehat{W}_{\mathcal{A}_{(x,|v\rangle)}}.$$

The two potentials in $\widehat{\Delta}_{\mathcal{A}}(x, |v\rangle)$ use the sequences associated with $F$ and $F \cup \{x\}$, respectively.

**Definition 4.7** (Approximate bad sets)**.** Given such sequences and a local assignment $\mathcal{A}_T$ with support $F_T$, let $\mathcal{A}_t$ denote its history at time $t$, for $0 \le t \le T$. For $1 \le i \le J$, define the approximate bad set

$$\widetilde{R}_i(\mathcal{A}_T) = \left\{ y \in V \setminus F_T : \begin{array}{l} \widehat{\Delta}_{\mathcal{A}_t}(y, |v\rangle) \ge (3\tau_i/4) n^{k-1} \\ \text{for every } 0 \le t \le T \text{ and every state } |v\rangle \in L_{\mathcal{A}_t}(\{y\}) \end{array} \right\}.$$

We also record the following rank criterion and a criterion for the existence of nonzero solutions to homogeneous polynomial systems.

**Fact 4.8.** *For a complex matrix $M$ with $q$ columns, set $G = M^*M$. For $1 \le r \le q$, define*

$$E_r(M) = \sum_{\substack{I \subseteq [q] \\ |I| = r}} \det G_{I,I},$$

*where $G_{I,I}$ is the principal submatrix indexed by $I$. We also set $E_0(M) := 1$ and $E_{q+1}(M) := 0$. Then*

$$E_r(M) > 0 \quad \Longleftrightarrow \quad \operatorname{rank} M \ge r \qquad (1 \le r \le q). \tag{7}$$

**Lemma 4.9** ([Laz77, Corollary to Thm. 1])**.** *Given homogeneous polynomials $f_1, \dots, f_s \in \mathbb{C}[v_1, \dots, v_d]$ of degree at most $N \geq 1$, let $\boldsymbol{\gamma}$ be the real vector obtained by listing the real and imaginary parts of their coefficients in a fixed order. Then there exists a nonnegative real polynomial $\Phi$ of degree at most $2\binom{dN}{d-1}$ such that*

$$\Phi(\boldsymbol{\gamma}) = 0 \quad \Longleftrightarrow \quad \exists\, v \in \mathbb{C}^d \setminus \{0\} : \; f_j(v) = 0 \quad (\forall j \in [s]).$$

**Proof Organization** Central to our proof is the representation of extension spaces as kernels of polynomial matrices, established in Lemma 4.10. This allows us to express their dimensions in terms of the ranks of polynomial matrices.

Using this representation, we first construct a uniform approximation of the potential using a number of sampled terms independent of $n$. We use a standard fact from linear algebra (Fact 4.8) to express the dimensions of extension spaces through polynomial signs. By applying the Goldberg–Jerrum theorem (Theorem 4.5) to this representation, we bound the pseudo-dimension of the corresponding function family. We then prove the uniform approximation bound in Lemma 4.11 by combining this bound with the uniform convergence theorem (Theorem 4.4). The approximation is uniform for all local assignments on each fixed support. We use this uniform approximation to construct approximate bad sets $\widetilde{R}_i(\mathcal{A})$ satisfying

$$R_i(\mathcal{A}) \subseteq \widetilde{R}_i(\mathcal{A}) \subseteq R_{i+1}(\mathcal{A}).$$

These inclusions preserve the connection between large bad sets while simplifying the structure of the family of bad sets.

Next, we bound the growth function of the family of approximate bad sets. This requires determining whether a given qudit belongs to an approximate bad set. Using the polynomial matrix representation above, we can answer this question by checking whether a family of polynomial equations has a common nonzero solution. By Lemma 4.9, the existence of such a solution can be determined from the signs of real polynomials. The approximation ensures that the number of polynomials, their degrees, and the number of variables are all independent of $n$. Applying the Goldberg–Jerrum theorem (Theorem 4.5) to these polynomials yields the growth-function bound in Lemma 4.12.

Finally, we apply the sampling bound (Lemma 4.2). We combine this bound with our growth-function bound (Lemma 4.12) and account for all relevant scales and ordered supports contained in the sample. This shows that the sample simultaneously hits all large approximate bad sets. The inclusions above then imply the required hitting property for the original bad sets, proving Lemma 3.2.

### 4.2.2 Uniform approximation of the potential

We first observe that the extension space $L_{\mathcal{A}}(X)$ can be represented as the kernel of a matrix. This representation, established in Lemma 4.10, is central to our proof.

