THE FUNDAMENTAL LIMITS OF RECOVERING PLANTED SUBGRAPHS

DANIEL Z. LEE°, FRANCISCO PERNICE°, AMIT RAJARAMAN°, AND ILIAS ZADIK†

ABSTRACT. Given an arbitrary subgraph $H = H_n$ and $p = p_n \in (0, 1)$, the planted subgraph model is defined as follows. A statistician observes the union of the "signal," which is a random "planted" copy H^* of H, together with random noise in the form of an instance of an Erdős–Rényi graph $G(n, p)$. Their goal is to then recover the planted H^* from the observed graph. Our focus in this work is to understand the minimum mean squared error (MMSE), defined in terms of recovering the edges of H^*, as a function of p and H, for sufficiently large n.

A recent paper [MNS+23] characterizes the graphs for which the limiting (as n grows) MMSE curve undergoes a sharp phase transition from 0 to 1 as p increases, a behavior known as the all-or-nothing phenomenon, up to a mild density assumption on H. However, their techniques fail to describe the MMSE curves for graphs that do not display such a sharp phase transition. In this paper, we provide a formula for the limiting MMSE curve for *any* graph $H = H_n$, up to the same mild density assumption. This curve is expressed in terms of a variational formula over pairs of subgraphs of H, and is inspired by the celebrated subgraph expectation thresholds from the probabilistic combinatorics literature [KK07]. Furthermore, we give a polynomial-time description of the optimizers of this variational problem. This allows one to efficiently approximately compute the MMSE curve for any dense graph H when n is large enough. The proof relies on a novel graph decomposition of H as well as a new minimax theorem which may be of independent interest.

Our results generalize to the setting of minimax rates of recovering arbitrary monotone boolean properties planted in random noise, where the statistician observes the union of a planted minimal element $A \subseteq [N]$ of a monotone property and a random $\mathrm{Ber}(p)^{\otimes N}$ vector. In this setting, we provide a variational formula inspired by the so-called "fractional" expectation threshold [Tal10], again describing the MMSE curve (in this case up to a multiplicative constant) for large enough n.

°CSAIL, Massachusetts Institute of Technology; †Department of Statistics and Data Science, Yale University. Email: {lee_d,fpernice,amit_r}@mit.edu, ilias.zadik@yale.edu.

Contents

1. Introduction

Over the past two decades, several influential papers have addressed the limiting (normalized) minimum mean squared error (MMSE) in high-dimensional statistical models, including [DMK$^+$16, LM17, MT06, RP19] among others. In this paper, we focus on understanding the limiting MMSE in the fundamental problem of recovering planted discrete structures from noisy random environments.

A motivating example: planted clique.   A very well-studied discrete high-dimensional model is the planted clique model, initially defined independently by Jerrum [Jer92] and Kučera [Kuč95]. To define it, we need two parameters, $k = k_n \in \mathbb{N}$ and $p = p_n \in (0, 1)$. We assume that the statistician observes the union of a "signal," which is a k-clique chosen uniformly at random from the complete graph K_n, and the "noise," which is an instance of an Erdős–Rényi graph $G(n, p)$. The goal of the statistician is to recover the planted clique from the observed graph. It should be noted that there is a large body of work dedicated to understanding the special case where $p = 1/2$ and $k = O(\sqrt{n})$. However, it is beneficial for us here to consider the more general case where $k = k_n$ and $p = p_n$ can scale arbitrarily with n. This generality has recently proven useful in circuit lower bounds [GMZ23] and low-degree lower bounds [YZZ24], among others.

The planted clique model has been studied extensively due to its conjectured computational hardness in certain regimes, and has interesting implications for other high-dimensional statistical models (see, e.g., [BBH18] and references therein). From a purely statistical perspective, [MNS$^+$23] recently computed the limiting MMSE of the model as n grows to infinity, proving that it converges to a step function, a behavior known as the All-or-Nothing (AoN) phenomenon (initially discovered in the context of Gaussian sparse regression, first conjectured to hold in [GZ22] and proven in [RXZ20]).

The planted subgraph model.   Motivated by the extensive study of models similar to the planted clique model, such as the planted matching model [MMX21, DWXY23] and the planted dense subgraph model [DMW25], the authors of [MNS$^+$23] initiated the study of the limiting MMSE for general planted subgraphs (also independently introduced in [Hul22] and [ABI$^+$23] by the cryptography community). To define this model, for each $n \in \mathbb{N}$, fix *any* (unlabeled) subgraph $H = H_n$ of the complete graph K_n and some $p = p_n \in (0, 1)$. Next, draw a uniformly random copy H^* of H in the complete graph. The statistician observes the graph G on n vertices which is the edge union of the "signal" H^* with "noise," which is an instance of $G_0 \sim G(n, p)$, i.e., $G = H^* \cup G_0$. It is easy to verify that this more general model subsumes the problems of planted clique and planted perfect matching as special cases.

Following [MNS$^+$23], the main object of interest for us is the (normalized) MMSE of the model, defined as

$$\mathrm{MMSE}_n(p) = \frac{1}{|H|} \min_{\hat{A}} \mathbb{E}\left[\|\mathbf{1}_{H^*} - \hat{A}(G)\|_2^2 \right] = \frac{1}{|H|} \mathbb{E}\left[\|\mathbf{1}_{H^*} - \mathbb{E}[\mathbf{1}_{H^*}|G]\|_2^2 \right], \qquad (1.1)$$

where by $\mathbf{1}_{H*}$ we refer the indicator vector of the edges of H^*, $\mathbf{1}_{H*} \in \{0, 1\}^{\binom{n}{2}}$ [1]. Observe that $\mathrm{MMSE}_n(p) \in [0, 1]$, since the trivial estimator $\hat{A} = 0$ attains mean squared error 1.

The main result of [MNS$^+$23] is that for any "weakly dense"[2] $H = H_n$, the planted subgraph model for $H = H_n$ satisfies the AoN phenomenon (i.e., the MMSE converges to a step function) if and only if the graph H is balanced, in that $\max_{J \subseteq H} \frac{|J|}{v(J)} = \frac{|H|}{v(H)}$.[3] The main goal of this paper is to go beyond the AoN phenomenon, and characterize the MMSE curve for an *arbitrary* weakly dense H, for sufficiently large n.

Kahn–Kalai thresholds and Bayesian approaches. An easy application of Bayes' rule shows that the posterior distribution of the planted subgraph model is the uniform distribution among all the copies of $H = H_n$ in the observed graph $G = H^* \cup G_0$, $G_0 \sim G(n, p)$. As a result, to understand the MMSE one needs to accurately calculate appropriate subgraph counts in G; specifically, those of the copies of H in G that have a given overlap with the planted copy H^*.

Because the planted graph $H = H_n$ may be arbitrary, naïve first and second moment method arguments fail to provide good estimates for these subgraph counts. To circumvent this issue, we derive inspiration from probabilistic combinatorics in the study of the "null" model $G(n, p)$, specifically the "Kahn–Kalai" conjectures on the tightness of expectation thresholds [KK07, Tal10, FKNP21, PP24].

The expectation thresholds relate to the following question about subgraph counts: given an (unlabeled) graph H, at what critical value of $p = p_c(H)$ does a copy of H begin to occur in $G(n, p)$ with high probability?[4] This has been studied for specific choices of H as far back as the original works of Erdős and Rényi [ER59, ER60], and has spurred a rich body of work. A series of conjectures by Kahn and Kalai, later refined by Talagrand, provide simple formulas that are conjectured to yield the critical threshold p_c up to a multiplicative $O(\log |H|)$ factor.

In particular, a conjecture of Kahn and Kalai, which remains unproven, asserts that the critical threshold for any H is essentially given by its *subgraph expectation threshold*, $\max_{S \subseteq H} p_{1M}(S)$, up to a multiplicative $O(\log |H|)$ factor. Here, $p_{1M}(S)$ is the first moment threshold for the appearance of the subgraph S. A straightforward first moment calculation implies that $p_{1M}(S)$ is approximately $n^{-v(S)/|S|}$ for weakly dense S. More sophisticated first moment methods result in the *expectation threshold* [KK07] and the *fractional expectation* threshold [Tal10], which are now known to yield the critical threshold up to a multiplicative $O(\log |H|)$ due to the breakthrough works [FKNP21, PP24]. It should also be noted that these results generalize well beyond subgraph inclusion properties, and identify the critical thresholds for arbitrary monotone properties on a random discrete universe of points under the product Bernoulli measure.

---

[1]Throughout the paper we identify a graph H with its edge set. Hence, by $J \subseteq H$ we refer to the edge-induced subgraph J of H. Moreover, note that the MMSE (1.1) in the planted subgraph model is measured in terms of number of edges recovered.

[2]This result concerns graphs H with $|H|$ edges and $v(H)$ vertices that satisfy $|H| = \omega(v(H) \log v(H))$.

[3]Technically, [MNS$^+$23] proves this equivalence for an asymptotic definition of balancedness, but we omit the details for the introduction.

[4]More concretely, what is the value of p_c such that $\mathbb{P}_{G \sim G(n, p_c)}$ [a copy of H appears in G] $= 1/2$?

While the original proof of tightness of the fractional expectation threshold in [FKNP21] proceeds by careful combinatorial arguments, [MNSZ22b] provides an alternate Bayesian proof using an approach similar to that employed in [MNS$^+$23] to prove the AoN characterization of the MMSE in the planted subgraph model. In fact, [MNSZ22b] proves the tightness of the fractional expectation threshold for any H, by bounding the limiting MMSE of the planted subgraph model corresponding to H from below. This leads to the following question:

Can one go the other direction, and use the expectation thresholds to characterize the MMSE for a large family of subgraph models?

Indeed, in this work, we prove that for large n, the MMSE curve is approximately given by a piecewise-constant function, whose discontinuity points are precisely modifications of the (fractional/subgraph) expectation thresholds of H.

1.1. **Main Contributions.**

As mentioned, our main result is a characterization of the MMSE for the planted subgraph model for all weakly dense graphs $H = H_n$ (for large enough n). To present the theorem, recall, as mentioned above, that the subgraph expectation threshold of any weakly dense H from [KK07] is (approximately) given by $\max_{J \subseteq H} n^{-v(J)/|J|}$ (see Lemma A.3 for an exact statement and a proof of this simple fact). Motivated by this, we define, for any $q \in [0, 1)$, the q-modified subgraph expectation threshold of H:

$$\varphi_q = (\varphi_q)_n = \min \left\{ \max \left\{ n^{-\frac{|V(J)\setminus V(S)|}{|J \setminus S|}} : J \subseteq H, J \supsetneq S \right\} : S \subseteq H, |S| \leqslant q|H| \right\}. \qquad (1.2)$$

Also set $\varphi_1 := 0$. In words, we define φ_q as the minimum possible subgraph expectation threshold of $H \setminus S$ among all choices of S with at most $q|H|$ edges.[5]

Why are the thresholds $(\varphi_q)_{q \in [0,1]}$ potentially relevant for the MMSE curve? Assuming the validity of the subgraph version of the Kahn–Kalai conjecture, φ_q approximately corresponds to the critical density p such that for some $S \subseteq H$ with at most $q|H|$ edges, the graph $H \setminus S$ appears in $G(n, p)$ with high probability. In particular, in the planted model, this is precisely the threshold at which a copy of H appears in the noise that intersects H^* at some subgraph S with at most $q|H|$ edges. This suggests that φ_q could be the threshold such that, as soon as $p > \varphi_q$, it is impossible to recover more than a q fraction of the edges of H^* because another copy of H appears in G with $|H \cap H^*| \leqslant q|H|$. Then, using standard Bayesian techniques, the above can be equivalently phrased as $p = \varphi_q$ being the threshold at which the MMSE hits $1 - q$.

However, this line of argument cannot be applied directly to conclude the desired result. Perhaps the most crucial issue is that the subgraph version of the Kahn–Kalai conjecture remains open for a growing $H = H_n$ (even restricted to weakly dense H); moreover, even if proved, it is only expected to approximate $p_c(H)$ to within a multiplicative $O(\log|H|)$ factor of the true critical threshold. Despite such barriers, we are able to prove that the threshold $(\varphi_q)_{q \in [0,1]}$ indeed characterize (with no multiplicative slack) the points that the MMSE crosses the values $(1 - q)_{q \in [0,1]}$ in the planted subgraph model, confirming the above intuition.

More specifically, our main result can be informally stated as follows.

---

[5]Here, by $H \setminus S$, we refer to the "graph cut" constructed after we delete all edges from S from H: see Definition 2.2 for a precise statement.

Theorem 1.1 (Informal version of Theorem 3.1 and Theorem 3.6(c))**.** *Let $\varepsilon > 0$, and $H = H_n$ an arbitrary weakly dense graph. Then, for sufficiently large n, there exists an integer $1 \leqslant M \leqslant |H|$ and thresholds $1 = q_M > \cdots > q_1 > q_0 = 0$, such that the following holds for the planted subgraph model corresponding to H.*

- *For all $i = 0, 1, \ldots, M - 1$, if $p \in ((1 + \varepsilon)\varphi_{q_{i+1}}, (1 - \varepsilon)\varphi_{q_i})$, it holds that*
$$\mathrm{MMSE}_n(p) = 1 - q_{i+1} + o(1).$$

    For $p \geqslant (1 + \varepsilon)\varphi_{q_0}$, $\mathrm{MMSE}_n(p) = 1 - o(1)$.
- *The "discontinuity" points $(\varphi_{q_i})_{i=1,2,\ldots,M}$ can be computed in $\mathrm{poly}(n)$ time.*

Our proof is in fact able to provide an intuitive characterization of the "critical" subgraphs of H^* that are recoverable from G as one gradually decreases the noise level p from one to zero, revealing that they are in fact highly structured. The first subgraph of H^* that is recoverable is (perhaps naturally) the densest one, i.e., the subgraph J^1 attaining the maximum $\max_{J \subseteq H} |J|/v(J)$, which happens at the critical threshold $p = \varphi_0 = n^{-v(J^1)/|J^1|}$. This results in the first "jump" of the MMSE from 1 to $1 - q_1$ where $q_1 = \frac{|J^1|}{|H|}$; in particular, if $p < \varphi_{q_1}$, a q_1-fraction of the edges of H (corresponding to the subset $J^{(1)} := J^1$) can be recovered. As we continue decreasing p, the next critical subgraph is the densest one after deleting the graph J^1 from H. This produces the next jump in the MMSE at critical probability $p = \varphi_{q_1} = \max_{J \subseteq H \backslash J^1} n^{-v(J)/|J|} = n^{-v(J^2)/|J^2|}$ from $1 - q_1$ to $1 - q_2$ for $q_2 = \frac{|J^2| + |J^1|}{|H|}$; in particular, if $p < \varphi_{q_2}$, a q_2-fraction of the edges of H (corresponding to $J^{(2)} := J^1 \cup J^2$) can be recovered. Continuing this "peeling process" of H produces a sequence of these critical subgraphs $J^{(i)}, i = 1, 2, \ldots, M$ which we define later as the "onion decomposition" of H (see Figure 1 for a pictorial depiction of these thresholds, and Definition 3.3 for more precise details). We direct the reader to Corollary 3.8 and the discussion after it for more precise statements.

Our proof of Theorem 1.1, in addition to employing Bayesian statistical tools such as the planting trick [ACO08], is combinatorial in nature and involves a novel minimax argument (see Lemma 6.3)—we are able to show that the min and max in the definition of φ_q can be swapped, which is crucial in establishing our main result and to the best of our knowledge not implied by any standard minimax result. We believe that the onion decomposition, and the aforementioned minimax result we prove with it, may be of independent interest. We also remark that we give an efficient algorithm to compute the onion decomposition of any graph (see Theorem 3.6), and hence its approximate MMSE curve for large n. A pictorial depiction of this decomposition, and its relation to the MMSE, is shown in Figure 1.

Although we do not prove the subgraph Kahn–Kalai conjecture (also called the "second" Kahn–Kalai conjecture in [MNSZ22a]), our work gives a clear *statistical meaning* to (variants of) the subgraph expectation thresholds from probabilistic combinatorics. For example, our main result implies that as long as $q_1 = \frac{|J^1|}{|H|} = \Omega(1)$ for $J^{(1)}$ the densest subgraph of H, the original subgraph expectation threshold of H, $\max_{J \subseteq H} n^{-v(J)/|J|}$, is exactly (without any logarithmic multiplicative error!) the critical noise level p at which the MMSE jumps from being trivial (equal to 1), to a non-trivial value (i.e., a value strictly less than 1). The location of the noise level at which the MMSE ceases to be trivial is a well-studied object in the literature known as the "weak recovery"
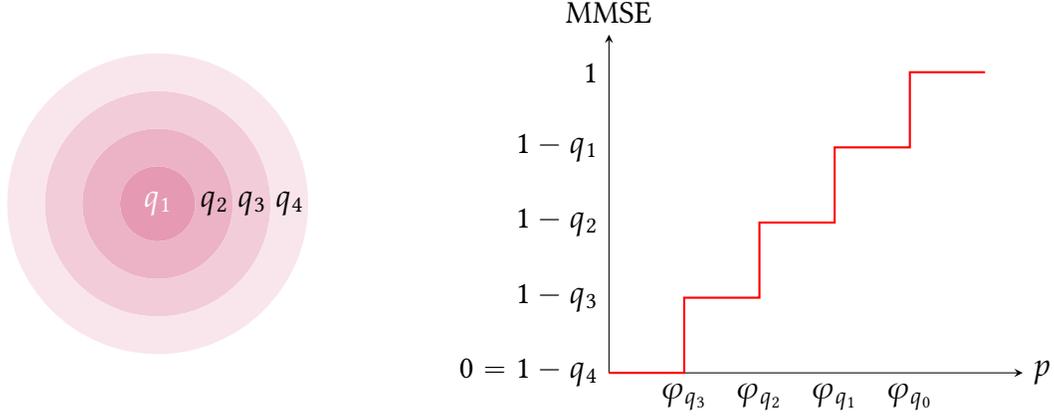
FIGURE 1. A pictorial representation of the onion decomposition of a graph (left) and the corresponding MMSE curve (right). Each q_i represents the cumulative fraction of the graph included in all layers starting from the center till the ith layer; hence $q_4 = 1$. The critical thresholds described above are denoted by φ, following the notation in our theorem statement below.

threshold of H in the planted subgraph model and there has been an intense study of the weak recovery thresholds for various planted models in the statistical physics, information theory and statistics communities (see e.g., [COKPZ17] and references therein).

Let us now define the weak recovery threshold in our context.

Definition 1.2. For any $H = H_n$ and $\varepsilon \in (0, 1)$ we define the ε-recovery recovery threshold $p_\varepsilon = p_\varepsilon(H)$ as

$$p_\varepsilon = p_\varepsilon(H_n) := \sup\{p \in [0, 1] : \mathrm{MMSE}_n(p) \leqslant 1 - \varepsilon\}.$$

The planted subgraph model for H is said to have a *(sharp) weak recovery threshold* if for some $\varepsilon_0 > 0$, for all $0 < \varepsilon < \varepsilon_0$, if n is sufficiently large, it holds that

$$p_\varepsilon = (1 + o(1))p_{\varepsilon_0},$$

that is, the MMSE "jumps" at the value p_{ε_0} from $1 - o(1)$ to $1 - \Omega(1)$. In that case, we call the sharp weak recovery threshold of $H = H_n$ to be

$$p_{\mathrm{WR}} = p_{\mathrm{WR}}(H) := p_{\varepsilon_0}.$$

We state this result in the following informal corollary.

Corollary 1.3 (Informal, based on Theorem 3.6 and Corollary 3.8). *Let $\varepsilon > 0$, and $H = H_n$ an arbitrary weakly dense graph. Suppose that for $J^{(1)} = \arg\max_{J \subseteq H} |J|/v(J)$ it holds that $q_1 = |J_1|/|H| = \Omega(1)$; a property referred to as H being delocalized in [MNS$^+$23, Definition 4.3]. Then, the planted subgraph model for H has a sharp weak recovery threshold. Moreover, this threshold $p_{\mathrm{WR}}(H)$ is equal to the subgraph expectation threshold of H, that is,*

$$p_{\mathrm{WR}}(H) = (1 + o(1)) \max_{J \subseteq H} n^{-v(J)/|J|}.$$

Notably, the weak recovery threshold has also been called "the condensation threshold" of the model in the literature of statistical physics [COKPZ17], due to an underlying I-MMSE relation linking the discontinuities of the MMSE and the first-order discontinuities of the free energy of the model. Interestingly, no such I-MMSE relation is currently known for the planted subgraph model for all H, despite multiple partial results for specific choices of H. To our knowledge, the best such general result is a one-sided I-MMSE relation proven in [MNS+23, Lemma 5.9].

Beyond (weakly dense) graphs $H = H_n$: planting arbitrary monotone properties. Recall that the theory of *fractional* expectation thresholds has been established in [FKNP21] (as conjectured in [Tal10]) to correctly predict the critical threshold p for *any* H up to a multiplicative logarithmic factor. More generally, it predicts the critical threshold p at which any monotone property is satisfied in an instance of Bernoulli$(p)^{\otimes N}$ (again, up to a multiplicative logarithmic factor). Given this, it is natural to ask whether our result can be generalized to understand the limiting MMSE curve when planting arbitrary monotone properties in Bernoulli$(p)^{\otimes N}$ noise.

We succeed in doing so, albeit at the cost of a constant multiplicative loss in the threshold. Our results hold under a worst-case prior over the planted element of the property, a standard setting known as the study of minimax estimation rates in statistics. As a special case, our result applies to the planted subgraph model for *any* subgraph $H = H_n$ and yields the order of the density p under which the MMSE of the planted subgraph model for H crosses any specific value in $(0, 1)$. We direct the reader to Section 4 for details on this general case.

The proof of this result is based on a modification of the fractional expectation thresholds (analogous to how φ_q is a modification of the subgraph expectation threshold), a strong duality trick first observed by Talagrand [Tal10], and finally a modification of the Bayesian proof of the spread lemma in [MNS+23]. We refer the reader to Theorem 4.5 for a precise statement of this result.

Related work. The limiting MMSE of various high dimensional statistical models has been analyzed in a number of recent influential works. Notable characterizations of the limiting MMSE include the exact computation for the spiked rank-one matrix model under Gaussian noise [DMK+16, LM17] (conjectured in [LKZ15]) and the random linear estimation problem with a Gaussian feature matrix [MT06, RP19] (conjectured in [Tan02, GV05]). In many cases, the approach begins with a non-rigorous yet remarkably precise calculation of the model's free energy via the "replica" or "cavity" methods, which also yields a conjectured exact formula for the limiting MMSE. In some cases, such as those mentioned above, the validity of these conjectured formulas has later been rigorously proven. However, our techniques fundamentally differ from those used here, and operate in a setting where typical assumptions in statistical physics literature do not hold. For instance, for our main result Theorem 1.1, the planted subgraph $H = H_n$ is only assumed to be weakly dense for each n, and in particular no product-like structure is assumed on its entries as is often the case for statistical physics techniques. Moreover, for our general results on monotone properties (Theorem 4.5), no assumption whatsoever is made on the planted discrete structure other than it being of a known cardinality.

In another related line of work, the goal is to prove properties of the limiting MMSE curve for specific planted models. Notable examples include [RXZ20] that understood the limiting MMSE

curve in sparse linear regression and proved the AoN phenomenon, and [DWXY23, GSXY25] that studied the MMSE for the planted matching problem. Our work again differs from these results in that it applies to a larger family of priors with much less structure—for instance, the planted matching model is a special case of our planted subgraph model with H a perfect matching so Theorem 4.5 is directly applicable (as the perfect matching is not a weakly dense graph, we cannot apply the stronger Theorem 3.1). Moreover, our proof crucially differs in the proof techniques we employ to the previous works. Most of the above works are based on involved second moment methods (e.g., in [RXZ20]) or algorithmic construction arguments (e.g., in [DWXY23]). Yet, in this work we analyze the limiting MMSE curves by leveraging Bayesian techniques, such as the planting trick [ACO08], alongside state-of-the-art combinatorics results and arguments using the structure of the (fractional) expectation thresholds, such as the duality trick of Talagrand [Tal10]. We hope that the combinatorics techniques introduced in this work will be useful in understanding more minimax rates or MMSE curves in the high-dimensional statistics literature.

2. Getting started

We now define several notions that will be useful in presenting and proving our main results below.

2.1. Graph Theory.
We start by introducing a few definitions from graph theory.

First, as mentioned in the introduction, in this paper we identify a graph T with its edge set $T \subseteq \binom{[n]}{2}$. Moreover, its vertex set $V(T)$ is the set of all vertices in K_n adjacent to at least one edge of T. Hence, we use the notations $|T|$ to refer to the number of edges in T and $v(T) := |V(T)|$ to refer to the number of vertices in T. Finally, we define it's *density* as $\rho(T) := |T|/v(T)$.

We now define the notion of a weakly dense graph.

Definition 2.1. Consider a sequence of graphs $H = H_n, n \in \mathbb{N}$. We say that a graph $H = H_n$ is weakly dense if

$$\liminf_n \frac{|H_n|}{v(H_n) \log v(H_n)} = +\infty.$$

We often simplify notation in this work and refer to $H = H_n$ as a graph, instead of a sequence of graphs indexed by n.

Now, we introduce a non-standard definition in graph theory, which we term a "graph-cut," which can intuitively be viewed as a graph "modulo" a specific subset of its vertices. This definition will be useful in understanding the properties of the onion decomposition of a graph H (see Section 5).

Definition 2.2. A "graph-cut" on n vertices is defined by a triplet $J = (V, S, E)$, where $S \subseteq V \subseteq [n]$, and

$$E \subseteq K_V \backslash K_S = \{\{u, v\} : u, v \in V \text{ and at most one of } u, v \text{ is in } S\}.$$

We further denote its number of edges $|J| := |E|$, its number of vertices $v(J) := |V \backslash S|$, and if $S \subsetneq V$, its density

$$\rho(J) := \frac{|J|}{v(J)} = \frac{|E|}{|V \backslash S|}. \tag{2.1}$$

If $V \backslash S = \varnothing$, we define $\rho(J) := +\infty$, using the convention that $\frac{1}{\rho(J)} = \frac{1}{\infty} = 0$. Note that any graph can trivially be considered a graph-cut by choosing $S = \varnothing$.

Towards succinctness, given a graph T_1 and subgraph $T_2 \subseteq T_1$, we use $T_2|T_1$ to denote the graph-cut $(V(T_1), V(T_2), T_1 \backslash T_2)$. In particular,

$$\rho(T_1|T_2) := \frac{|T_1| - |T_2|}{|V(T_1) \backslash V(T_2)|}. \tag{2.2}$$

2.2. Monotone Properties and Expectation Thresholds.

Now we turn to some required background from combinatorics. Let $\mathcal{X}$ be a finite universe of points. We start with defining a monotone property.

Definition 2.3. A monotone property $\tilde{A}$ in the finite universe $\mathcal{X}$ is a collection of subsets of $\mathcal{X}$, such that for any $A \in \tilde{A}$ if $A \subseteq B$ then $B \in \tilde{A}$. An element $A \in \tilde{A}$ is called a minimal element of $\tilde{A}$ if there is no $A' \in \tilde{A}$ with $A' \subsetneq A$.

Observe that a monotone property $\tilde{A}$ is fully characterized by the set of minimal elements of $\tilde{A}$. We can define the fractional expectation threshold of a monotone property as in [Tal10].

Definition 2.4 (Fractional expectation threshold). Fix any monotone property $\tilde{A}$ in a finite universe $\mathcal{X}$. We define

$$p_{\mathrm{FE}}(\tilde{A}) := \max_{w \in W} \left\{ \mathcal{E}_w^{-1}(1/2) \right\},$$

where $\mathcal{E}_w(p) := \sum_{T \subseteq \mathcal{X}} w(T) p^{|T|}$, and W is the set of all "fractional covers" w which are functions $w : 2^{\mathcal{X}} \to [0, 1]$ such that for any minimal element A of $\tilde{A}$, $\sum_{T \subseteq A} w(T) \geqslant 1$.

It turns out that that the fractional expectation threshold is a natural lower bound for the critical probability $p_c(\tilde{A})$ (via a relatively simple "fractional" union bound argument) for the property of interest to be satisfied under the product $\mathrm{Ber}(p)$ measure on $\mathcal{X}$ [Tal10]. Importantly, because of its fractional nature, a strong duality trick observed by Talagrand in [Tal10] leads to a representation of $p_{\mathrm{FE}}(\tilde{A})$ in terms of so-called "spread" measures on the minimal elements of $\tilde{A}$ [ALWZ20]. Then, the breakthrough work [FKNP21] used this representation and proved that $p_{\mathrm{FE}}(\tilde{A})$ also upper bounds $p_c(\tilde{A})$, up to a logarithmic multiplicative factor. As the critical probability $p_c(\tilde{A})$ is not the focus of this work, we refer the interested reader to [FKNP21] for more details. We only highlight that the same duality trick is in fact crucial in our work in order to obtain our general connection between minimax rates and variants of the fractional expectation thresholds in Section 7.

3. Main Results I: Planting a Weakly Dense H

3.1. Characterizing the MMSE curve via the φ_q thresholds.

We start with our first and main result in this section. For any $q \in [0, 1)$, we identify the noise level p (up to $1 + o(1)$ error) at which the MMSE curve crosses the threshold $1 - q$. Recall the definition of φ_q from (1.2).

Theorem 3.1. *Consider any (sequence of) weakly dense $H = H_n$ and $q \in [0, 1)$. Then, for the planted subgraph model corresponding to H, the following holds as n grows to infinity.*

(a) If $\liminf_n p/\varphi_q > 1$, *then*

$$\liminf_n \mathrm{MMSE}_n(p) \geqslant 1 - q.$$

(b) If $\limsup_n p/\varphi_q < 1$, *then*

$$\limsup_n \mathrm{MMSE}_n(p) \leqslant 1 - q.$$

In particular, for any constant $q \in [0, 1]$*, there exists some sequence of noise levels* $p = p_n$ *satisfying*

$$\lim_n p/\varphi_q = 1$$

for which it holds

$$\lim_n \mathrm{MMSE}_n(p) = 1 - q.$$

The proof of the theorem is deferred to Section 6.

Remark 3.2 (The MMSE curve using the discontinuities of φ_q). Theorem 3.1 can be used to extract non-trivial structure for the $\mathrm{MMSE}_n(p)$ curve from the behavior of the points $(\varphi_q)_{q \in [0,1)}$. Observe that φ_q is a piecewise constant function of q, since for any $\varepsilon = \varepsilon(H) > 0$ small enough, $\varphi_q = \varphi_{q+\varepsilon}$. For example, this holds if $\lfloor (q + \varepsilon)|H| \rfloor = \lfloor q|H| \rfloor$ since in both these cases the same collection of subsets $S \subseteq H$ is considered inside the minimum operation in the definition of φ_q.

This implies that one can choose the minimal possible integer $0 \leqslant M \leqslant |H|$ for which there exist points $1 = q_M > q_{M-1} > \ldots > q_1 > q_0 = 0$ such that φ_q is constant on the intervals $[q_{i+1}, q_i), i = 0, 1, \ldots, M - 1$. In terms of the MMSE, Theorem 3.1 implies that if for some $i = 0, 1, \ldots, M - 1$ we have $\varphi_{q_i} = (1 + \Omega(1))\varphi_{q_{i+1}}$, then for any $\delta > 0$ if p satisfies

$$(1 + \delta)\varphi_{q_{i+1}} \leqslant p \leqslant (1 - \delta)\varphi_{q_i}$$

then

$$\mathrm{MMSE}_n(p) = 1 - q_{i+1} + o(1),$$

and if $p \geqslant (1 + \delta)\varphi_{q_0}$, $\mathrm{MMSE}_n(p) = 1 - o(1)$.

Indeed, if $p \leqslant (1 - \delta)\varphi_{q_i}$ then for any $\varepsilon > 0$ we have $p \leqslant (1 - \delta)\varphi_{q_{i+1}-\varepsilon}$ and therefore $\mathrm{MMSE}_n(p) \leqslant 1 - q_{i+1} + \varepsilon + o(1)$ by Theorem 3.1(b). Since $\varepsilon > 0$ can be made arbitrarly small, we conclude $\mathrm{MMSE}_n(p) \leqslant 1 - q_{i+1} + o(1)$. For the other direction, we have $(1 + \delta)\varphi_{q_{i+1}} \leqslant p$ and thus $\mathrm{MMSE}_n(p) \geqslant 1 - q_{i+1} - o(1)$ again by Theorem 3.1(a) and the fact that $\mathrm{MMSE}_n(p)$ is a non-increasing function of p. The result follows. The proof when $p \geqslant (1 + \delta)\varphi_{q_0}$ is analogous.

Based on Remark 3.2, to further understand the MMSE curve one needs to understand the discontinuity points of the piecewise constant function, (φ_q). At first sight, φ_q is a minimax problem over exponentially many subgraphs, suggesting that such an analysis would be very challenging. Despite this, our next result is a full combinatorial characterization of φ_q based on a peeling process which we term the *onion decomposition* of H.

3.2. **Simplifying the φ_q thresholds via the onion decomposition of H.** We start with formally defining the onion decomposition here. Recall from Definition 2.2 that $\rho(J_1|J_2) := \frac{|J_1 \backslash J_2|}{|V(J_1) \backslash V(J_2)|}$ for any graphs $J_2 \subseteq J_1$.

Definition 3.3 (Onion decomposition of H). Let $H = H_n$ be an arbitrary graph $H = H_n$. Consider the following procedure generating an increasing sequence of subgraphs $J^{(0)}, J^{(1)}, \cdots$ of H.

(i) Set $J^{(0)} = \varnothing$.
(ii) For each $t > 0$, let $J^{(t)}$ be a maximal subgraph in $\arg\max_{J \supsetneq J^{(t-1)}} \rho(J|J^{(t-1)})^6$.
(iii) If $J^{(t)} = H$, stop.

Let $M = M(H) \leqslant |H|$ be the number of steps required until the above process terminates. We refer to the outcome of this process, $J^{(t)}, t = 0, 1, 2, \ldots, M$, as the *onion decomposition* of H.

In words, the decomposition proceeds by setting $J^{(1)}$ as the densest subgraph of H, then removes $J^{(1)}$ and denotes by $J^{(2)}$ the new densest subgraph, and so on.