**Lemma 4.10.** *Given a local assignment $\mathcal{A} = (F, |\psi\rangle)$ and any $X \subseteq V \setminus F$ with $1 \leq |X| \leq k-1$, there exists a matrix $M_{\mathcal{A},X}$ with $d^{|X|}$ columns satisfying $L_{\mathcal{A}}(X) = \ker M_{\mathcal{A},X}$. Moreover, write*

$$|\psi\rangle = \bigotimes_{z \in F} |\psi_z\rangle, \qquad |\psi_z\rangle = \sum_{j=0}^{d-1} (a_{z,j} + \mathrm{i} b_{z,j}) |j\rangle, \quad a_{z,j}, b_{z,j} \in \mathbb{R}.$$

*Then each entry of $M_{\mathcal{A},X}$ is a homogeneous polynomial of total degree at most $k$ in the variables $a_{z,j}, b_{z,j}$.*

*Proof.* Recall that a vector $|u\rangle \in L_{\mathcal{A}}(X)$ if and only if $\Pi_S(|u\rangle \otimes |\psi\rangle) = 0$ for every $X \subseteq S \subseteq F \cup X$ with $|S| \leq k$. For each constraint $\Pi_S(|u\rangle \otimes |\psi\rangle) = 0$, view $|u\rangle$ as an unknown vector with $d^{|X|}$ coordinates and write

$$|\psi\rangle = \bigotimes_{z \in F} |\psi_z\rangle, \qquad |\psi_z\rangle = \sum_{j=0}^{d-1} (a_{z,j} + \mathrm{i} b_{z,j})|j\rangle, \quad a_{z,j}, b_{z,j} \in \mathbb{R}.$$

Expanding each constraint $\Pi_S(|u\rangle \otimes |\psi\rangle) = 0$ on $\mathcal{H}_S$ in the computational basis yields a homogeneous linear system in the coordinates of $|u\rangle$, with coefficient matrix $M_{\mathcal{A},X}$. with coefficient matrix $M_{\mathcal{A},X}$. Thus, $L_{\mathcal{A}}(X) = \ker M_{\mathcal{A},X}$. Since each constraint acts on at most $k$ qudits and each assigned qudit contributes linearly through its amplitudes, every matrix entry is a homogeneous polynomial of total degree at most $k$ in the variables $a_{z,j}, b_{z,j}$. □

We use this representation to obtain a uniform approximation of the potential.

**Lemma 4.11.** *For fixed $k, d$, there is a constant $K_{k,d} > 0$ such that, for every support $F$ with $|F| \leq 4e^d/\tau_J + 1$, there is an approximating collection $\{X_\ell\}_{\ell=1}^{L}$ of size*

$$L = \left\lceil K_{k,d} \varepsilon^{-3} \log \frac{2}{\varepsilon} \right\rceil \tag{8}$$

*whose corresponding approximate potential satisfies*

$$\left| \widehat{W}_{\mathcal{A}} - W_{\mathcal{A}} \right| \leq \frac{\varepsilon}{256} n^{k-1} \tag{9}$$

*for every local assignment $\mathcal{A}$ with support $F$.*

*Proof.* Fix a support $F$. Define a probability distribution $P_F$ on $\{X \subseteq V \setminus F : |X| \leq k-1\}$ by

$$P_F(X) := \frac{d^{|X|}}{e^d n^{|X|}} \qquad (\varnothing \neq X \subseteq V \setminus F,\ |X| \leq k-1),$$

with the remaining probability mass assigned to $\varnothing$ and define

$$h_{\mathcal{A}}(X) := \begin{cases} e^d n^{k-1} d^{-|X|} \dim L_{\mathcal{A}}(X), & X \neq \varnothing, \\ 0, & X = \varnothing. \end{cases}$$

Then $\mathbb{E}_{X \sim P_F}[h_{\mathcal{A}}(X)] = W_{\mathcal{A}}$. Let $\mathcal{G}_F = \{h_{\mathcal{A}} : \mathcal{A} \text{ is a local assignment on } F\}$. We next bound the pseudo-dimension of $\mathcal{G}_F$.

Suppose that $\mathrm{Pdim}(\mathcal{G}_F) = m$. By definition, there are points $X_1, \ldots, X_m$ and thresholds $a_1, \ldots, a_m \in \mathbb{R}$ such that

$$\left\{ \left( \mathbf{1}\{g(X_\ell) > a_\ell\} \right)_{\ell=1}^{m} : g \in \mathcal{G}_F \right\} = \{0,1\}^m. \tag{10}$$

For $X \neq \varnothing$, Lemma 4.10 gives $h_{\mathcal{A}}(X) = e^d n^{k-1} d^{-|X|} \big( d^{|X|} - \operatorname{rank} M_{\mathcal{A},X} \big)$. Thus, each comparison $h_{\mathcal{A}}(X) > a$ is determined by $\operatorname{rank} M_{\mathcal{A},X}$. By Fact 4.8, for $0 \leq r \leq d^{|X|} \leq d^{k-1}$, we have

$$\operatorname{rank} M_{\mathcal{A},X} = r \quad \Longleftrightarrow \quad E_r(M_{\mathcal{A},X}) > 0 \ \wedge \ E_{r+1}(M_{\mathcal{A},X}) = 0,$$