Remark 3.4. Observe for intuition that if the graph is balanced [ER60], that is $\max_{J \subseteq H} |J|/v(H) = |H|/V(H)$, then $M = 1$ and $J^{(0)} = \varnothing, J^{(1)} = H$.

Remark 3.5. Note that we refer to Definition 3.3 as *the* onion decomposition of H. This is because, as stated in Theorem 3.6 below, the choice of $J^{(t)}$ turns out to be unique for all t.

The following key combinatorial lemma establishes an important relation between the onion decomposition and the function $(\varphi_q)_{q \in [0,1]}$.

Theorem 3.6. *For any graph $H = H_n$, let $J^{(t)}, t = 0, 1, 2, \ldots, M$ be its onion decomposition from Definition 3.3. The following holds.*

(a) *For any $t = 0, 1, 2, \ldots, M - 1$, given $J^{(t)}$, the choice of $J^{(t+1)}$ is unique in step (ii) of the procedure. That is, the onion decomposition of a given graph H is unique.*
(b) *For any $q \in [0, 1)$, let $t = t(q)$ be such that $|J^{(t)}|/|H| \leqslant q < |J^{(t+1)}|/|H|$. Then*

$$\varphi_q = n^{-\frac{|V(J^{(t+1)}) \backslash V(J^{(t)})|}{|J^{(t+1)} \backslash J^{(t)}|}} = n^{-\frac{1}{\rho(J^{(t+1)}|J^{(t)})}},$$

i.e., the pair $(S = J^{(t)}, J = J^{(t+1)})$ is a minimax optimal pair in the definition of φ_q in (1.2).
(c) *The onion decomposition of H, and in particular, the function $(\varphi_q)_{q \in [0,1]}$ can be computed in time which is polynomial in $|H|$.*

The proof of these parts are based on a novel minimax duality principle that could be of independent interest. We defer the proof to Section 5.

Remark 3.7. (On the computability of φ_q.) Recall that φ_q is defined using the minimax optimization problem where both the minimum and the maximum operation is defined over exponentially many subsets of H. Fortunately, the simple structural characterization of Theorem 3.6 (a) and (b) allows us to reduce the problem to a modification of the densest subgraph problem, for which a polynomial time algorithm is known due to [Cha00]. The fact that we are working with graph-cuts introduces some complications, see Section 5.3 for details.

---

^{6}Maximality here means that there is no $J' \supsetneq J^{(t)}$ such that J' also maximizes the density $\rho(J|J^{(t-1)})$.

Now notice that Remark 3.2 characterizes the MMSE using φ_q, and Theorem 3.6 character-izes φ_q using the onion decomposition of H. Hence, their combination allows us to obtain the following corollary directly, which characterizes the MMSE at a much finer detail than before.

Corollary 3.8. *For any weakly dense graph $H = H_n$ let $J^{(t)}, t = 0, 1, 2, \ldots, M$ be its onion decomposition from Definition 3.3. If for some $t = 0, 1, 2, \ldots, M - 1$ it holds that*

$$\frac{|V(J^{(t+1)})\setminus V(J^{(t)})|}{|J^{(t+1)}\setminus J^{(t)}|} - \frac{|V(J^{(t)})\setminus V(J^{(t-1)})|}{|J^{(t)}\setminus J^{(t-1)}|} = \Omega\left(\frac{1}{\log n}\right), \tag{3.1}$$

then for any $\delta > 0$ if

$$(1 + \delta)n^{-\frac{|V(J^{(t+1)})\setminus V(J^{(t)})|}{|J^{(t+1)}\setminus J^{(t)}|}} \leqslant p \leqslant (1 - \delta)n^{-\frac{|V(J^{(t)})\setminus V(J^{(t-1)})|}{|J^{(t)}\setminus J^{(t-1)}|}}, \tag{3.2}$$

where in (3.2) we treat for convenience $n^{-0/0} := 1$, it holds that

$$\mathrm{MMSE}_n(p) = 1 - |J^{(t)}|/|H| + o(1).$$

Moreover, for any $\delta > 0$ if $p \leqslant (1 - \delta)n^{-\frac{|V(J^{(M)})\setminus V(J^{(M-1)})|}{|J^{(M)}\setminus J^{(M-1)}|}}$ then

$$\mathrm{MMSE}_n(p) = 1 - |J^{(M)}|/|H| + o(1) = o(1).$$

The proof is deferred to Section 8. Some remarks are in order.

Remark 3.9. (On the "jump" assumption (3.1).) Assumption (3.1) asks that at the $(t+1)$th step of the onion decomposition, the maximum density $|J^{(t)}|/|V(J^{(t)})|$ reached has a non-trivial "jump" from the tth step. The role of the assumption is to ensure that the interval for p considered in (3.2) is non-vacuous. Besides being of a technical nature, we highlight that the assumption is in fact satisfied for many H of interest: for example, it can be straightforwardly checked to be satisfied for all steps t for any graph H with size sub-logarithmic, i.e., satisfying $|H| = o(\log n)$ (also see Section 3.4 for more examples of such H.)

The reason we require the assumption is that, in our main result Theorem 3.1, a small $1 + \Omega(1)$ room between p and φ_q is needed to derive the asymptotic bounds on $\mathrm{MMSE}_n(p)$. We believe studying the behavior of the MMSE for p being $(1 + o(1))$-close to the thresholds φ_q, i.e., understanding the "sharpness" of these transitions [FB+99, GMZ23], to be an interesting direction for future work.

Remark 3.10. (Interpretation of Corollary 3.8 and the subgraph expectation threshold) For this remark, assume a weakly dense H such that the assumption (3.1) holds for all steps t of the onion decomposition. Then, we highlight two interesting implications of Corollary 3.8 in terms of the different values the MMSE can take and the thresholds that are the MMSE discontinuity points.

First, notice that Corollary 3.8 tells us that for any noise level p satisfying (3.2), the Bayes-optimal performance is achieved by an estimator that recovers exactly the edges of the $J^{(t)}$ sub-graph in H^*, up to an additive $o(1)$ error. One can therefore interpret our result as saying that for all noise levels p, the maximal possible subgraphs of H that can be recovered is always a pre-scribed element of the onion decomposition of H. As the noise level p decreases from one to zero, the first subgraph that can be recovered should be the densest one, i.e., $J^{(1)}$, the second should be

the densest subgraph containing $J^{(1)}$, i.e., $J^{(2)}$, and so on, until we can recover all of $J^{(M)} = H$. It is arguably surprising that these critical subgraphs for recovery are so well structured for any weakly dense H.

Moreover, Corollary 3.8 implies that the thresholds at which the MMSE "jumps" are exactly given as functions of the onion decomposition via the formula

$$p = n^{-\frac{|V(J^{(t+1)}\setminus V(J^{(t)})|}{|J^{(t+1)}\setminus J^{(t)}|}} = n^{-\frac{1}{\rho(J^{(t+1)}|J^{(t)})}}, t = 1, \ldots, M.$$

In particular, assuming the mild condition $|J^{(1)}|/|H| = \Omega(1)$ on H—a small variant of this natural condition is referred to as begin "delocalized" in [MNS$^+$23]—the above observation implies that for $t = 1$ the MMSE jumps from being trivial, i.e. taking the value $1 - o(1)$, to its first non-trivial value $1 - |J^{(1)}|/|H| + o(1) = 1 - \Omega(1)$ exactly at the threshold

$$p = n^{-\frac{|V(J^{(1)})|}{|J^{(1)}|}} = \max_{J\subseteq H} n^{-\frac{|V(J)|}{|J|}},$$

which is the subgraph expectation threshold for H defined by Kahn and Kalai in [KK07], thus proving Corollary 1.3. In other words, as mentioned in the introduction, the weak recovery threshold equals to the subgraph expectation threshold of H for any such weakly dense graph H. In particular, our work gives a statistical meaning to the *exact* subgraph expectation threshold of H from probabilistic combinatorics in Bayesian statistics, without the need to account for any additional $O(\log |H|)$-factor.

Remark 3.11 (A characterization of the AoN phenomenon). Recall that AoN happens for the planted subgraph model for some $H = H_n$ if the $\mathrm{MMSE}_n(p)$ curve converges to a step function from 0 to 1 as n grows to infinity. An immediate consequence of Theorem 3.1 is that the AoN phenomenon happens for a weakly dense graph if and only if for every $q, q' \in (0, 1)$,

$$\lim_n \frac{\varphi_q}{\varphi_{q'}} = 1.$$

Interestingly, this implication can be seen as a generalization of one of the main results in [MNS$^+$23], which requires an assumption of H being "delocalized" (see [MNS$^+$23, Definition 4.3]) to obtain a similar such AoN characterization.

3.3. **On the convergence of the MMSE curve.** Observe that all the above results discuss the behavior of the MMSE curve when n is finite but large. This is slightly diverging compared to the study of the limiting MMSE of other high dimensional statistical models, like compressed sensing [RP19] or the planted clique problem [MNS$^+$23], where the MMSE curve is (pointwise) converging as n grows to infinity to a specific curve and the curve is characterized, often via methods from statistical physics. This discrepancy is fundamental in our result as we consider a much larger family of priors compared to the above results. Furthermore, we do not assume any joint structure on H_n across different values of n; in other words, the planted subgraph H_n, and hence our prior, may change arbitrarily with n. For example, H could be a $10 \log n$-clique for even n, and a $\log \log n$-clique for odd n, leading to two convergent MMSE subsequences with different limits. More specifically, in this case, if say $p = 1/2$, it can be easily proven via the AoN theory from [MNS$^+$23] that the (normalized) MMSE equals $o(1)$ for even n and equals to $1 - o(1)$

for odd n, hence the MMSE as a curve is *not converging* as n grows to infinty. Yet, under the assumption that the sequence of graphs H_n is such that the pointwise limit of the MMSE *does* exist, we have the following asymptotic formula as a direct corollary of Theorem 3.1.

Corollary 3.12. *Let* $H = H_n$ *be a weakly dense graph and* $p = p_n \in (0, 1)$. *Let the function* $G : (0, 1] \rightarrow [0, +\infty)$ *defined by* $G(q) = \varphi_q$. *Consider the sequence* $\alpha_n \in (0, 1)$ *defined by*

$$\alpha_n = \inf G^{-1}([0, p]), n \in \mathbb{N}.$$

Then any limit point $x \in [0, 1]$ *of* $\{\alpha_n : n \in \mathbb{N}\}$, *corresponds in an one-to-one fashion to a limit point of the sequence* $\mathrm{MMSE}_n(p)$. *In particular, if the* $\lim_n \mathrm{MMSE}_n(p)$ *exists, then the* $\lim_n \alpha_n$ *exists and it holds,*

$$\lim_n \mathrm{MMSE}_n(p) = 1 - \lim_n \alpha_n. \tag{3.3}$$

The proof is based on simple calculus and is deferred to Section 8.

3.4. **A simple example.** We now work through a simple example to apply our established MMSE theory on weakly dense $H = H_n$ that illustrates the diversity of possible MMSE curves that one can obtain as the planted graph $H = H_n$ is varied. We emphasize at the outset that, while the simple example below can be solved directly by other methods, the calculation illustrates how our main theorem gives a general recipe that can be applied to compute the MMSE curve of any planted graph which is weakly dense.

A union of disjoint cliques.   Let $m = m_n$ and $r_1 > \cdots > r_m \geqslant 1$, be an arbitrary sequence of integers with

$$1/r_{i+1} - 1/r_i \geqslant \delta / \log n$$

for some constant $\delta > 0$. Such sequences include the cases $r_i = (C_i + o(1)) \log n, i = 1, 2, \ldots, m$ for some constants $C_1 > C_2 > \ldots > C_m > 0$, or $r_i, i = 1, \ldots, n$ is an arbitrary decreasing sequence of positive integers with $r_1 = o(\sqrt{\log n})$.

We then consider H to be the *disjoint* union of m cliques of sizes $r_1 > \cdots > r_m \geqslant 1$, so that

$$H = \bigcup_{s \leqslant m} K_{r_s}$$

and $k := |H| = \sum_{s \leqslant m} \binom{r_s}{2}$.

In terms of the onion decomposition of H (Definition 3.3), we have $J^{(0)} = \varnothing$ and the densest subgraph of H can be straightforwardly checked to be $J^{(1)} = K_{r_1}$. Similarly, a simple inductive argument implies that $M = M(H) = m$ and

$$J^{(t)} = \bigcup_{s \leqslant t} K_{r_s}, t = 1, 2, \ldots, m.$$

Moreover, from Theorem 3.6, we know that for any $1 > q \geqslant 0$, if $m - 1 \geqslant t \geqslant 0$ is maximal such that $|J^{(t)}|/|H| \leqslant q$, it holds

$$\varphi_q = p_t := n^{-\frac{|V(J^{(t+1)}) \setminus V(J^{(t)})|}{|E(J^{(t+1)}) \setminus E(J^{(t)})|}} = n^{-\frac{2}{r_{t+1}-1}}.$$

We also set $p_m := \varphi_1 = 0$ and $p_{-1} := 1$. Then using Corollary 3.8 we have that for any $\delta > 0$ and for each $-1 \leqslant t \leqslant m - 1$, if $(1 + \delta)p_{t+1} \leqslant p \leqslant (1 - \delta)p_t$, then

$$\mathrm{MMSE}_n(p) = 1 - \frac{\sum_{s \leqslant t+1} \binom{r_s}{2}}{k} + o(1).$$

For example, if $r_i = (C_i + o(1)) \log n$, $i = 1, 2, \ldots, m$ for some constants $C_1 > C_2 > \ldots > C_m > 0$ we conclude that $p_t = e^{-2/C_{t+1}} + o(1)$ for all $0 \leqslant t \leqslant m - 1$ and $p_m = 0, p_{-1} = 0$. Hence, we have that for any $\delta > 0$ and for each $-1 \leqslant t \leqslant m - 1$, if

$$(1 + \delta)e^{-2/C_{t+2}} \leqslant p \leqslant (1 - \delta)e^{-2/C_{t+1}},$$

then it holds

$$\mathrm{MMSE}_n(p) = 1 - \frac{\sum_{s \leqslant t+1} C_s^2}{\sum_{s \leqslant m} C_s^2} + o(1),$$

where set $e^{-2/C_{m+1}} := 0, e^{-2/C_0} := 1$.

If $m = 1$ and $C_1 = 2/\log 2$, we recover the AoN phenomenon for the planted clique model from [MNS+23] where the the clique is of size $k = (2 + o(1)) \log_2 n$ and the critical AoN noise level is $p = e^{-2/C_1} = 1/2$.

4. Main Results II: Planting general monotone properties

We now turn to a more general model, referred to as the planted subset model in [MNS+23]. In this general setting, we provide a weaker characterization of the MMSE curve as n grows, but under *no* assumptions on the planted structure. In particular, for the planted subgraph model discussed previously, this more general result applies to the MMSE curve for any $H = H_n$, even if not weakly dense.

To define this, let $N \in \mathbb{N}$ and fix some arbitrary family of $k = k_N$-sets $\mathcal{A} = \{A_i\}_{i=1}^M$ in a finite universe $\mathcal{X}$ of N elements, and some noise level $p = p_N$. Now, the "signal" is an element A from $\mathcal{A}$ which is sampled from an (arbitrary but known) prior π on $\mathcal{A}$. Then, the statistician observes

$$Y = A \cup X_p$$

where the "noise" X_p is a sample from the p-biased product measure on $\mathcal{X}$, i.e., $X_p \sim \mathrm{Ber}(p)^{\otimes \mathcal{X}}$, and aims to recover A from Y.

Remark 4.1. One may want to imagine $\mathcal{A}$ as the set of minimal elements of some monotone property of interest (see Section 2.2 for the standard definitions on monotone properties). Then the statistician aims to recover the planted "certificate" of this property in the presence of Bernoulli noise. The question of interest is: for what noise levels p is it possible to infer the certificate of the monotone property?

For example, the planted subgraph model is a special case of the planted subset model. Consider the universe $\mathcal{X}$ to correspond to the $N = \binom{n}{2}$ edges of K_n, the complete graph on n vertices. In that case, the Bernoulli noise X_p equals in distribution to $G(n, p)$. Now, for any subgraph $H = H_n$ of K_n, let $\mathcal{A}$ to be the collection of all subsets of $\mathcal{X}$ (i.e., all subgraphs of K_n) which are copies of H in K_n. Under this choice of $\mathcal{X}$ and $\mathcal{A}$, the planted subset model becomes the planted

subgraph model (it is easy to prove that the "worst" prior for planted subgraph models is the uniform one, see Lemma A.4 for the details).

Of course, the planted subset model generalizes well beyond the planted subgraph model to other interesting models, such as planting a connectivity certificate (i.e., a spanning tree) in $G(n, p)$ or planting arbitrary monotone properties on random hypergraphs.

As in the previous section, we focus on the MMSE of the model as a function of p, defined as

$$\text{MMSE}_N(p, \pi) = \frac{1}{k} \min_{\hat{A}} \mathbb{E}_{A \sim \pi, Y = A \cup X_p, \hat{A}} \left[\| \mathbf{1}_A - \hat{A}(Y) \|_2^2 \right] = \frac{1}{k} \mathbb{E}_{A \sim \pi, Y = A \cup X_p} \left[\| \mathbf{1}_A - \mathbb{E}[\mathbf{1}_A | Y] \|_2^2 \right],$$

where the minimum is over all randomized estimators $\hat{A} : 2^X \to \mathbb{R}^N$. As the prior is arbitrary in this model, we focus on the MMSE of estimating A under the "worst" possible prior π, which using the strong duality of linear programs can be written as

$$\text{MinMax}(\mathcal{A}, p) := \max_{\pi} \text{MMSE}_N(p, \pi) = \frac{1}{k} \min_{\hat{A}} \max_{\pi} \mathbb{E}_{A, Y, \hat{A}} \left[\| \hat{A}(Y) - \mathbf{1}_A \|_2^2 \right]. \qquad (4.1)$$

Remark 4.2. We note that the study of minimax rates of estimation, like $\text{MinMax}(\mathcal{A}, p)$, where one accounts for the worst-case prior is a topic of intense study for many decades in the literature of theoretical statistics. Moreover, the optimizers $\pi, \hat{A}$ in the definition of (4.1) are called the "least favorable prior" and the "minimax estimator", respectively.

Our main result is to identify the order of the noise level p for which $\text{MinMax}(\mathcal{A}, p)$ achieves a given fixed value of interest in $(0, 1]$. Recall that in the planted subgraph model for a weakly dense H and under a uniform prior, Theorem 3.1 implies that $\text{MMSE}_n(p) = 1 - q + o(1)$ for some $p = (1 + o(1))\varphi_q$ where the threshold φ_q is a modified subgraph expectation threshold. In the planted subset model, for any $q \in [0, 1)$, the desired threshold is given by a modification of the so-called *fractional expectation* thresholds of the monotone property generated by $\mathcal{A}$.

Definition 4.3. Fix any family of k-sets $\mathcal{A}$ in a finite universe X of N elements. For any $q \in [0, 1)$, we define

$$\psi_q := (\psi_q)_N = \min_{S \subseteq X, |S| \leqslant qk} \left\{ \max_{w \in W_S} \left\{ \mathcal{E}_w^{-1}(1/2) \right\} \right\},$$

where given $S \subset X$, $\mathcal{E}_w(p) := \sum_{T \subseteq X} w(T) p^{|T \setminus S|}$, and W_S is the set of all "fractional covers" w which are functions $w : 2^X \to [0, 1]$ satisfying the following two properties.

(a) ("Support condition") Unless $|T| \geqslant qk$ and $S \subseteq T$, it holds $w(T) = 0$.
(b) ("Fractional cover condition") For all $A_i, i = 1, 2, \ldots, M$ where $S \subseteq A_i$ we have

$$\sum_{T \subseteq A_i} w(T) \geqslant 1.$$

We finally set $\psi_1 := 0$.

Remark 4.4. We remark that the case $q = 0$ and $S = \varnothing$ corresponds exactly to the fractional expectation threshold [Tal10] of the monotone property with minimal elements $\mathcal{A}$, see Definition 2.4. In particular, the value of the fractional expectation threshold has been calculated

up to constants for a number of monotone properties, including the H-inclusion property for a subgraph $H = H_n$ of the complete graph K_n [FKNP21, MNSZ22b].

4.1. Characterizing the Minimax curve via the ψ_q thresholds.
We now present our general result in this setting.

Theorem 4.5. *Let $q \in [0, 1)$, and $\mathcal{A}$ be an arbitrary k-uniform family of subsets of $[N]$. Then, as $N \to \infty$,*

(a) If $\lim_N p/\psi_q = +\infty$, then

$$\liminf_N \text{MinMax}(\mathcal{A}, p) \geqslant 1 - q.$$

(b) If $\lim_N p/\psi_q = 0$, then

$$\limsup_N \text{MinMax}(\mathcal{A}, p) \leqslant 1 - q.$$

The proof is deferred to Section 7.

Remark 4.6. We remark that while we are able to efficiently compute φ_q in the setting of weakly dense graphs using the onion decomposition of H, we do not yet have a way to efficiently compute an optimizer for ψ_q for arbitrary monotone properties. It is an interesting question for future work if a similar onion-peeling process exists for any monotone property.

Despite this, given the success of probabilistic combinatorics in computing the fractional expectation thresholds for various different properties [FKNP21], we find the calculation of ψ_q for specific monotone properties an interesting direction—by Theorem 4.5, this would have implications for the minimax rates of such problems. We next go over such a calculation for the well-studied case where H is a perfect matching.

4.2. A simple example.
We now work through a simple example in this more general case.

Perfect matching. Let us consider the case that H is the perfect matching on n vertices (n even). This is a model that has received significant interest in the community (see e.g., [DWXY23, GSXY25] and references therein). In this setting, the minimax rate is equal to the MMSE for the uniform distribution over planted matchings (see Lemma A.4).

We will show via direct computation that for any $q \in (0, 1)$,

$$\psi_q = (1 + o(1)) \frac{e}{n}.$$

Applying Theorem 4.5, we can conclude the following.

- If $p = o\left(\frac{1}{n}\right)$, then the MMSE will be $o(1)$ for any prior. We note that this is a result already established in [DWXY23], by a direct analysis in this setting.
- If $p = \omega\left(\frac{1}{n}\right)$, then there exists a prior (the uniform distribution) for which the MMSE will be $1 - o(1)$. We note that this was very recently independently proved by [GSXY25], using a direct analysis.

Recall that

$$\psi_q = \min_{S \subseteq \mathcal{X}, |S| \leqslant qk} \left\{ \max_{w \in W_S} \left\{ \mathcal{E}_w^{-1}(1/2) \right\} \right\}.$$

First, we observe that for any choice of S, the max is the fractional Kahn–Kalai threshold τ_k for a perfect matching on $k = n - v(S)$ vertices. This is because any S will be disconnected from the rest of H, so the resulting graph-cut is essentially just a normal graph. Therefore, we have that

$$
\psi_q = \min_{S \subseteq \mathcal{X}, |S| \leqslant qk} \left\{ \tau_{n-v(S)} \right\}
$$
$$
= \min_{k \geqslant (1-q)n} \tau_k
$$

It remains to compute τ_k, the fractional Kahn Kalai threshold for a perfect matching in a size k graph.

From [MNSZ22b], for subgraph inclusion properties, the fractional Kahn–Kalai threshold is equal to the following:

$$
\tau_n = \max_{J \subseteq H} \left(\frac{H[J]}{K_n[J]} \right)^{1/|J|}
$$

where $G[J]$ for a graph G refers to the number of unlabeled copies of J in G.

In the case of perfect matchings, τ_k can be expressed as

$$
\tau_k = \max_{0 \leqslant r \leqslant k/2} \left(\frac{M_{k/2}[M_r]}{K_k[M_r]} \right)^{2/r} = \max_{1 \leqslant r \leqslant k/2} \left(\frac{\binom{k/2}{r}}{\binom{k}{2r} \cdot (2r-1)!!} \right)^{2/r},
$$

with M_t denoting the matching with t edges. We shall show that the above is maximized for $r = k/2$. Letting a_r be the expression inside the parentheses, we have for $r \geqslant 1$,

$$
a_r = \frac{\binom{k/2}{r}}{\binom{k}{2r} \cdot (2r-1)!!} = \frac{\binom{k/2}{r-1}}{\binom{k}{2r-2} \cdot (2r-3)!!} \cdot \frac{\frac{k/2-r}{r}}{\frac{(k-2r)(k-2r+1)}{2r(2r-1)} \cdot (2r-1)} = a_{r-1} \cdot \frac{1}{k-2r+1}.
$$

A simple inductive argument then implies that $a_r \leqslant \left(\frac{1}{k-2r-1} \right)^r$. Indeed, this is trivially true for $r = 1$. For larger r, we have by induction that $a_r = \frac{1}{k-2r+1} \cdot a_{r-1} \leqslant \frac{1}{k-2r+1} \cdot \left(\frac{1}{k-2r+1} \right)^{r-1} \leqslant \left(\frac{1}{k-2r-1} \right)^r$ as desired. Given this,

$$
a_r^{1/r} = \left(\frac{1}{k-2r+1} \cdot a_{r-1} \right)^{1/r} \geqslant \left(f(r-1)^{1/r-1} \cdot a_{r-1} \right)^{1/r} = a_{r-1}^{1/(r-1)}.
$$

It follows that the maximum is attained for $r = k/2$. Thus,

$$
\tau_k = \left(\frac{1}{(k-1)!!} \right)^{2/k}
$$

Observe that this is minimized at $k = n$. Approximating $(n-1)!! = \left(\frac{n}{e} \right)^{n/2+o(1)}$, we have that

$$
\tau_n = \left((e/n)^{n/2+o(1)} \right)^{2/n}
$$
$$
= (1 + o(1)) \frac{e}{n}.
$$

Hence we conclude that

$$\psi_q = (1 + o(1))\frac{e}{n}$$

for all $q \in [0, 1]$, as desired.

5. Onion Decomposition: Proof of Theorem 3.6

In this section, we establish some key results on the onion decomposition that will be useful in multiple proofs throughout the paper and also prove Theorem 3.6.

We start with recalling the ρ notation from Section 2. Notice that our key quantity (φ_q) can be defined via ρ; if we define for $q \in [0, 1]$ the quantity

$$\rho_q := \min\left\{\max\left\{\rho(J|S) : J \supseteq S\right\} : S \subseteq H, |S| \leqslant q|H|\right\} \tag{5.1}$$

then it clearly holds

$$\varphi_q = n^{-\frac{1}{\rho_q}}.$$

5.1. **Key lemmas.** We now establish four useful combinatorial lemmas for the "density" ρ function, that follow from elementary graph theory arguments. We defer their proofs to Section 8.

Lemma 5.1. *Let A, B be subgraphs. Then,*

$$\rho(A \cup B|A) \geqslant \rho(B|A \cap B).$$

Lemma 5.2. *Let $K|T$ be a graph-cut that is "balanced", in that for all T' such that $T \subseteq T' \subseteq K$, $\rho(T'|T) \leqslant \rho(K|T)$. Then, for all T' such that $T \subseteq T' \subseteq K$ it also holds*

$$\rho(K|T') \geqslant \rho(K|T).$$

The third lemma implies that modulo any subgraph A, there exists a unique maximizer of the density $\rho(J|A), J \supseteq A$.

Lemma 5.3. *Let A be a subgraph and consider the map φ_A which maps any subgraph J containing A to $\varphi_A(J) = \rho(J|A)$. If J' and J'' are two subgraphs containing A and at least one of them is balanced, then $\varphi_A(J' \cup J'') \geqslant \min\{\varphi_A(J'), \varphi_A(J''')\}$. In particular, if J', J'' are two maximizers of φ_A, then $J' \cup J''$ is also a maximizer of φ_A, so there is a unique maximal maximizer of φ_A.*

The fourth lemma asserts that the densities in the onion decomposition are non-increasing.

Lemma 5.4. *Let A be a subgraph and $B = \arg\max_{J \supseteq A} \rho(J|A)$. Then for all $C \supseteq B$,*

$$\rho(C|B) \leqslant \rho(B|A).$$

In particular, for $J^{(t)}, t = 0, 1, \ldots, M$ the onion decomposition of any subgraph H the sequence $\rho(J^{(t+1)}|J^{(t)}), t = 0, 1, \ldots, M - 1$ is non-increasing.

5.2. **Simplifying φ_q: Proofs of Theorem 3.6(a) and (b).** The above lemmas allow us to conclude parts (a) and (b) of Theorem 3.6, restated for convenience.

Theorem 3.6. *For any graph $H = H_n$, let $J^{(t)}, t = 0, 1, 2, \ldots, M$ be its onion decomposition from Definition 3.3. The following holds.*

(a) *For any $t = 0, 1, 2, \ldots, M - 1$, given $J^{(t)}$, the choice of $J^{(t+1)}$ is unique in step (ii) of the procedure. That is, the onion decomposition of a given graph H is unique.*

(b) *For any $q \in [0, 1)$, let $t = t(q)$ be such that $|J^{(t)}|/|H| \leqslant q < |J^{(t+1)}|/|H|$. Then*

$$\varphi_q = n^{-\frac{|V(J^{(t+1)})\setminus V(J^{(t)})|}{|J^{(t+1)}\setminus J^{(t)}|}} = n^{-\frac{1}{\rho(J^{(t+1)}|J^{(t)})}},$$

i.e., the pair $(S = J^{(t)}, J = J^{(t+1)})$ is a minimax optimal pair in the definition of φ_q in (1.2).

(c) *The onion decomposition of H, and in particular, the function $(\varphi_q)_{q\in[0,1]}$ can be computed in time which is polynomial in $|H|$.*

Lemma 5.3 applied to $A = J^{(t)}$ immediately implies (a).

Proof of (b). Fix some $q \in [0, 1)$ and let (S^*, J^*) be a minimax pair in the definition of ρ_q, which in particular also satisfies

$$\varphi_q = n^{-1/\rho(J^*|S^*)}.$$

In particular, our goal is to show that for t such that $|J^{(t)}| \leqslant q|H| < |J^{(t+1)}|$, it holds that $\rho(J^{(t+1)}|J^{(t)}) = \rho(J^*|S^*)$.

We start with choosing t to be the step of the onion decomposition such that $J^{(t)} \subseteq S^*$ but $J^{(t+1)} \nsubseteq S^*$. We first prove that for this choice of t,

$$\rho(J^{(t+1)}|J^{(t)}) = \rho(J^*|S^*) \tag{5.2}$$

By the definition of the onion decomposition, $J^{(t+1)} \in \arg\max_{J \supseteq J^{(t)}} \rho(J|J^{(t)})$. Combining this with the optimality of S^*, that means we have

$$\rho\left(J^{(t+1)}|J^{(t)}\right) = \max_{J \supseteq J^{(t)}} \rho(J|J^{(t)}) \geqslant \rho(J^*|S^*).$$

Next, defining $\widetilde{J} = J^{(t+1)} \cup S^*$ and $\widetilde{S} = J^{(t+1)} \cap S^*$, Lemma 5.1 implies that

$$\rho(\widetilde{J}|S^*) \geqslant \rho(J^{(t+1)}|\widetilde{S}).$$

Finally, by Lemma 5.2,

$$\rho(J^{(t+1)}|\widetilde{S}) \geqslant \rho(J^{(t+1)}|J^{(t)}).$$

Hence, combining the last three displayed inequalities, we have

$$\rho(\widetilde{J}|S^*) \geqslant \rho(J^*|S^*).$$

But by the optimality of J^* given S^*, $\max_{J \supseteq S^*} \rho(J|S^*) = \rho(J^*|S^*)$, all the above four displayed inequalities should be equalities, proving (5.2).

Thus, for all $q \in [0, 1]$,

$$\rho_q \in \{\rho(J^{(t+1)}|J^{(t)}) : 0 \leqslant t \leqslant q|H|\}.$$

But again as by definition for all t, $\rho\left(J^{(t+1)}|J^{(t)}\right) = \max_{J \supseteq J^{(t)}} \rho(J|J^{(t)})$ it also holds by setting $S = J^{(t)}, t \leqslant q|H|$ and the definition of ρ_q that

$$\rho_q \leqslant \min\{\rho(J^{(t+1)}|J^{(t)}) : 0 \leqslant t \leqslant q|H|\}.$$

Hence,

$$\rho_q = \min\{\rho(J^{(t+1)}|J^{(t)}) : 0 \leqslant t \leqslant q|H|\}.$$

But by Lemma 5.4 we have that $\rho(J^{(t+1)}|J^{(t)})$ is a non-increasing function of t. Therefore the last displayed equation gives $\rho_q = \rho(J^{(t+1)}|J^{(t)})$ for the maximum possible $t \leqslant q|H|$, from which (b) follows. $\qquad\square$

5.3. Computability of φ_q: Proof of Theorem 3.6(c). We further have the following lemma, which shows that this procedure can be implemented in polynomial-time.

Proof. We first show that there is an efficient algorithm to find the densest subgraph in a graph-cut. Specifically, given a graph-cut $G|S$, we show how to find a $J \subseteq G|S$ attaining the value $f(G|S) = \max_{J \subseteq G|S} \rho(J)$.

The algorithm and its analysis are based on Charikar's Linear Programming (LP) approach to find a densest subgraph in a graph [Cha00]. We consider the following LP:

$$
\begin{aligned}
\text{maximize} \quad & \sum_{ij \in G \setminus S} x_{ij} \\
\text{subject to} \quad & x_{ij} \leqslant y_i && \forall i \in V(G) \setminus V(S) \\
& x_{ij} \leqslant y_i && \forall j \in V(G) \setminus V(S) \\
& \sum_{i \in V(G) \setminus V(S)} y_i \leqslant 1 \\
& x_{ij}, y_i \geqslant 0 && \forall ij \in G|S.
\end{aligned}
$$

Let v^* be the optimal value of this LP. To see that $v^* \geqslant f(G|S)$, note that for every $J \subseteq G|S$, letting $y_i = \frac{1}{v(J)}\mathbf{1}\{i \in V(J)\}$ and $x_{ij} = \frac{1}{v(J)}\mathbf{1}\{ij \in J\}$, we get $\sum_{ij \in G|S} x_{ij} = \rho(J)$. Moreover, $\sum_{i \in V(G) \setminus V(S)} y_i = 1$, and the other two constraints are trivially satisfied.