Since each entry of $M_{\mathcal{A},X}$ is a polynomial of total degree at most $k$, each entry of $M^*M$ has total degree at most $2k$. Consequently, $E_r(M)$, being a sum of $r \times r$ principal minors of $M^*M$ with

$0 \le r \le d^{|X|}$, is a real polynomial in the $2d|F|$ real variables $a_{z,j}, b_{z,j}$, of total degree at most $2kr \le 2kd^{k-1}$.

To determine $\operatorname{rank} M_{\mathcal{A},X}$, we determine the signs of $E_r(M_{\mathcal{A},X})$ for $0 \le r \le d^{|X|}$. Combining this observation with the preceding bounds, the vector $\left(\mathbf{1}\{g(X_\ell) > a_\ell\}\right)_{\ell=1}^{m}$ is determined by the signs of $md^{k-1}$ real polynomials in $2d|F|$ variables, each of total degree at most $2kd^{k-1}$. Since $m = 0$ when $F = \varnothing$ and $md^{k-1} < 2d|F|$ already implies $m = O_{k,d}(|F|)$, we may assume $md^{k-1} \ge 2d|F| \ge 1$. Applying Theorem 4.5 with $D = 2kd^{k-1}$, $r = 2d|F|$ and $s = md^{k-1}$, we obtain at most

$$\left(\frac{8eDs}{r}\right)^r = \left(\frac{16ekmd^{2k-2}}{2d|F|}\right)^{2d|F|}$$

distinct label vectors. Combining this bound with (10) gives

$$2^m \le \left(\frac{16ekmd^{2k-2}}{2d|F|}\right)^{2d|F|}.$$

Solving this inequality yields $m \le 4d|F|\log_2(16ekd^{2k-2}) = O_{k,d}(|F|)$. Thus,

$$\operatorname{Pdim}(\mathcal{G}_F) = O_{k,d}(|F|) = O_{k,d}(\varepsilon^{-1}).$$

Applying Theorem 4.4 to $\mathcal{G}_F$ with $H = e^d n^{k-1}$, $\eta = \varepsilon n^{k-1}/256$, and $\delta = 1/2$, we obtain that a collection of $L = \left\lceil K_{k,d}\varepsilon^{-3}\log\frac{2}{\varepsilon}\right\rceil$ independent samples from $P_F$, with $K_{k,d}$ sufficiently large, satisfies

$$\left|\widehat{W}_{\mathcal{A}} - W_{\mathcal{A}}\right| \le \frac{\varepsilon}{256}n^{k-1}$$

simultaneously for every local assignment $\mathcal{A}$ on $F$, with probability at least $1/2$. □

### 4.2.3 Growth-function bounds for approximate bad sets

Fix one approximating collection for each support $F$ by Lemma 4.11. Then we have the following key inequality:

$$\left|\Delta_{\mathcal{A}}(x,|v\rangle) - \widehat{\Delta}_{\mathcal{A}}(x,|v\rangle)\right| \le \left|W_{\mathcal{A}} - \widehat{W}_{\mathcal{A}}\right| + \left|W_{\mathcal{A}_{(x,|v\rangle)}} - \widehat{W}_{\mathcal{A}_{(x,|v\rangle)}}\right| \le \frac{\varepsilon}{128}n^{k-1}.$$

Since $\varepsilon/128 < \tau_J/4 \le \tau_i/4$, the inequality yields

$$R_i(\mathcal{A}) \subseteq \widetilde{R}_i(\mathcal{A}) \subseteq R_{i+1}(\mathcal{A}). \tag{11}$$

For an ordered support $F_t = (x_1, \ldots, x_t)$, define the family of approximate bad sets as follows:

$$\mathcal{R}_{i,F_t} := \{\widetilde{R}_i(\mathcal{A}_t) : \mathcal{A}_t \text{ is a local assignment on } F_t\}. \tag{12}$$

We next bound the growth function for the family of approximate bad sets.

**Lemma 4.12** (Growth-function bound for intermediate ranges)**.** *Fix approximating collections of size $L$ as in* (8) *and set $N = Ld^{k-1}$. Then there exists a constant $C_d > 0$ depending only on $d$ such that*

$$\pi_{\mathcal{R}_{i,F_t}}(m) \le (C_d k N^{d+1} m)^{2dt} \tag{13}$$

*for all $i \in [J]$, integers $m \ge 1$ and ordered supports $F_t = (x_1, \ldots, x_t)$ with $0 \le t \le 4e^d/\tau_J$.*

*Proof.* Fix a scale $i$ and an ordered support $F_T = (x_1, \ldots, x_T)$. We divide the proof into the following steps.