Next, given a solution to the LP with value v^*, we show how to extract a choice of J such that $\rho(J) \geqslant v^*$. Let $\{\overline{x}_{ij}\}_{ij} \cup \{\overline{y}_i\}_i$ be such a solution. Note that we can assume without loss of generality that $\overline{x}_{ij} = \min\{\overline{y}_i, \overline{y}_j\}$ when $i, j \in V(G) \setminus V(S)$, $\overline{x}_{ij} = \overline{y}_i$ when $j \in V(S)$, and $\overline{x}_{ij} = \overline{y}_j$ when $i \in V(S)$. For each $r \in \mathbb{R}_{\geqslant 0}$, consider the subgraph $E(r) = \{ij : \overline{x}_{ij} \geqslant r\} \subseteq G \setminus S$ and denote by $V(r) := \{i : \overline{y}_i \geqslant r\}$. Note that for each r, by our assumption, for any edge of the subgraph $E(r)$, $ij \in E(r)$ it can be easily checked to hold $\{i, j\} \setminus V(S) \subseteq V(r)$, and therefore

$$V(E(r)) \setminus V(S) \subseteq V(r). \tag{5.3}$$

Now, we claim that there exists r such that $\rho(V(r), E(r)) \geqslant v^*$. Indeed, since $\int_0^\infty |E(r)| dr = \sum_{ij \in G|S} \overline{x}_{ij}$ and $\int_0^\infty |V(r)| dr = \sum_{i \in V(G) \setminus V(S)} \overline{y}_i \leqslant 1$, we get

$$
\begin{aligned}
v^* &\leqslant \frac{\int_0^\infty |E(r)| dr}{\int_0^\infty |V(r)| dr} \\
&\leqslant \sup_r \frac{|E(r)|}{|V(r)|} \\
&\leqslant \sup_r \frac{|E(r) \setminus S|}{|V(E(r)) \setminus V(S)|}.
\end{aligned}
$$

But then note that we can efficiently enumerate all combinatorially distinct $E(r)$ to find a $J = E(r)$ such that

$$
\rho(J) = \rho(E(r)) = \frac{|E(r) \setminus S|}{|V(E(r)) \setminus V(S)|} \geqslant v^*,
$$

as desired. Given the above, we can efficiently construct sequences $\varnothing = \widetilde{J}^0, \widetilde{J}^1, \ldots, \widetilde{J}^\ell$ and $\widetilde{J}^{(t)} = \cup_{s \leqslant t} \widetilde{J}^s$ so that, for each $t \leqslant \ell - 1$, $\widetilde{J}^{t+1}$ is a maximizer of $J \mapsto \rho(J | \widetilde{J}^{(t-1)})$ among $J \supseteq \widetilde{J}^{(t-1)}$.

To efficiently construct the onion decomposition, we need yet to show something slightly stronger: we need to efficiently construct the sequence $\varnothing = J^0, J^1, \ldots, J^m$ such that, for $J^{(t)} = \cup_{s \leqslant t} J^s$, we have that J^{t+1} is the *maximal* J maximizing the function $J \mapsto \rho(J | \widetilde{J}^{(t-1)})$ among $J \supseteq \widetilde{J}^{(t-1)}$ (the uniqueness of the maximal J follows by Theorem 3.6).

We do this in a greedy fashion as follows. Let $s_1 \geqslant 1$ be maximal such that

$$
\rho(\widetilde{J}^1 | \widetilde{J}^{(0)}) = \cdots = \rho(\widetilde{J}^{s_1} | \widetilde{J}^{(s_1 - 1)}),
$$

and let $J^1 = \cup_{s \leqslant s_1} \widetilde{J}^s$. Then let s_2 be maximal such that

$$
\rho(\widetilde{J}^{s_1 + 1} | \widetilde{J}^{(s_1)}) = \cdots = \rho(\widetilde{J}^{s_2} | \widetilde{J}^{(s_2 - 1)}),
$$

and let $J^2 = \cup_{s_1 < s \leqslant s_2} \widetilde{J}^s$, and so on. Notice that using Lemma 5.4 we have for all t,

$$
\rho(\widetilde{J}^{s_t + 1} | \widetilde{J}^{(s_t)}) < \rho(\widetilde{J}^{s_t} | \widetilde{J}^{(s_t - 1)}). \tag{5.4}
$$

Now, it is straightforward to check that for all t J^t maximizes $J \mapsto \rho(J | J^{(t-1)})$. Hence, it remains to show that J^t is in fact a *maximal* such maximize for all. Suppose that this is not the case for some t. Then by Lemma 5.3, there exists $J' \nsubseteq J^t$ so that $J' \cup J^t$ is the maximal maximizer of $J \mapsto \rho(J | J^{(t-1)})$. In particular, it holds

$$
\rho(J' \cup J^t | J^{(t-1)}) = \rho(J^t | J^{(t-1)}) = \rho(\widetilde{J}^{s_t} | \widetilde{J}^{s_t - 1})
$$

which implies for $J'' := J' \setminus J^t$ that

$$
\rho(J'' | \widetilde{J}^{s_t}) = \rho(\widetilde{J}^{s_t} | \widetilde{J}^{s_t - 1}).
$$

We can therefore guarantee that

$$
\rho(\widetilde{J}^{s_t + 1} | \widetilde{J}^{(s_t)}) \geqslant \rho(J'' | \widetilde{J}^{s_t}) = \rho(\widetilde{J}^{s_t} | \widetilde{J}^{(s_t - 1)}),
$$

a contradiction with (5.4).   □

6. The MMSE for weakly dense H: Proof of Theorem 3.1

In this section, we give an overview of the main technical steps involved in the proof of our main result, Theorem 3.1. Our first step is a "positive" result, showing that if p is smaller than a threshold $\widetilde{\varphi}_q$, then the noise is small enough so that the limiting MMSE is at least $1 - q$. The second step is a "negative" result, proving that if p is larger than φ_q then the noise is high enough so that the limiting MMSE is at most $1 - q$. Our final step shows that these two thresholds, $\widetilde{\varphi}_q$ and φ_q, are in fact equal.

6.1. **Two lemmas: the importance of the order of the min and max operators.** Our first lemma, proved in Section 6.4, shows that for p below a modified version of the φ_q threshold defined in (1.2) (where importantly the min and max operators have been exchanged) it is possible to achieve MMSE at most $1 - q$.

Lemma 6.1 (Possibility of Recovery). *Suppose $H = H_n$ is a weakly dense H. Given $q \in [0, 1)$, define the threshold*

$$\widetilde{\varphi}_q = (\widetilde{\varphi}_q)_n = \max \left\{ \min \left\{ n^{-\frac{|V(J) \setminus V(S)|}{|J \setminus S|}} : S \subseteq J, |S| \leqslant q|H| \right\} : J \subseteq H \right\}.$$

Then for any $p = p_n$ with $\limsup_n p / \widetilde{\varphi}_q < 1$, it holds

$$\limsup_n \mathrm{MMSE}_n(p) \leqslant 1 - q.$$

The lemma is proven by showing that for p below this new $\widetilde{\varphi}_q$ threshold, with high probability there does not exist an isomorphic copy H' of the planted graph H^* in the observed graph $G = H^* \cup G_0, G_0 \sim G(n, p)$ which overlaps with H^* in fewer than $(q - o(1))|H|$ edges. We then prove via the Nishimori identity that the MMSE is equal to one minus the expected overlap between a random isomorphic copy of H and H^*, from which we can directly upper bound the MMSE by $1 - q + o(1)$. The proof is deferred to Section 6.4.

The intuition behind the variational formula in Lemma 6.1 for $\widetilde{\varphi}_q$ is as follows. In order to avoid having a copy H' of H in G with overlap $\leqslant q|H|$ with the planted H^*, we need to understand for which noise levels p we can rule out such structures. Let us denote by S this potential intersection set of edges between a copy H' of H, and H^*. Once we fix such an S, a bound on the largest value of p such that no copy of H appears in the observed graph G overlapping with H^* exactly on S can be obtained by a standard first moment method (i.e, a vanilla union bound). Then optimizing the bound to account for the worst-case S, leads to a bound on p to have MMSE at most $1 - q$. However, in many cases this bound will not be tight. To improve it, using the intuition from the tightness of the expectation thresholds from [KK07], one can fix a subgraph $J \subseteq H$ and instead upper bound for any p the probability that a copy H' of H in G overlaps H^* on S by the probability that *a copy of J* in G overlaps with H^* in at most $\leqslant q|H|$ edges (simply because $J \subseteq H$). Then, the same standard first moment method argument as above can be performed with J in place of H to obtain a new upper bound on the probability that a copy H' of H in G overlaps H^* on S, which depends on J and S. Now, notice that by optimizing in a worst-case sense over S (to guarantee an MMSE at most $1 - q$), and then optimizing over all subgraphs J (now, in favor

of the statistician so choosing the "best-case" J), we arrive exactly at the max-min formula for $\widetilde{\varphi}_q$ giving Lemma 6.1.

We next describe our second lemma, proved in Section 6.5, which shows that above the φ_q threshold, the MMSE must be worse than $1 - q$.

Lemma 6.2 (Impossibility of recovery). *Suppose $H = H_n$ is a weakly dense H. Given $q \in [0, 1)$, recall the threshold φ_q from (1.2). Then for any $p = p_n$ with $\liminf_n p/\varphi_q > 1$, it holds*

$$\liminf_n \mathrm{MMSE}_n(p) \geqslant 1 - q.$$

The proof shows that when p crosses φ_q, the posterior distribution of H^*, which is the uniform distribution over the copies of H in the observed graph, is concentrated on copies with overlap less than $q|H|$ with H^*. Our proof follows from an appropriate leverage of the planting trick [ACO08] similar to its application in the recent Bayesian proof of the spread lemma [MNSZ22b]. The proof is deferred to Section 6.5.

Here, the intuition for φ_q is as follows. Given $S \subseteq H$ with $|S| \leqslant q|H|$, one can define a statistically easier version of the estimation task of interest, in which the statistician is also given as side information the S part of the planted graph H^*. An application of the data processing inequality implies that the inability to recover a given fraction of H^* in this easier problem implies the same inability in the original statistical problem (Such techniques are often called "genie" lower bound arguments, e.g, [RG13]).

To prove the MMSE lower bound it suffices to determine the noise levels p for which the MMSE is at least $1 - q$ when the statistician is given S as side information, and then optimizing over S to find the minimum such noise level. For this reason, we start by fixing any S with $|S| \leqslant q|H|$ and assume that the S part of H is given to the statistician. Then a simple posterior calculation implies that MMSE is at least $1 - |S|/|H| \geqslant 1 - q$, if a typical copy H' (of H) in $G = H^* \cup G_0$ that contains S does not overlap an $\Omega(1)$-fraction of edges in $H^* \backslash S$. Equivalently, this is true if most copies H' containing S are such that $H' \backslash S$ is (almost) fully included in G_0.

For this reason, it makes sense to study the threshold of p at which many (near-complete) copies of $H \backslash S$ start appearing in $G_0 \sim G(n, p)$. The intuition from the literature on Kahn–Kalai thresholds suggests that this threshold is given by the subgraph expectation threshold of $H \backslash S$, which is precisely the maximizing problem in the definition of φ_q. While the tightness of this threshold is still at a conjectural phase and expected to hold only up to a logarithmic error, we prove that the approximate version of it (recall we want $H \backslash S$ to only approximately appear in G) that holds exactly above the subgraph expectation threshold. Optimizing over all subgraphs S to determine the best such bound results in φ_q, the minimum possible subgraph expectation threshold of $H \backslash S$ among all S.

6.2. **Exchanging the max and min operators.** To conclude Theorem 3.1, we must relate the thresholds in the two lemmas above. It is somewhat dissatisfying that the natural probabilistic way to argue about Lemma 6.1 and Lemma 6.2 leads to almost the same threshold for p, but with the max and the min operators being exchanged in the two results. While there are plenty of "minimax results" in the literature for exchanging the min and the max operators (e.g., in the literature of game theory and of optimization), no such result appears applicable in our discrete

vertex-symmetric setting. Yet, a careful combinatorial argument allows us to indeed exchange the operators. This minimax result is the content of the next lemma, proved in Section 6.6, and is proved by employing the onion decomposition of H.

Lemma 6.3 (Minimax result). *For any n, $q = q_n \in [0, 1)$, and $H = H_n$, $\widetilde{\varphi}_q = \varphi_q$.*

We remark that the duality lemma above is an exact, non-asymptotic equality: it holds for all graphs $H = H_n$ and for all n. In order to prove it, our main tool is a structural result for the max-min and min-max optimal pairs (S, J) in φ_q. Observe that the pair (S, J) can take exponentially many values. Yet, by considering the onion decomposition $(J^{(t)})$ of H (Definition 3.3), it turns out that, as explained in Section 5, the minimax optimization defining φ_q has a simple optimizing pair: it is exactly of the form $(S = J^{(t)}, J = J^{(t+1)})$, where t is maximal such that $|J^{(t)}| \leqslant q|H|$. For instance, for q close to 1, the optimal pair is to choose $S = J^{(1)}$, the densest subgraph of H and $J = J^{(2)}$ the "second" densest subgraph in the onion decomposition. The knowledge of the optimizers greatly simplifies φ_q and is key to establishing Lemma 6.3.

6.3. **Putting it all together.** Notice that the above lemmas immediately imply Theorem 3.1.

Proof of Theorem 3.1. Parts (a) and (b) follow immediately from Lemmas 6.1 to 6.3.

The last part follows from the fact that $\mathrm{MMSE}_n(p)$ is a non-increasing polynomial function p (see Lemma A.2) together with a simple analysis argument to construct the desired sequence. Indeed, for any $\varepsilon > 0$, and for large enough n, parts (a), (b) imply

$$\mathrm{MMSE}((1 - \varepsilon)\varphi_q) \leqslant 1 - q + \varepsilon$$

and

$$\mathrm{MMSE}((1 + \varepsilon)\varphi_q) \geqslant 1 - q - \varepsilon.$$

That means that for any $\varepsilon > 0$ there exists $n_0 = n_0(\varepsilon)$ large enough such that for all $n \geqslant n_0$ there exists $p = p_n \in [(1 - \varepsilon)\varphi_q, (1 + \varepsilon)\varphi_q]$ for which

$$1 - q - \varepsilon \leqslant \mathrm{MMSE}(p) \leqslant 1 - q + \varepsilon.$$

Iterating this argument for $\varepsilon = 1/m, m = 1, 2, \ldots$ leads to (potentially distinct) sequences of noise levels $p^{(m)} = p_n^{(m)}$ such that for all m, there exists a (without loss of generality strictly increasing) $G(m) \in \mathbb{N}$ such that for all $n \geqslant G(m)$, $p_n^{(m)} \in [(1 - 1/m)\varphi_q, (1 + 1/m)\varphi_q]$ and

$$1 - q - 1/m \leqslant \mathrm{MMSE}(p_n^{(m)}) \leqslant 1 - q + 1/m.$$

Now, "diagonalizing" this argument, consider the sequence defined as follows: for each $n \in [G(m), G(m + 1))$, let

$$p = p_n := p_n^{(m)}$$

(for $n < G(1)$, set $p_n = 0$). Then notice that clearly for this sequence it holds $\lim_n p_n/\varphi_q = 1$ and also $\lim_n \mathrm{MMSE}_n(p_n) = 1 - q$ as desired. $\square$

6.4. Proof of Lemma 6.1: Recovery for $p \leqslant (1 - \varepsilon)\widetilde{\varphi}_q$. In this section, we prove Lemma 6.1. In particular, we assume that for some $\varepsilon > 0$ and n large enough it holds $p \leqslant (1 - \varepsilon)\varphi_q$.

Proof. Fix $\delta > 0$ to be a constant and set $k := |H|$. It suffices to show that the $\mathrm{MMSE}_n(p)$ is at most

$$1 - q + \delta + o(1).$$

Since δ is arbitrary, we may eventually take δ to 0 to obtain the desired statement.