**Step 1: Express the approximations by matrix ranks.** For a local assignment $\mathcal{A} = \mathcal{A}_T = (F_T, |\psi\rangle)$ supported on $F_T$, Lemma 4.10 gives

$$\dim L_{\mathcal{A}}(X_\ell) = d^{|X_\ell|} - \operatorname{rank} M_{\mathcal{A},X_\ell} \qquad (X_\ell \neq \varnothing).$$

For each nonempty sample $X_\ell$, set $w_\ell = d^{k-1-|X_\ell|}$. We form the block-diagonal matrix

$$A(|\psi\rangle) = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_L \end{pmatrix},$$

where $A_\ell = I_{d^{k-1}}$ if $X_\ell = \varnothing$, and otherwise

$$A_\ell = \underbrace{\begin{pmatrix} M_{\mathcal{A},X_\ell} & & 0 \\ & \ddots & \\ 0 & & M_{\mathcal{A},X_\ell} \end{pmatrix}}_{w_\ell \text{ diagonal blocks}}.$$

We regard the real and imaginary parts of the amplitudes of the state $|\psi\rangle$ as $2d|F_T|$ real variables. By Lemma 4.10, each entry of $A(|\psi\rangle)$ is a polynomial of total degree at most $k$ in these variables. By construction, $A(|\psi\rangle)$ has $N = Ld^{k-1}$ columns. Therefore,

$$\widehat{W}_{\mathcal{A}} = \frac{e^d n^{k-1}}{L} \sum_{\ell : X_\ell \neq \varnothing} d^{-|X_\ell|} \dim L_{\mathcal{A}}(X_\ell) = \frac{e^d n^{k-1}}{N} \big(N - \operatorname{rank} A(|\psi\rangle)\big).$$

Applying the same construction to the extended local assignment $\mathcal{A}_{(x,|v\rangle)}$ yields a matrix $B(|\psi\rangle, |v\rangle)$ with $N$ columns such that

$$\widehat{W}_{\mathcal{A}_{(x,|v\rangle)}} = \frac{e^d n^{k-1}}{N} \big(N - \operatorname{rank} B(|\psi\rangle, |v\rangle)\big).$$

Each entry of $B(|\psi\rangle, |v\rangle)$ is a multivariate polynomial of total degree at most $k$ in $2d(|F_T| + 1)$ real variables. Therefore,

$$\widehat{\Delta}_{\mathcal{A}}(x, |v\rangle) = \frac{e^d n^{k-1}}{N} \big(\operatorname{rank} B(|\psi\rangle, |v\rangle) - \operatorname{rank} A(|\psi\rangle)\big). \tag{14}$$

**Step 2: Growth bounds via polynomial signs.** Fix a candidate qudit $x \in V \setminus F_T$ and set $\theta := 3\tau_i/4$. By Definition 4.7, $x \notin \widetilde{R}_i(\mathcal{A}_T)$ if and only if there exist a time $t \in \{0, \dots, T\}$ and a state $|v\rangle \in L_{\mathcal{A}_t}(\{x\})$ such that $\widehat{\Delta}_{\mathcal{A}_t}(x, |v\rangle) < \theta n^{k-1}$. Here $\mathcal{A}_t = (F_t, |\psi^{(t)}\rangle)$ is the history of $\mathcal{A}_T$ at time $t$, with $F_t = (x_1, \dots, x_t)$. For each $t$, we construct $A$ and $B$ as in Step 1 using the approximating collections for $F_t$ and $F_t \cup \{x\}$, respectively.

Using (14) and Lemma 4.10, this is equivalent to asking whether there exists a state $|v\rangle$ satisfying

$$M_{\mathcal{A}_t,\{x\}}|v\rangle = 0, \quad \operatorname{rank} B(|\psi^{(t)}\rangle, |v\rangle) < \frac{N\theta}{e^d} + \operatorname{rank} A(|\psi^{(t)}\rangle)$$

for some $t \in \{0, \dots, T\}$.

We now show how the signs of certain polynomials determine whether such a vector $|v\rangle$ exists. Fix a history $\mathcal{A}_t$, suppose that $\operatorname{rank} A(|\psi^{(t)}\rangle) = r$, and set $s = r + \lceil N\theta/e^d \rceil$. For $s \leq N$, the rank criterion gives

$$\operatorname{rank} B(|\psi^{(t)}\rangle, |v\rangle) < \frac{N\theta}{e^d} + r \quad \Longleftrightarrow \quad \det\Big(B(|\psi^{(t)}\rangle, |v\rangle)_{I,K}\Big) = 0 \quad \text{for all } |I| = |K| = s,$$

where $B(|\psi^{(t)}\rangle, |v\rangle)_{I,K}$ denotes the submatrix with rows indexed by $I$ and columns indexed by $K$. If $s > N$, the rank inequality holds automatically since $B$ has $N$ columns, and we impose no determinant conditions.

The vector $|v\rangle$ must also satisfy the compatibility condition $M_{\mathcal{A}_t,\{x\}}|v\rangle = 0$. The determinant conditions above and $M_{\mathcal{A}_t,\{x\}}|v\rangle = 0$ can be viewed as a system of polynomial equations

$$f_j(|\psi^{(t)}\rangle; |v\rangle) = 0, \qquad j = 1, \ldots, q.$$

By construction, each row of $B(|\psi^{(t)}\rangle, |v\rangle)$ is homogeneous in $|v\rangle$ of degree zero or one, so every determinant above is homogeneous. Since the compatibility equations are also homogeneous and linear, all $f_j$ are homogeneous in $|v\rangle$.

We regard the $d$ coordinates $v_1, \ldots, v_d$ of $|v\rangle$ as complex variables and the real and imaginary parts of the local amplitudes in $|\psi^{(t)}\rangle$ as $2dt$ real parameters. The construction involves no complex conjugates of the $v_j$, so the equations form a homogeneous system in $\mathbb{C}[v_1, \ldots, v_d]$ of degree at most $N$, with coefficients polynomial in these real parameters. For each $f_j$ of degree zero in $|v\rangle$, we replace the equation $f_j = 0$ by $f_j v_1 = \cdots = f_j v_d = 0$. This replacement preserves the common nonzero roots, as well as the degree bound $N$ and the degree bound on the coefficients. Thus, the problem is equivalent to asking whether a family of homogeneous polynomials of positive degree at most $N$ has a common root $|v\rangle \neq 0$. By Lemma 4.9, such a solution exists if and only if $P_r(\boldsymbol{\gamma}) = 0$, where $P_r$ is a nonnegative real polynomial in at most $2dt$ variables satisfying

$$\deg P_r \leq 4kN \binom{dN}{d-1} = O_d(kN^d).$$

The degree bound follows because $P_r$ is a polynomial in $\boldsymbol{\gamma}$, the coefficient vector of the $f_j$, and each entry of $\boldsymbol{\gamma}$ has degree at most $2kN$ in the local amplitude parameters. Substituting these expressions therefore introduces an additional factor of $2kN$ in the degree bound.

**Step 3: Bound the growth function.** In Step 2, we showed that for each history $\mathcal{A}_t$, once $\operatorname{rank} A$ is known, the existence of the desired vector is determined by the sign of one real polynomial of degree $O_d(kN^d)$. Determining $\operatorname{rank} A$ requires the signs of $N$ real polynomials, each of degree at most $2kN$. Taking all polynomials $P_r$ with $0 \leq r \leq N$ into account, we therefore need the signs of at most $N + (N + 1) = 2N + 1$ real polynomials.

Since there are $T + 1$ histories and $m$ candidate qudits, the resulting procedure checks the signs of at most $m(T + 1)(2N + 1)$ real polynomials in at most $2dT$ variables, each of degree $O_d(kN^d)$, counting all possible ranks.

Applying Theorem 4.5 to the combined polynomial family over all possible ranks gives, for $T \geq 1$,

$$\pi_{\mathcal{R}_{i,F_T}}(m) \leq \left(C'_d k N^{d+1} m \frac{T+1}{T}\right)^{2dT} \leq \left(C_d k N^{d+1} m\right)^{2dT},$$

where $C'_d$ and $C_d$ depend only on $d$. For $T = 0$, the growth function is at most 1. □

### 4.2.4 Completing the proof

We now combine Lemmas 4.11 and 4.12 with the sampling bound in Lemma 4.2 to prove Lemma 3.2.

*Proof of Lemma 3.2.* Fix a scale $i \in [J]$ and an ordered support $F_t$, where $t = |F_t|$, and consider the corresponding family $\mathcal{R}_{i,F_t}$ of approximate bad sets defined in (12).