Observe that by Bayes' rule, the posterior of the model is simply the uniform distribution over all copies of H in the observed graph. Using the Nishimori identity (Lemma A.1), we have

$$1 - \mathrm{MMSE}_n(p) = \frac{1}{k} \mathop{\mathbb{E}}_{\substack{G \sim G(n,p) \\ H^* \sim \text{uniform copy of } H \text{ in } K_n}} \left[\mathbb{E}_{H' \sim \text{copies of } H \text{ in } G \cup H^*} |H' \cap H^*| \right].$$

We bound this as follows,

$$1 - \mathrm{MMSE}_n(p) \geqslant (q - \delta) - \frac{1}{k} \mathop{\mathbb{E}}_{\substack{G \sim G(n,p) \\ H^* \sim \text{uniform copy of } H \text{ in } K_n}} \left[\mathbb{E}_{H' \sim \text{copies of } H \text{ in } G \cup H^*} |H' \cap H^*| \cdot \mathbf{1}_{|H' \cap H^*| \leqslant (q-\delta)k} \right]$$

$$\geqslant (q - \delta) - \mathop{\mathbb{E}}_{\substack{G \sim G(n,p) \\ H^* \sim \text{uniform copy of } H \text{ in } K_n}} \left[\mathbb{E}_{H' \sim \text{copies of } H \text{ in } G \cup H^*} \mathbf{1}_{|H' \cap H^*| \leqslant (q-\delta)k} \right]$$

$$\geqslant (q - \delta) - \mathbb{P}_{G,H^*} \left[\exists \text{ a copy } H' \text{ of } H \text{ in } G \text{ such that } |H^* \cap H'| \leqslant (q - \delta)k \right]$$

It suffices to show that

$$\mathbb{P}_{G,H^*} \left[\exists \text{ a copy } H' \text{ of } H \text{ in } G \text{ such that } |H^* \cap H'| \leqslant (q - \delta)k \right] = o(1).$$

For starters, let J^* be an optimal choice of J in the definition of $\widetilde{\varphi}_q$. Then, using the notation $\cong$ to denote the graph isomorphism relation, we have

$$\mathbb{P} \left[\exists\, H' \subseteq G \cup H^* \text{ such that } H' \cong H \text{ and } |H^* \cap H'| \leqslant (q - \delta)k \right]$$
$$\leqslant \mathbb{P} \left[\exists\, J' \subseteq G \cup H^* \text{ such that } J' \cong J^* \text{ and } |H^* \cap J'| \leqslant (q - \delta)k \right]$$

This can be bounded further by Markov's inequality and the definition of the $G(n, p)$ noise to give,

$$\mathbb{P} \left[\exists\, J' \subseteq G \cup H^* \text{ such that } J' \cong J^* \text{ and } |H^* \cap J'| \leqslant (q - \delta)k \right]$$
$$\leqslant \sum_{\ell \leqslant (q-\delta)k} \mathbb{E}[|J' \subseteq G \cup H^* \text{ such that } J' \cong J^* \text{ and } |H^* \cap J'| = \ell|]$$
$$\leqslant \sum_{\ell \leqslant (q-\delta)k} \mathbb{P}_{J',H' \sim K_n} \left[|J' \cap H'| = \ell \right] \cdot M_J p^{L-\ell},$$

where M_J is the number of copies of J in K_n, and $L = |J| \geqslant qk$ (by Theorem 3.6(b)) and by $\mathbb{P}_{J',H' \sim K_n}$ we refer to the joint measure of J', and copy of J^* in K_n and H' an independent copy of H in K_n.

By employing [MNS$^+$23, Lemma 4.4],

$$\mathbb{P}_{J',H' \sim K_n} \left[|J' \cap H'| = \ell \right] \leqslant v(H)^{O(v(H))} \cdot n^{-v_\ell(J)},$$

where $v_\ell(J)$ is the minimal number of vertices in a subgraph of J with ℓ edges. Therefore,

$$\mathrm{MMSE}_n(p) - (1 - q + \delta) \leqslant v(H)^{O(v(H))} \sum_{\ell \leqslant (q-\delta)k} \cdot n^{-v_\ell(J)} \cdot M_J p^{L-\ell}$$

Now observing that $M_J = (n)_{v(J)}/|\mathrm{Aut}(H)| \leqslant n^{v(J)}$, we get that

$$\mathrm{MMSE}_n(p) - (1 - q + \delta) \leqslant \sum_{\ell \leqslant (q-\delta)k} v(H)^{O(v(H))} \cdot n^{v(J)-v_\ell(J)} \cdot p^{L-\ell}.$$

By the definitions of $\widetilde{\varphi}_q$ and J^*, we have that for all $\ell \leqslant (q-\delta)k$,

$$\widetilde{\varphi}_q \leqslant n^{-\frac{v(J)-v_\ell(J)}{L-\ell}}.$$

Thus, for large enough n, since $p \leqslant (1 - \varepsilon)\widetilde{\varphi}_q$, we also have

$$\mathrm{MMSE}_n(p) - (1 - q + \delta) \leqslant v(H)^{O(v(H))} \sum_{\ell \leqslant (q-\delta)k} (1 - \varepsilon)^{L-\ell} \leqslant v(H)^{O(v(H))} k(1 - \varepsilon)^{\delta k} = o(1),$$

as desired.

$\square$

6.5. **Proof of Lemma 6.2: Impossibility of recovery for $p \geqslant (1 + \varepsilon)\varphi_q$.** In this section, we prove Lemma 6.2. The proof of this lemma will essentially be a formalization of the intuition mentioned above.

By definition of φ_q, there exists some $S \subseteq H$ with $|S| \leqslant qk$, for which we have

$$p \geqslant (1 + \varepsilon) \cdot \max \left\{ n^{-\frac{|V(T)\setminus V(S)|}{|T\setminus S|}} : T \subseteq H \right\}.$$

Let $\delta > 0$ be a fixed but arbitrary small constant, which we will eventually shrink it to 0. In particular, it must hold

$$p \geqslant (1 + \varepsilon) \cdot \max \left\{ n^{-\frac{|V(T)\setminus V(S)|}{|T\setminus S|}} : T \subseteq H, |T\setminus S| \geqslant \delta k. \right\}.$$

Our goal will be to show that, for all p satisfying the inequality above, it will be impossible to recover better than a δ-fraction of the set of edges S^c. This will imply by standard Bayesian tools that

$$\mathrm{MMSE}_n(p) \geqslant 1 - q - O(\delta) - o(1),$$

and thus Lemma 6.2, since δ can be taken arbitrarily small.

More concretely, our technique will be to show that even if the statistician is given the S^*-part of the planted graph H^* as side information, it will be impossible for the statistician to recover an $\Omega(1)$ fraction of the rest of $H^*\setminus S^*$; in particular, with side information the statistician must miss at least a $1 - q - o(1)$ fraction of H_*, and then via a data processing argument, we can conclude the same for the original setting without side information.

To define this alternate problem, consider the following inference problem defined on slightly modified graphs where the edges of a particular subgraph of H is removed from the observations.

Definition 6.4 (Planted S^c-graph). Let $H = H_n$ be a graph and H^* be a uniformly random draw from the isomorphism class of H in K_n. Let also S be a subgraph of H and S^* an arbitrary copy of S in H^*. The *planted S^c-graph* problem is the task to infer $H^* \backslash S^*$ given a draw from $G = (H^* \backslash S^*) \cup G_0$, where G_0 is the product Bernoulli measure with parameter p only on the edges of $K_n \backslash S^*$. Moreover, we define the MMSE of the problem,

$$\mathrm{MMSE}_{n,S}(p) = \frac{1}{|H| - |S|} \mathbb{E}[\|1_{H^* \backslash S^*} - \mathbb{E}[1_{H^* \backslash S^*}\|_2^2 | G].$$

Notice that in the planted S^c graph problem, the observed G is only supported on the edges of $K_n \backslash S^*$.

We start with the following reduction from the original problem to the planted graph-cut problem for S, at an MMSE loss of a multiplicative factor $1 - q$.

Lemma 6.5 (Reduction). *Let $S \subseteq H$ with $|S| \leqslant q|H|$. Recall $\mathrm{MMSE}_n(p)$ is the MMSE of the original planted problem and $\mathrm{MMSE}_{n,S}(p)$ is the MMSE in the planted S^c-graph-cut problem. Then it holds*

$$\mathrm{MMSE}_{n,S}(p)(1 - q) \leqslant \mathrm{MMSE}_n(p).$$

Proof. Given a copy $\tilde{H}$ of H, let us denote by $S(\tilde{H})$ the part of $\tilde{H}$ corresponding to a copy of S (resolving ties arbitrarily). Notice that for any fixed copy S_0 of S in K_n, by the law of total variance we have

$$\mathrm{MMSE}_n(p) = \frac{1}{|H|} \mathbb{E}\mathrm{Var}(H^*|G) \geqslant \frac{1}{|H|} \mathbb{E}\mathrm{Var}(H^*|G, \widetilde{S(H)} = S_0)$$

But now consider the statistical model where the prior is uniform over all copies H' of H so that $S(H') = S_0$, and the noise is as usual an independent instance of $G(n, p)$. The posterior $\mathbb{P}'_{S_0}$ in this model is uniform among all copies H' of H in the observed graph with $S(H') = S_0$, and therefore the *unnormalized* MMSE in this model satisfies

$$\mathrm{MMSE}'_{S_0,n}(p) = \mathbb{E}\mathrm{Var}(H^*|G, \widetilde{S(H)} = S_0).$$

In particular,

$$\frac{1}{|H|} \mathrm{MMSE}'_{S_0,n}(p) \leqslant \mathrm{MMSE}_n(p).$$

Moreover clearly the posterior in the planted S_0^c-graph model is the marginal of $\mathbb{P}'_{S_0}$ on $K_n \backslash S_0$. Therefore by two applications of the Nishimori identity [Lemma A.1](identifying graphs on n vertices with vectors in $\{0, 1\}^{\binom{n}{2}}$),

$$\mathrm{MMSE}'_{S_0,n}(p) = |H| - \mathbb{E}_{v,v' \sim \mathrm{i.i.d.} \mathbb{P}_{S_0}}\langle v, v'\rangle = |H| - (|S| + \mathbb{E}_{v,v' \sim \mathrm{i.i.d.} \mathbb{P}_S}\langle v, v'\rangle) = (|H| - |S|) \mathrm{MMSE}_{S_0,n}(p)$$

The second equality follows because the bits about the edges of S_0 are always equal to one under $\mathbb{P}'_{S_0}$. The result follows since for any copy S_0 of S, clearly $\mathrm{MMSE}_{S_0,n}(p) = \mathrm{MMSE}_{S,n}(p)$, and of course $|H| - |S| \geqslant (1 - q)|H|$. $\qquad \square$

Given [Lemma 6.5](), to complete the proof of impossibility of recovery, we require the following impossibility result for the planted graph-cut problem.

Lemma 6.6. *Let $q \in [0, 1)$, $H = H_n$ a weakly dense graph and $S = S_n \subseteq H$, $|S| \leqslant q|H|$. Consider the planted S^c-graph problem for the planted graph H. Then, assume that for some constants $\varepsilon, \delta > 0$, we have for large enough n,*

$$p \geqslant (1 + \varepsilon) \max \left\{ n^{-\frac{|V(T) \backslash V(S)|}{|T \backslash S|}} : T \subseteq H, |T \backslash S| \geqslant \delta|H|. \right\}.$$

Then

$$\mathrm{MMSE}_{S,n}(p) \geqslant 1 - \delta/(1 - q) - o(1).$$

Proof. Let $k := |H| - |S|$. Let also $\widetilde{G} = (H^* \backslash S) \cup G_0$ be a sample from the planted S^c-graph distribution as defined in Definition 6.4. In particular, let H^* the planted copy of H and S^* the (arbitrarily chosen) copy of S in H^*.

First notice that it suffices to show that

$$\mathbb{P}\left[\frac{|(H' \backslash S^*) \cap (H^* \backslash S^*)|}{k} \geqslant \delta/(1 - q) \right] = o_N(1), \tag{6.1}$$

where H' is a uniformly random isomorphic copy of H containing S^* that is present in the graph $\widetilde{G} \cup S^*$. Indeed, a simple application of Baye's rule implies that the posterior of the planted S^c-graph problem for H is uniform on all $H' \backslash S^*$ where H' is a copy of H in $\widetilde{G} \cup S^*$ containing S^*. Given that (6.1) and Nishimori's identity (Lemma A.1) imply the result.

In the calculation below, we always condition on S^*. For convenience, we introduce some new notation. We denote by $(S^*)^c$-graph, any graph on edges only in $K_n \backslash S^*$. We define G' to be an independent sample from the product Bernoulli(p) measure on $K_n \backslash S^*$ (that is G' has no planted copy $H^* \backslash S^*$) and $\delta' := \delta/(1 - q)$. We also denote for an $(S^*)^c$-graph G,

(a) by $Z(G)$ the number of isomorphic copies H' of H containing S^* in $G \cup S^*$,
(b) by $Z_\ell(H^*, G)$ the number of copies H' of H containing S^* in $G \cup S^*$ with

$$|(H' \backslash S^*) \cap (H^* \backslash S^*)| = \ell,$$

and

(c) by $Z^2(\ell, G)$ is the number of pairs of copies H', H'' of H containing S^* in $G \cup S^*$ with

$$|(H' \backslash S^*) \cap (H'' \backslash S^*)| = \ell.$$

For any $\varepsilon > 0$ We have that by direct calculations (similar to the proof of [MNS$^+$23, Lemma 3.11] where the universe are all edges minus the edges in S)

$$\mathbb{P}\left[\frac{|(H'\backslash S^*) \cap (H^*\backslash S^*)|}{k} \geqslant \delta'\right] = \mathbb{E}\left[\frac{1}{Z(\widetilde{G})}\sum_{\ell \geqslant k\delta'} Z_\ell(H', \widetilde{G})\right]$$

$$\leqslant \mathbb{P}[Z(\widetilde{G}) \leqslant \varepsilon\mathbb{E}Z(G')] + \frac{1}{\varepsilon\mathbb{E}Z(G')}\mathbb{E}\sum_{\ell \geqslant k\delta'} Z_\ell(H^*, \widetilde{G})$$

$$\leqslant \varepsilon + \frac{1}{\varepsilon\mathbb{E}Z(G')}\mathbb{E}\sum_{\ell \geqslant k\delta'} Z_\ell(H^*, \widetilde{G})$$

$$\leqslant \varepsilon + \frac{1}{\varepsilon(\mathbb{E}Z(G'))^2}\mathbb{E}\sum_{\ell \geqslant k\delta'} Z^2(\ell, G')$$

$$= \varepsilon + \frac{1}{\varepsilon}\sum_{\ell \geqslant k\delta'} \frac{\mathrm{P}_{H'}(|(H'\backslash S^*) \cap (H^*\backslash S^*)| = \ell)}{p^\ell}$$

where in the two last inequalities we used a Bayesian inference idea called the "planting trick" [ACO08], as executed in [MNS$^+$23, Lemmas 3.9, 3.10], which allows us to pass from the planted observation graph $\widetilde{G}$ to the "null" graph G'. Moreover, in the last equality under $\mathrm{P}_{H'}$, H' is a uniform copy random isomorphic copy of H in K_n containing S^* and H^* is without loss of generality any fixed arbitrary copy of H in K_n containing S_*. Hence it suffices to show that

$$\sum_{\ell \geqslant k\delta'} \frac{\mathrm{P}_{H'}(|(H'\backslash S^*) \cap (H^*\backslash S^*)| = \ell)}{p^\ell} = o_N(1).$$

Now we employ an idea [MNS$^+$23, Lemma 4.4] to show this, by adjusting it to our S^c-graph setting. Fix some $\ell \geqslant \delta'k$. Notice that for a copy H' to satisfy $|(H'\backslash S^*) \cap (H^*\backslash S^*)| = \ell$ it must hold $|V(H' \cap H^*)\backslash V(S^*)| \geqslant v_{\ell,S*}(H)$ where $v_{\ell,S*}(H)$ is the minimum value of $|V(T)\backslash V(S^*)|$ among all subsets T of $(H^*\backslash S^*)$ with $|T| = \ell$ edges. But to choose the random copy H' of H containing S^*, $|V(H' \cap H^*)\backslash V(S^*)|$ is simply an Hypergeometric distribution with parameters $(v(H) - v(S^*), v(H) - v(S^*), n - v(S^*))$. Hence,

$$\mathrm{P}_{H'}(|(H'\backslash S^*) \cap (H^*\backslash S^*)|) \leqslant \binom{n - v(S^*)}{v(H) - v(S^*)}^{-1}\sum_{w=v_{\ell,S*}(H)}^{v(H)-v(S^*)}\binom{v(H) - v(S^*)}{w}\binom{n - v(H)}{v(H) - v(S^*) - w}$$

$$\leqslant v(H)^{O(v(H))}\sum_{w=v_{\ell,S*}(H)}^{v(H)-v(S^*)}\frac{(n - v(H))_{v(H)-v(S^*)-w}}{(n - v(S^*))_{v(H)-v(S^*)}}$$

$$\leqslant v(H)^{O(v(H))}\frac{(n - v(H))_{v(H)-v(S^*)-v_{\ell,S*}(H)}}{(n - v(S^*))_{v(H)-v(S^*)}}$$

$$\leqslant v(H)^{O(v(H))}(n - v(H))^{-v_{\ell,S*}(H)}$$

$$\leqslant v(H)^{O(v(H))}n^{-v_{\ell,S*}(H)}.$$

Notice let $T_1 \subseteq H^*\backslash S^*$ be such that $|T_1| = \ell$ and $|V(T)\backslash V(S^*)| = v_{\ell,S*}(H)$. Hence for $T = T_1 \cup S^*$ it must hold $|T\backslash S^*| = \ell$ and $|V(T)\backslash V(S^*)| = v_{\ell,S*}(H)$. Now by assumption on p since

$\ell \geqslant \delta' k = \delta(|H| - |S|)/(1 - q) \geqslant \delta|H|$ it must hold that $p^{|T \setminus S^*|} n^{|V(T) \setminus V(S^*)|} \geqslant (1 + \varepsilon)^{|T \setminus S^*|}$ and therefore

$$p^\ell n^{v_{\ell, S^*}(H)} \geqslant (1 + \varepsilon)^\ell.$$

Combining the last two displayed equations we have for all $\ell \in [\delta' k, k]$,

$$\frac{\mathrm{P}_{H'}(|(H' \setminus S^*) \cap (H^* \setminus S^*)|)}{p\ell} \leqslant v(H)^{O(v(H))}(1 + \varepsilon)^{-\ell},$$

hence

$$\sum_{\ell \geqslant k\delta'} \frac{\mathrm{P}_{H'}(|(H' \setminus S^*) \cap (H^* \setminus S^*)|)}{p\ell} \leqslant \sum_{\ell \geqslant k\delta} \frac{v(H)^{O(v(H))}}{(1 + \varepsilon)^\ell}$$

$$= v(H)^{O(v(H))}(1 + \varepsilon)^{-\delta' k} = o(1)$$

where we have used that $\ell \geqslant k\delta' = \delta'(|H| - |S|)/(1 - q) \geqslant \delta|H| = \omega(v(H) \log v(H))$ where the last inequality is by the weakly dense assumption on H. This concludes the proof. $\qquad \square$

6.6. **Proof of Lemma 6.3: A minimax result.** In this section, we proof Lemma 6.3, which is the last missing part of proving Theorem 3.1. To do so we establish the following stronger claim.