Set $\alpha_t = \frac{\varepsilon(t+1)}{128e^d}$, and let $C$ be a uniformly random subset of $V$ with $|C| = c$. Conditional on $F_t \subseteq C$, the remaining sample $C \setminus F_t$ is a uniformly random $m$-subset of $V \setminus F_t$, where $m = c - t$. By Lemma 4.2, the growth-function bound (13), and (11), we have

$$\Pr_C[\exists R \in \mathcal{R}_{i,F_t} : |R| > \alpha_t n,\ C \cap R = \varnothing \mid F_t \subseteq C] \le 2(2C_d k N^{d+1} m)^{2dt} e^{-\alpha_t m/4}$$

whenever $0 < \alpha_t < 1$ and $m \ge 8/\alpha_t$. To determine how large $c$ needs to be, we require the preceding bound to be less than $q_{i,t}$, where

$$q_{i,t} = \frac{1}{8J(4e^d/\tau_i + 1)t!\binom{c}{t}}.$$

Here $J = \lceil \log_2(16/\varepsilon) \rceil$ is as defined in Section 2.1. It suffices to choose $c$ such that

$$2(2C_d k N^{d+1} m)^{2dt} e^{-\alpha_t m/4} \le q_{i,t}.$$

Taking logarithms, it suffices to show that

$$\frac{\alpha_t m}{4} \ge \log \frac{2}{q_{i,t}} + 2dt \log(2C_d k N^{d+1} m).$$

For $c \ge 8e^d/\tau_J$, we have $m = c - t \ge c/2$. Using $t!\binom{c}{t} \le c^t$ and $m \le c$, a sufficient condition is

$$\frac{\varepsilon c(t+1)}{1024e^d} \ge \log\bigl(16J(4e^d/\tau_i + 1)\bigr) + t \log c + 2dt \log(2C_d k N^{d+1} c).$$

Since

$$\log N = O_{k,d}\left(\log \frac{2}{\varepsilon}\right), \qquad \log\bigl(16J(4e^d/\tau_i + 1)\bigr) = O_d\left(\log \frac{2}{\varepsilon}\right),$$

we have

$$D_{k,d}(t+1)\left(\log \frac{2}{\varepsilon} + \log c\right) \ge \log\bigl(16J(4e^d/\tau_i + 1)\bigr) + t \log c + 2dt \log(2C_d k N^{d+1} c),$$

for some constant $D_{k,d} > 0$ depending only on $k, d$. Thus the preceding inequality holds whenever

$$\frac{\varepsilon c(t+1)}{1024e^d} \ge D_{k,d}(t+1)\left(\log \frac{2}{\varepsilon} + \log c\right).$$

This condition is satisfied by

$$c = \left\lceil A_{k,d} \varepsilon^{-1} \log \frac{2}{\varepsilon} \right\rceil = O_{k,d}(\varepsilon^{-1} \log(2/\varepsilon))$$

for a sufficiently large constant $A_{k,d} > 0$ depending only on $k, d$. If $\alpha_t \ge 1$, no range satisfies $|R| > \alpha_t n$; otherwise, this choice of $A_{k,d}$ also ensures $\alpha_t(c - t) \ge \varepsilon c/(256e^d) \ge 8$.

It remains to show that $C$ simultaneously intersects every range $R \in \mathcal{R}_{i,F_t}$ with $|R| > \alpha_t n$ for all scales $i \in [J]$ and all ordered supports $F_t \subseteq C$ with probability at least $7/8$.

For each scale $i$ and ordered support $F_t$, let $E_{i,F_t}$ be the event that $F_t \subseteq C$ and there exists $R \in \mathcal{R}_{i,F_t}$ with $|R| > \alpha_t n$ and $C \cap R = \varnothing$. The conditional bound above gives

$$\Pr_C[E_{i,F_t}] = \Pr_C[F_t \subseteq C] \Pr_C[E_{i,F_t} \mid F_t \subseteq C] \le \frac{\binom{c}{t}}{\binom{n}{t}} q_{i,t}.$$

A union bound over all scales $i \in [J]$ and all ordered supports $F_t$ in $V$ with $0 \le t \le 4e^d/\tau_i$ gives

$$\Pr_C\left[\bigcup_{i,F_t} E_{i,F_t}\right] \le \sum_{i=1}^{J} \sum_{t=0}^{\lfloor 4e^d/\tau_i \rfloor} \sum_{F_t} \Pr_C[E_{i,F_t}].$$

Using the preceding bound on $\Pr_C[E_{i,F_t}]$ and the fact that there are $t!\binom{n}{t}$ ordered supports of size $t$, the last expression is at most

$$\sum_{i=1}^{J} \sum_{t=0}^{\lfloor 4e^d/\tau_i \rfloor} t!\binom{n}{t}\frac{\binom{c}{t}}{\binom{n}{t}} q_{i,t} = \frac{1}{8J}\sum_{i=1}^{J} \frac{\lfloor 4e^d/\tau_i \rfloor + 1}{4e^d/\tau_i + 1} \le \frac{1}{8}.$$

This proves Lemma 3.2, since $7/8 > 3/4$. □

## AI Disclosure

At the start of this work, we asked Codex whether the container method could be used to improve the upper bound in the original [MSV25]. The model initially provided a proof that did not use the container method and gave a bound exponential in $1/\varepsilon$. After further prompting, it improved the bound to $(1/\varepsilon)^3$. We then specifically asked whether the bound could be improved to $(1/\varepsilon)\log(1/\varepsilon)$, and the model provided a complete proof.

We interpreted the AI-generated arguments and rewrote the proof to make it clearer and easier for human readers to follow. We also used AI for grammatical editing.

# A Proof of Corollary 3.5

*Proof of Corollary 3.5.* Our proof consists of the following four steps.

**Step 1: Potential budget.** For an ordered proper partial coloring $\mathcal{A}$ on $F$, the potential sums only over singleton extension spaces, since $k = 2$. Each such space has dimension at most $r$, and hence

$$W_{\mathcal{A}} = \sum_{v \in V \setminus F} \dim L_{\mathcal{A}}(\{v\}) \leq r(n - |F|) \leq rn.$$

We use the bad sets $R_i(\mathcal{A})$ from Definition 2.5. Extending $\mathcal{A}$ at a vertex in $R_i(\mathcal{A})$ with any compatible color decreases the potential by at least $\tau_i n$. Thus, if $m_i$ such extensions are made at scale $i$, telescoping the potential drops gives

$$\sum_{i=1}^{J} m_i \tau_i \leq r. \tag{15}$$

**Step 2: Simultaneous hitting by direct counting.** Set $c = \lceil Kr\varepsilon^{-1} \log(2r^2/\varepsilon) \rceil$, where $K$ is a sufficiently large absolute constant. For $n < c$, inspect the entire graph; hence assume $n \geq c$. View a uniform $c$-vertex sample $C$ as an ordered sample without replacement. Fix a scale $i \in [J]$, an ordered choice of $t \leq c/2$ sample positions, and their colors. Conditional on the corresponding vertices $F$, the remaining $c - t$ vertices are uniform outside $F$. If the resulting partial coloring $\mathcal{A}$ is proper and $|R_i(\mathcal{A})| > \varepsilon n(t+1)/(16r)$, then

$$\Pr_C[C \cap R_i(\mathcal{A}) = \varnothing \mid F] \leq \exp\left(-\frac{\varepsilon c(t+1)}{32r}\right).$$

There are at most $c^t r^t = (cr)^t$ choices of positions and colors. Averaging over $F$ and taking a union bound over these choices and all scales, the probability of missing any such bad set is at most

$$J \sum_{t=0}^{\lfloor c/2 \rfloor} (cr)^t \exp\left(-\frac{\varepsilon c(t+1)}{32r}\right) \leq \frac{1}{4}.$$

Indeed, writing $L = \log(2r^2/\varepsilon)$, we have $\log(cr) = O(L + \log K)$, $J = O(L)$, and $\varepsilon c/(32r) \geq KL/32$. For sufficiently large $K$, the geometric ratio is at most $1/2$ and the prefactor $J \exp(-\varepsilon c/(32r))$ is at most $1/8$. Consequently, with probability at least $3/4$,

$$|R_i(\mathcal{A})| > \frac{\varepsilon n(|\mathcal{A}| + 1)}{16r} \quad \Longrightarrow \quad C \cap R_i(\mathcal{A}) \neq \varnothing \tag{16}$$

holds simultaneously for every $i \in [J]$ and every proper partial coloring $\mathcal{A}$ on $F \subseteq C$ with $|\mathcal{A}| \leq c/2$.

**Step 3: Contradiction for a colorable sample.** Fix a sample satisfying (16) and suppose that $G[C]$ has a proper $r$-coloring. Starting from the empty assignment, process scales $i = 1, \ldots, J$ in order. While $|R_i(\mathcal{A})| > \varepsilon n(|\mathcal{A}| + 1)/(16r)$, choose $x \in C \cap R_i(\mathcal{A})$ and extend $\mathcal{A}$ using its color in the chosen coloring of $G[C]$. This color is compatible, and the extension decreases the potential by at least $\tau_i n$.

To justify each use of (16), the completed extensions give

$$|\mathcal{A}| \tau_J n \leq W_{\mathcal{A}_0} - W_{\mathcal{A}} \leq rn.$$

Since $\tau_J > \varepsilon/32$ and $K$ is sufficiently large, $|\mathcal{A}| < 32r/\varepsilon \le c/2$. Thus the hitting property applies whenever the stopping condition fails, and the potential bound forces termination.