Theorem 6.7. *For $q \in (0, 1)$, define*

$$\rho_q = \min\{\max\{\rho(J|S) : J \supseteq S\} : S \subseteq E(H), |S| \leqslant qk\}$$

and

$$\widetilde{\rho}_q = \max\{\min\{\rho(J|S) : S \subseteq J, |S| \leqslant qk\} : J \subseteq E(H)\}.$$

Then,

(a) $\rho_q = \widetilde{\rho}_q$. Equivalently, $\varphi_q = \widetilde{\varphi}_q$.
(b) There exists an optimizer (S^*, J^*) in the definition of ρ_q (and $\widetilde{\rho}_q$) such that $|J^*| \geqslant qk$.

Clearly, Theorem 6.7(a) immediately yields Lemma 6.3.

Proof of Theorem 6.7. It is not difficult to see that Theorem 3.6(b), in conjunction with Theorem 6.7(a), immediately yields Theorem 6.7(b). It remains to prove (a).

Let us first show the easier direction, that $\widetilde{\rho}_q \leqslant \rho_q$. Let (S^*, J^*) be an optimizer in the definition of $\widetilde{\rho}_q$, and $S \subseteq H$ be arbitrary. Our goal is to show that there exists $J' \supseteq S$ such that $\rho(J'|S) \geqslant \rho(J^*|S^*)$. Set $J' = J^* \cup S$ and $T = J^* \cap S$. Then, by Lemma 5.1,

$$\rho(J'|S) \geqslant \rho(J^*|T) \geqslant \rho(J^*|S^*),$$

where the second inequality is because $\rho(J^*|S^*) = \min_{S \subseteq J} \rho(J^*|S)$, so $\widetilde{\rho}_q \leqslant \rho_q$.

Next, suppose that $(J^{(t)}, J^{(t+1)})$ is an optimizer in the definition of ρ_q, using Theorem 3.6(b). We shall show that

$$\widetilde{\rho}_q \geqslant \min_{S \subseteq J^{(t+1)}} \rho(J^{(t+1)}|S) \geqslant \rho(J^{(t+1)}|J^{(t)}) = \rho_q.$$

The first inequality follows from the definition of $\widetilde{\rho}_q$ and that $|J^{(t+1)}| > q|H|$, so it remains to show the middle inequality. Let $S \subseteq J^{(t+1)}$ be arbitrary, and for each $1 \leqslant i \leqslant t + 1$, set $S_i =$

$S \cap J^{(i)}$. First off, note that $\left(\rho(J^{(r+1)}|J^{(r)})\right)_{r \geqslant 0}$ is a non-increasing sequence in r by Lemma 5.4. By Lemma 5.2, for all $0 \leqslant i \leqslant t$, $\rho(J^{(i+1)}|S_{i+1} \cup J^{(i)}) \geqslant \rho(J^{(i+1)}|J^{(i)})$. As a result, for all $i \leqslant t$,

$$\rho(J^{(i+1)}|S_{i+1} \cup J^{(i)}) \geqslant \rho(J^{(t+1)}|J^{(t)}) = \rho_q.$$

But by leveraging the increasing structure of the $J^{(i)}$, $0 \leqslant i \leqslant t+1$, it holds

$$
\begin{aligned}
\rho(J^{(t+1)}|S) &= \frac{|J^{(t+1)} \backslash S|}{|V(J^{(t+1)})| - |V(S)|} \\
&= \frac{\sum_{i=0}^{t} |(J^{(i+1)} \backslash J^{(i)}) \cap (J^{(t+1)} \backslash S)|}{\sum_{i=0}^{t} |V(J^{(i+1)})| - |V(J^{(i)} \cup S_{i+1})|} \\
&= \frac{\sum_{i=0}^{t} |(J^{(i+1)} \backslash J^{(i)}) \cap (J^{(i+1)} \backslash S)|}{\sum_{i=0}^{t} |V(J^{(i+1)})| - |V(J^{(i)} \cup S_{i+1})|} \\
&\geqslant \frac{\sum_{i=0}^{t} |J^{(i+1)}| - |J^{(i)} \cup S_{i+1}|}{\sum_{i=1}^{t} |V(J^{(i+1)})| - |V(J^{(i)} \cup S_{i+1})|} \geqslant \min_{0 \leqslant i \leqslant t} \rho(J^{(i+1)}|S_{i+1} \cup J^{(i)}) \geqslant \rho_q,
\end{aligned}
$$

as desired. □

7. The Minimax rate for any monotone property: Proof of Theorem 4.5

We now prove Theorem 4.5. We do this by first establishing the positive result (part (b)), and then the negative result (part (a)).

7.1. The positive result: Proof of Theorem 4.5(b).
The proof follows by an appropriate application of a "fractional" first moment method. While this method has been used in the probabilistic combinatorics literature, we are not aware of any such application in the high dimensional statistics literature.

Proof. We establish the following property. For any $\delta > 0$, we have that with high probability for any $A_i \in \mathcal{A}$ with $A_i \subseteq Y$ it must hold $|A_i \cap A| \geqslant (q - \delta)k$.

Notice that this property is sufficient to conclude the result. Indeed, Baye's rule implies that for an arbitrary prior π on $\mathcal{A}$ the posterior is assign mass to each $A_i \in \mathcal{A}$ given by the formula $\hat{\pi}(A_i) \propto \pi(A_i)\mathbf{1}(A_i \subseteq Y)$. In particular, the suggested property implies that with high probability the posterior outputs with high probability an A_i with $|A_i \cap A| \geqslant (q - \delta)k$. Therefore by Nishimori's identity (Lemma A.1),

$$\mathrm{MMSE}_N(p, \pi) \leqslant 1 - (q - \delta) + o(1).$$

Since $\delta > 0$ is arbitrary we conclude the desired result.

Now to prove the desired property, first observe that since $Y = A \cup \mathcal{X}_p$,

$$
\begin{aligned}
\mathbb{P}_{A,Y} &\left[\exists\, A_i \in \mathcal{A} \text{ such that } A_i \subseteq Y, |A \cap A_i| \leqslant (q - \delta)k\right] \\
&= \mathbb{E}_A \sum_{S \subseteq A, |S| \leqslant (q-\delta)k} \mathbb{P}_Y\left[\exists\, A_i \in \mathcal{A} \text{ such that } A_i \subseteq Y, A \cap A_i = S\right] \\
&= \mathbb{E}_A \sum_{S \subseteq A, |S| \leqslant (q-\delta)k} \mathbb{P}_{\mathcal{X}_p}\left[\exists\, A_i \in \mathcal{A} \text{ such that } A_i \backslash S \subseteq \mathcal{X}_p, A_i \subseteq S\right]. \qquad (7.1)
\end{aligned}
$$

Now for any S we can choose $w_S \in W_S$ with $\sum_T w_S(T)(\psi_q)^{|T\setminus S|} \leqslant 1/2$. Notice that since by definition of w_S we have $\min_{A_i \in \mathcal{A}, S \subseteq A_i} \sum_{T \subseteq A_i} w_S(T) \geqslant 1$, it also holds

$$\mathbf{1}\left[\exists\, A_i \in \mathcal{A} \text{ such that } A_i \setminus S \subseteq \mathcal{X}_p\right] \leqslant \sum_T w_S(T) \mathbf{1}(T\setminus S \subseteq \mathcal{X}_p)$$

and therefore

$$\mathbb{E}_A \sum_{S \subseteq A, |S| \leqslant (q-\delta)k} \mathbb{P}_{\mathcal{X}_p}\left[\exists\, A_i \in \mathcal{A} \text{ such that } A_i \setminus S \subseteq \mathcal{X}_p\right] \leqslant \mathbb{E}_A \sum_{S \subseteq A, |S| \leqslant (q-\delta)k} \sum_T w_S(T) p^{|T\setminus S|}.$$

Combining with (7.1) gives,

$$\mathbb{P}_{A,Y}\left[\exists\, A_i \in \mathcal{A} \text{ such that } A_i \subseteq Y, |A \cap A_i| \leqslant qk\right] \leqslant \mathbb{E}_A \sum_{S \subseteq A, |S| \leqslant (q-\delta)k} \sum_T w_S(T) p^{|T\setminus S|}.$$

Since w is only supported on T containing S, for all T in the inside summation we have $|T\setminus S| = |T| - |S| \geqslant \delta k$, and therefore using also the $p < \psi_q$ and the definition of w_S,

$$\mathbb{P}_{A,Y}\left[\exists\, A_i \in \mathcal{A} \text{ such that } A_i \subseteq Y, |A \cap A_i| \leqslant qk\right] \leqslant 2^{qk}(p/\psi_q)^{\delta k} \sum_T w_S(T) \psi_q^{|T\setminus S|} \leqslant 2^{qk}(p/\psi_q)^{\delta k}.$$

Since $p = o(\psi_q)$ we conclude the result. $\qquad\square$

7.2. The negative result: Proof of Theorem 4.5(a).
We now prove the negative result of the theorem. The proof follows by an interesting use of a duality trick of Talagrand between fractional covers and spread measures.

Proof. By definition, there exists an $S \subseteq \mathcal{X}, |S| \leqslant qk$ such that

$$\max_{w \in W_S} \sum_{T \subseteq \mathcal{X}} w(T) \psi_q^{|T\setminus S|} \leqslant 1/2.$$

Using the strong duality of linear programs, we conclude that there exists some function π' from the set $\mathcal{A}(S) := \{A_i \in \mathcal{A} : S \subseteq A_i\}$ to $[0, 1]$ such that

(a) $\displaystyle\sum_{A_i \in \mathcal{A}(S)} \pi'(A_i) = 1/2.$

(b) For all T such that $S \subseteq T$ and $|T| \geqslant qk$,

$$\sum_{A_i \in \mathcal{A}(S), T \subseteq A_i} \pi'(A_i) \leqslant (\psi_q)^{|T\setminus S|}.$$

Hence, setting $\pi(A_i) = 2\pi'(A_i)$ for all $A_i \in \mathcal{A}(S)$ we conclude the existence of a measure on $\mathcal{A}(S)$ such that for all T such that $S \subseteq T$ and $|T| \geqslant qk$ it holds

$$\pi_{X \sim \pi}(T \subseteq X) \leqslant 2(\psi_q)^{|T\setminus S|}. \tag{7.2}$$

Remark 7.1. In the language of probabilistic combinatorics, if (7.2) was true for all T, the measure would be called ψ_q-spread. Yet, in our setting we only know that the "spread" condition holds for a subset of T's. Interesting, it is this spread-type of condition that will allow us to use π as a "non-favorable enough" prior to conclude our tight minimax rate lower bound.

We use the "spread" π as our prior and prove that $\liminf_{N\to\infty} \mathrm{MMSE}_N(p, \pi) \geqslant 1 - q$ which concludes the proof. By the Nishimori identity (Lemma A.1), it suffices to prove that for any $1 - q > \delta > 0$ it holds $\mathbb{P}(|A' \cap A| \geqslant (q + \delta)k) = o(1)$ where $A \sim \pi$ and A' is sampled from the posterior of A given $Y = A \cup X_p$.

For any $A \in \mathcal{A}(S), Y \in 2^{\mathcal{X}}$ and $t \in [q + \delta, 1]$, let

$$Z_Y(A, t) := \sum_{A_i \in \mathcal{A}(S)} \pi(A_i)\mathbf{1}\{A_i \subseteq Y, |A \cap A_i| \geqslant tk\}. \tag{7.3}$$

Also denote for simplicity $Z_Y := Z_Y(A, 0)$. A quick application of Baye's rule gives that the posterior distribution is $\mathbb{P}(A|Y) = \pi(A)\mathbf{1}\{A \subseteq Y\}/Z_Y, A \in \mathcal{A}(S)$, and therefore we have the identity

$$\mathbb{P}(|A' \cap A| \geqslant (q + \delta)k) = \mathbb{E}_{A,Y} Z_Y(A, q + \delta)/Z_Y.$$

Now we claim that if we establish $p^{|S|-k}\mathbb{E}_{A,Y} Z_Y(A, q + \delta) = o(1)$, then we can conclude $\mathbb{E}_{A,Y} Z_Y(A, q + \delta)/Z_Y = o(1)$.

This implication follows by a careful "change of measure" argument. To see this denote by $\mathbb{Q}_p$ the product law of $Y = X_p \setminus S$, and by $\mathbb{P}_p$ the "planted" law of $Y = (A \setminus S) \cup (X_p \setminus S)$ for independent $A \sim \pi, X_p \sim \mathbb{Q}_p$. Also recall that π is always supported on sets containing S. Combining these, it holds for all $Y \subseteq \mathcal{X}$ that

$$\mathbb{P}_p(Y \setminus S|A) = p^{|Y|-k}\mathbf{1}(A \subseteq Y) = p^{|S|-k}\mathbf{1}(A \subseteq Y)\mathbb{Q}_p(Y \setminus S),$$

and therefore the likelihood ratio between $\mathbb{P}_p$ and $\mathbb{Q}_p$ take the following form for all $Y \subseteq \mathcal{X}$,

$$\frac{\mathbb{P}_p(Y \setminus S)}{\mathbb{Q}_p(Y \setminus S)} = Z_Y p^{|S|-k}.$$

In particular, for all $\varepsilon > 0$,

$$\mathbb{P}_p(Z_Y \leqslant \varepsilon p^{k-|S|}) = \mathbb{E}_{\mathbb{Q}_p}\left[\frac{\mathbb{P}_p(Y \setminus S)}{\mathbb{Q}_p(Y \setminus S)}\mathbf{1}(Z_Y p^{|S|-k} \leqslant \varepsilon)\right] = \mathbb{E}_{\mathbb{Q}_p}[Z_Y p^{|S|-k}\mathbf{1}(Z_Y p^{|S|-k} \leqslant \varepsilon)] \leqslant \varepsilon.$$

Hence, for any $\varepsilon > 0$, since almost surely $Z_Y(A, q + \delta)/Z_Y \leqslant 1$, we have

$$\mathbb{E}_{A,Y} Z_Y(A, q + \delta)/Z_Y \leqslant \frac{1}{\varepsilon}p^{|S|-k}\mathbb{E}_{A,Y} Z_Y(A, q + \delta) + \varepsilon.$$

From the last inequality by optimizing over $\varepsilon > 0$ we conclude

$$\mathbb{E}_{A,Y} Z_Y(A, q + \delta)/Z_Y \leqslant 2\sqrt{p^{|S|-k}\mathbb{E}_{A,Y} Z_Y(A, q + \delta)},$$

from which the claimed implication follows.

Now by linearity of expectation and (7.3) (for the first equality), (7.2) (for the second inequality) and direct counting arguments,

$$
\begin{aligned}
p^{|S|-k}\mathbb{E}_{A,Y}Z_Y(A,q+\delta) &= \sum_{\ell \geqslant (q+\delta)k} \pi^{\otimes 2}(|A \cap A'| = \ell)p^{|S|-\ell} \\
&\leqslant \sum_{\ell \geqslant (q+\delta)k} \mathbb{E}_A \sum_{T \subseteq A, |T|=\ell} \pi(T \subseteq A')p^{|S|-\ell} \\
&\leqslant 2 \sum_{\ell \geqslant (q+\delta)k} \binom{k-|S|}{\ell-|S|}(\psi_q/p)^{\ell-|S|} \\
&\leqslant 2 \sum_{\ell \geqslant (q+\delta)k} (ke/(\ell-|S|)^{\ell-|S|}(\psi_q/p)^{\ell-|S|} \\
&\leqslant 2 \sum_{\ell \geqslant (q+\delta)k} (e/\delta)^{\ell-|S|}(\psi_q/p)^{\ell-|S|} \\
&= O\left(((\psi_q e)/(\delta p))^{\delta k}\right).
\end{aligned}
$$

Since $p = \omega(\psi_q)$ the last quantity is $o(1)$ and the proof is complete. $\qquad\square$

8. Omitted proofs

8.1. **Proof of Corollary 3.8.**

Proof of Corollary 3.8. Consider a $t = 0, 1, \ldots, M-1$ for which (3.1) holds and any p satisfying (3.2). Now by Theorem 3.6 for any sufficiently small $\varepsilon = \varepsilon_n > 0$ if $q_1 = |J^{(t)}|/|H| + \varepsilon, q_2 = |J^{(t)}|/|H| - \varepsilon$ it holds

$$
\varphi_{q_1} = n^{-\frac{|V(J^{(t+1)})\setminus V(J^{(t)})|}{|J^{(t+1)}\setminus J^{(t)}|}}
$$

and

$$
\varphi_{q_2} = n^{-\frac{|V(J^{(t)})\setminus V(J^{(t-1)})|}{|J^{(t)}\setminus J^{(t-1)}|}}.
$$

Hence p must satisfy

$$
(1+\delta)\varphi_{q_1} \leqslant p \leqslant (1-\delta)\varphi_{q_2}.
$$

By Theorem 3.1 it therefore holds

$$
1 - |J^{(t)}|/|H| - \varepsilon \leqslant \liminf_n \mathrm{MMSE}_n(p) \leqslant \limsup_n \mathrm{MMSE}_n(p) \leqslant 1 - |J^{(t)}|/|H| - \varepsilon.
$$

As $\varepsilon = \varepsilon_n > 0$ can be made arbitrarily small, and in particular $o(1)$, the first part follows.