For the resulting extension counts $m_i$, put $t_i = \sum_{j=1}^{i} m_j$. The stopping conditions and monotonicity imply

$$|R_i(\mathcal{A}_{\text{fin}})| \le \frac{\varepsilon n(t_i+1)}{16r} \qquad \text{for every } i \in [J].$$

Using $\sum_i \tau_i \le 1$, $\sum_{i=j}^{J} \tau_i \le 2\tau_j$, and (15), we obtain

$$\begin{aligned}\sum_{i=1}^{J} \tau_i |R_i(\mathcal{A}_{\text{fin}})| &\le \frac{\varepsilon n}{16r} \sum_{i=1}^{J} \tau_i (t_i+1) \\ &\le \frac{\varepsilon n}{16r}\left(1 + 2\sum_{i=1}^{J} m_i \tau_i\right) \le \frac{3\varepsilon n}{16}.\end{aligned}$$

This contradicts the lower bound $15\varepsilon n/16$ in Lemma 3.1. Hence every $\varepsilon$-far graph is rejected with probability at least $3/4$.

**Step 4: Simplifying the sample bound.** A uniformly random $r$-coloring has $|E(G)|/r \le n^2/(2r)$ monochromatic edges in expectation. Some coloring therefore has at most this many monochromatic edges, whose deletion makes $G$ $r$-colorable. An $\varepsilon$-far graph must consequently satisfy $\varepsilon \le 1/(2r)$. In this range,

$$\log \frac{2r^2}{\varepsilon} \le 3 \log \frac{2}{\varepsilon}, \qquad c = O\left(\frac{r}{\varepsilon} \log \frac{2}{\varepsilon}\right).$$

For $\varepsilon > 1/(2r)$, there are no $\varepsilon$-far graphs, and the tester may simply accept without sampling. □

# References


[AE02] Gunnar Andersson and Lars Engebretsen. Property testers for dense constraint satisfaction programs on finite domains. *Random Structures & Algorithms*, 21(1):14–32, 2002.

[AGR26] Marco Aldi, Sevag Gharibian, and Dorian Rudolph. An unholy trinity: TFNP, polynomial systems, and the quantum satisfiability problem. In *17th Innovations in Theoretical Computer Science Conference (ITCS 2026)*, volume 362 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 7:1–7:24. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2026.

[AK02] Noga Alon and Michael Krivelevich. Testing $k$-colorability. *SIAM Journal on Discrete Mathematics*, 15(2):211–227, 2002.

[AS03] Noga Alon and Asaf Shapira. Testing satisfiability. *Journal of Algorithms*, 47(2):87–103, 2003.

[BEHW89] Anselm Blumer, Andrzej Ehrenfeucht, David Haussler, and Manfred K. Warmuth. Learnability and the Vapnik–Chervonenkis dimension. *Journal of the ACM*, 36(4):929–965, 1989.

[Bra11] Sergey Bravyi. Efficient algorithm for a quantum analogue of 2-SAT. In *Cross Disciplinary Advances in Quantum Computing*, volume 536 of *Contemporary Mathematics*, pages 33–48. 2011.

[BS24] Eric Blais and Cameron Seth. New graph and hypergraph container lemmas with applications in property testing. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing (STOC)*, pages 1793–1804, 2024.

[BS25] Eric Blais and Cameron Seth. Testing graph properties with the container method. *SIAM Journal on Computing*, 54(6):1626–1642, 2025.

[GGR98] Oded Goldreich, Shafi Goldwasser, and Dana Ron. Property testing and its connection to learning and approximation. *Journal of the ACM*, 45(4):653–750, 1998.

[GJ95] Paul W. Goldberg and Mark R. Jerrum. Bounding the Vapnik–Chervonenkis dimension of concept classes parameterized by real numbers. *Machine Learning*, 18:131–148, 1995.

[GR17] Rishi Gupta and Tim Roughgarden. A PAC approach to application-specific algorithm selection. *SIAM Journal on Computing*, 46(3):992–1017, 2017.

[KS26] Nick Kushnir and Asaf Shapira. Towards tight bounds for testing $k$-colorability. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2026)*, volume 392 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 35:1–35:13. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2026.

[Laz77] Daniel Lazard. Algèbre linéaire sur $K[X_1, \ldots, X_n]$ et élimination. *Bulletin de la Société Mathématique de France*, 105(2):165–190, 1977.

[LLM+10] C. R. Laumann, A. M. Läuchli, R. Moessner, A. Scardicchio, and S. L. Sondhi. Product, generic, and random generic quantum satisfiability. *Physical Review A*, 81(6):062345, 2010.

[MSV25] Ashley Montanaro, Changpeng Shao, and Dominic Verdon. Testing quantum satisfiability. *Communications in Mathematical Physics*, 406:241, 2025.

[RS96] Ronitt Rubinfeld and Madhu Sudan. Robust characterizations of polynomials with applications to program testing. *SIAM Journal on Computing*, 25(2):252–271, 1996.

[Soh12] Christian Sohler. Almost optimal canonical property testers for satisfiability. In *53rd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 541–550, 2012.