The second part follows by a similar argument noticing that for any sufficiently small $\varepsilon_n = \varepsilon_n > 0$, it holds

$$
n^{-\frac{|V(J^{(M)})\setminus V(J^{(M-1)})|}{|J^{(M)}\setminus J^{(M-1)}|}} = \varphi_{1-\varepsilon}.
$$

$\qquad\square$

8.2. **Proof of Corollary 3.12.** We begin with a proof of Corollary 3.12, which inverts **??** to calculate the MMSE for a specified value of p.

Corollary 3.12. *Let* $H = H_n$ *be a weakly dense graph and* $p = p_n \in (0, 1)$. *Let the function* $G : (0, 1] \to [0, +\infty)$ *defined by* $G(q) = \varphi_q$. *Consider the sequence* $\alpha_n \in (0, 1)$ *defined by*

$$\alpha_n = \inf G^{-1}([0, p]), n \in \mathbb{N}.$$

Then any limit point $x \in [0, 1]$ *of* $\{\alpha_n : n \in \mathbb{N}\}$, *corresponds in an one-to-one fashion to a limit point of the sequence* $\mathrm{MMSE}_n(p)$. *In particular, if the* $\lim_n \mathrm{MMSE}_n(p)$ *exists, then the* $\lim_n \alpha_n$ *exists and it holds,*

$$\lim_n \mathrm{MMSE}_n(p) = 1 - \lim_n \alpha_n. \tag{3.3}$$

Proof. Consider $y \in [0, 1]$ to be the limit point of the subsequence of $\alpha_{k_n}, n \in \mathbb{N}$. Observe that G is a non-increasing right-continuous function. Hence for any sufficiently small constant $\varepsilon > 0$, for all large n, $G(\alpha_n) \leqslant p \leqslant G((1 - \varepsilon)\alpha_n)$. Now, if $y \notin \{0, 1\}$, then we can conclude by the fact that G is decreasing that for all $\varepsilon > 0$ it holds

$$\limsup_n p_{k_n}/(G(y - \varepsilon))_{k_n} \leqslant 1 \leqslant \liminf_n p_{k_n}/(G(y + \varepsilon))_{k_n}.$$

Using Theorem 3.1 (which trivially generalizes over subsequences) this allows us to conclude

$$1 - y - \varepsilon \leqslant \liminf_n \mathrm{MMSE}_{k_n}(p) \leqslant \limsup_n \mathrm{MMSE}_{k_n}(p) \leqslant 1 - y + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, (3.3) follows for $y \in (0, 1)$. If $y = 0$, then we know for each $\varepsilon > 0$, it must hold $\liminf_n p_{k_n}/(G(\varepsilon))_{k_n} \geqslant 1$. Therefore from Theorem 3.1,

$$1 - \varepsilon \leqslant \liminf_n \mathrm{MMSE}_{k_n}(p).$$

Since $\varepsilon > 0$ is arbitrary, (3.3) follows. The case $y = 1$ follows similarly to the case $y = 0$. $\qquad \square$

8.3. **Proofs of lemmas for the ρ function.** In this section, we prove some properties of the ρ function, restated for convenience.

Lemma 5.1. *Let* A, B *be subgraphs. Then,*

$$\rho(A \cup B | A) \geqslant \rho(B | A \cap B).$$

Proof. We have

$$\rho(A \cup B | A) = \frac{|(A \cup B) \backslash A|}{v(A \cup B) - v(A)}$$

$$= \frac{|B \backslash (A \cap B)|}{v(B) - |V(A) \cap V(B)|},$$

so the lemma is a consequence of the simple inequality

$$|V(A) \cap V(B)| \geqslant |V(A \cap B)| = v(A \cap B). \qquad \square$$

Lemma 5.2. *Let $K|T$ be a graph-cut that is "balanced", in that for all T' such that $T \subseteq T' \subseteq K$, $\rho(T'|T) \leqslant \rho(K|T)$. Then, for all T' such that $T \subseteq T' \subseteq K$ it also holds*

$$\rho(K|T') \geqslant \rho(K|T).$$

Proof. For any T' such that $T \subseteq T' \subseteq K$ we have

$$\rho(K|T') = \frac{|K \backslash T'|}{v(K) - v(T')} = \frac{(|K| - |T|) - (|T'| - |T|)}{(v(K) - V(T)) - (v(T') - v(T))} \geqslant \frac{|K \backslash T|}{v(K) - v(T)} = \rho(K|T),$$

where the final inequality is equivalently with $\rho(T'|T) \leqslant \rho(K|T)$. $\square$

Lemma 5.3. *Let A be a subgraph and consider the map φ_A which maps any subgraph J containing A to $\varphi_A(J) = \rho(J|A)$. If J' and J'' are two subgraphs containing A and at least one of them is balanced, then $\varphi_A(J' \cup J'') \geqslant \min\{\varphi_A(J'), \varphi_A(J''')\}$. In particular, if J', J'' are two maximizers of φ_A, then $J' \cup J''$ is also a maximizer of φ_A, so there is a unique maximal maximizer of φ_A.*

Proof. Let $Q' = J' \backslash A, Q'' = J'' \backslash A$. Then we have

$$\rho(J' \cup J''|A) = \frac{|Q'| + |Q''| - |Q \cap Q''|}{|V(Q')| + |V(Q'')| - |V(Q') \cap V(Q'')|} \tag{8.1}$$

$$\geqslant \frac{|Q'| + |Q''| - |Q \cap Q''|}{v(Q') + v(Q'') - v(Q' \cap Q'')}. \tag{8.2}$$

By the balancedness of one of J', J'',

$$\rho(J' \cap J''|A) = \frac{|Q' \cap Q''|}{v(Q' \cap Q'')} \leqslant \max\left\{\frac{|Q'|}{v(Q')}, \frac{|Q''|}{v(Q'')}\right\}.$$

Combining the two last displayed equations with (8.1) we conclude

$$\rho(J' \cup J''|A) \geqslant \min\left\{\frac{|Q'|}{v(Q')}, \frac{|Q''|}{v(Q'')}\right\} = \min\{\rho(J'|A), \rho(J''|A)\},$$

as desired. $\square$

Lemma 5.4. *Let A be a subgraph and $B = \arg\max_{J \supseteq A} \rho(J|A)$. Then for all $C \supseteq B$,*

$$\rho(C|B) \leqslant \rho(B|A).$$

In particular, for $J^{(t)}, t = 0, 1, \ldots, M$ the onion decomposition of any subgraph H the sequence $\rho(J^{(t+1)}|J^{(t)}), t = 0, 1, \ldots, M - 1$ is non-increasing.

Proof. For all $C \supseteq B$, we have

$$\rho(B|A) = \frac{|B| - |A|}{|V(B)| - |V(A)|} \geqslant \rho(C|A) = \frac{|C| - |A|}{|V(C)| - |V(A)|}$$

and therefore

$$\frac{|B| - |A|}{|V(B)| - |V(A)|} \geqslant \frac{|C| - |B| + (|B| - |A|)}{(|V(C)| - |V(B)|) + (|V(B)| - |V(A)|)},$$

which rearranging gives

$$\rho(C|B) = \frac{|C| - |B|}{|V(C)| - |V(B)|} \leqslant \rho(B|A) = \frac{|B| - |A|}{|V(B)| - |V(A)|}$$

Now, consider $J^{(t)}, t = 0, 1, \ldots, M$ the onion decomposition of a graph H. Then for any $t = 0, 1, \ldots, M-1$, applying the established inequality for $A := J^{(t-1)}$, which gives by definition $B := J^{(t)}$, and $C := J^{(t+1)} \supseteq J^{(t)}$, implies

$$\rho(J^{(t+1)}|J^{(t)}) \leqslant \rho(J^{(t)}|J^{(t-1)}). \qquad \qquad \square$$

REFERENCES

[ABI⁺23]   Damiano Abram, Amos Beimel, Yuval Ishai, Eyal Kushilevitz, and Varun Narayanan. Cryptography from planted graphs: security with logarithmic-size messages. In *Theory of Cryptography Conference*, pages 286–315. Springer, 2023.

[ACO08]   Dimitris Achlioptas and Amin Coja-Oghlan. Algorithmic barriers from phase transitions. In *2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 793–802. IEEE, 2008.

[ALWZ20]   Ryan Alweiss, Shachar Lovett, Kewen Wu, and Jiapeng Zhang. Improved bounds for the sunflower lemma. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 624–630, 2020.

[BBH18]   Matthew Brennan, Guy Bresler, and Wasim Huleihel. Reducibility and computational lower bounds for problems with planted sparse structure. In *Conference On Learning Theory*, pages 48–166. PMLR, 2018.

[Cha00]   Moses Charikar. Greedy approximation algorithms for finding dense components in a graph. In *International workshop on approximation algorithms for combinatorial optimization*, pages 84–95. Springer, 2000.

[COKPZ17]   Amin Coja-Oghlan, Florent Krzakala, Will Perkins, and Lenka Zdeborová. Information-theoretic thresholds from the cavity method. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 146–157, 2017.

[DMK⁺16]   Mohamad Dia, Nicolas Macris, Florent Krzakala, Thibault Lesieur, Lenka Zdeborová, et al. Mutual information for symmetric rank-one matrix estimation: A proof of the replica formula. *Advances in Neural Information Processing Systems*, 29, 2016.

[DMW25]   Abhishek Dhawan, Cheng Mao, and Alexander S Wein. Detection of dense subhypergraphs by low-degree polynomials. *Random Structures & Algorithms*, 66(1):e21279, 2025.

[DWXY23]   Jian Ding, Yihong Wu, Jiaming Xu, and Dana Yang. The planted matching problem: Sharp threshold and infinite-order phase transition. *Probability Theory and Related Fields*, 187(1-2):1–71, 2023.

[ER59]   Paul Erdős and Alfréd Rényi. On random graphs. *Publ. math. debrecen*, 6(290-297):18, 1959.

[ER60]   Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publ. math. inst. hung. acad. sci*, 5(1):17–60, 1960.

[FB⁺99]   Ehud Friedgut, Jean Bourgain, et al. Sharp thresholds of graph properties, and the k-sat problem. *Journal of the American mathematical Society*, 12(4):1017–1054, 1999.

[FKNP21]   Keith Frankston, Jeff Kahn, Bhargav Narayanan, and Jinyoung Park. Thresholds versus fractional expectation-thresholds. *Annals of Mathematics*, 194(2):475–495, 2021.

[GMZ23]   David Gamarnik, Elchanan Mossel, and Ilias Zadik. Sharp thresholds imply circuit lower bounds: from random 2-sat to planted clique. *arXiv preprint arXiv:2311.04204*, 2023.

[GSXY25]   Julia Gaudio, Colin Sandon, Jiaming Xu, and Dana Yang. "All-Something-Nothing" Phase Transitions in Planted k-Factor Recovery, 2025.

[GV05]   Dongning Guo and Sergio Verdú. Randomly spread CDMA: Asymptotics via statistical physics. *IEEE Transactions on Information Theory*, 51(6):1983–2010, 2005.

[GZ22]   David Gamarnik and Ilias Zadik. Sparse high-dimensional linear regression. estimating squared error and a phase transition. *The Annals of Statistics*, 50(2):880–903, 2022.

[Hul22]   Wasim Huleihel. Inferring hidden structures in random graphs. *IEEE Transactions on Signal and Information Processing over Networks*, 8:855–867, 2022.

[Jer92]   Mark Jerrum. Large cliques elude the metropolis process. *Random Structures & Algorithms*, 3(4):347–359, 1992.

[KK07]   Jeff Kahn and Gil Kalai. Thresholds and expectation thresholds. *Combinatorics, Probability and Computing*, 16(3):495–502, 2007.

[Kuč95]   Luděk Kučera. Expected complexity of graph partitioning problems. *Discrete Applied Mathematics*, 57(2-3):193–212, 1995.

[LKZ15]   Thibault Lesieur, Florent Krzakala, and Lenka Zdeborová. Mmse of probabilistic low-rank matrix estimation: Universality with respect to the output channel. In *2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 680–687. IEEE, 2015.

[LM17]   Marc Lelarge and Léo Miolane. Fundamental limits of symmetric low-rank matrix estimation. In *Conference on Learning Theory*, pages 1297–1301. PMLR, 2017.

[MMX21]   Mehrdad Moharrami, Cristopher Moore, and Jiaming Xu. The planted matching problem: Phase transitions and exact results. *The Annals of Applied Probability*, 31(6):2663–2720, 2021.

[MNS+23]   Elchanan Mossel, Jonathan Niles-Weed, Youngtak Sohn, Nike Sun, and Ilias Zadik. Sharp thresholds in inference of planted subgraphs. In *The Thirty Sixth Annual Conference on Learning Theory*, pages 5573–5577. PMLR, 2023.

[MNSZ22a]   Elchanan Mossel, Jonathan Niles-Weed, Nike Sun, and Ilias Zadik. On the second kahn–kalai conjecture. *arXiv preprint arXiv:2209.03326*, 2022.

[MNSZ22b]   Elchanan Mossel, Jonathan Niles-Weed, Nike Sun, and Ilias Zadik. A second moment proof of the spread lemma. *arXiv preprint arXiv:2209.11347*, 2022.

[MT06]   Andrea Montanari and David Tse. Analysis of belief propagation for non-linear problems: The example of CDMA (or: How to prove Tanaka's formula). In *2006 IEEE Information Theory Workshop-ITW'06 Punta del Este*, pages 160–164. IEEE, 2006.

[NZ23]   Jonathan Niles-Weed and Ilias Zadik. It was "all" for "nothing": sharp phase transitions for noiseless discrete channels. *IEEE Transactions on Information Theory*, 69(8):5188–5202, 2023.

[PP24]   Jinyoung Park and Huy Pham. A proof of the kahn–kalai conjecture. *Journal of the American Mathematical Society*, 37(1):235–243, 2024.

[RG13]   Galen Reeves and Michael C Gastpar. Approximate sparsity pattern recovery: Information-theoretic lower bounds. *IEEE Transactions on Information Theory*, 59(6):3451–3465, 2013.

[RP19]   Galen Reeves and Henry D Pfister. The replica-symmetric prediction for random linear estimation with gaussian matrices is exact. *IEEE Transactions on Information Theory*, 65(4):2252–2283, 2019.

[RXZ20]   Galen Reeves, Jiaming Xu, and Ilias Zadik. The all-or-nothing phenomenon in sparse linear regression. *Mathematical Statistics and Learning*, 3(3), 2020.

[Tal10]   Michel Talagrand. Are many small sets explicitly small? In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 13–36, 2010.

[Tan02]   Toshiyuki Tanaka. A statistical-mechanics approach to large-system analysis of CDMA multiuser detectors. *IEEE Transactions on Information theory*, 48(11):2888–2910, 2002.

[YZZ24]   Xifan Yu, Ilias Zadik, and Peiyuan Zhang. Counting stars is constant-degree optimal for detecting any planted subgraph: Extended abstract. In Shipra Agrawal and Aaron Roth, editors, *Proceedings of Thirty Seventh Conference on Learning Theory*, volume 247 of *Proceedings of Machine Learning Research*, pages 5163–5165. PMLR, 30 Jun–03 Jul 2024.

Appendix A. Auxiliary Lemmas

In this section we prove some standard results that are needed for the main body of this work. We start with the Nishimori identity we use often in this work.

Lemma A.1 (Nishimori's identity). *Suppose a Bayesian inference setting where the parameter θ is sampled from some prior π supported on S^{N-1}, i.e., $\theta \sim \pi$, and we have observations $Y \sim P_\theta$. Let also $\pi'(Y)$ be the posterior distribution of θ given Y. Then,*

$$\mathrm{MMSE}(\theta) := \mathbb{E}[\|\theta - \mathbb{E}[\theta|Y]\|_2^2] = 1 - \mathbb{E}_{\theta \sim \pi, Y \sim P_\theta, \theta' \sim \pi'(Y)}[\langle \theta, \theta' \rangle].$$

In particular, for the planted subgraph model for a graph H it holds

$$\mathrm{MMSE}_n(p) = 1 - \frac{1}{|H|}\mathbb{E}[|H^* \cap H'|],$$

where H' is a uniform random copy of H in the observed graph G.

Proof. See [NZ23, Lemma 2] for the proof of the identity. The application for the normalized MMSE of the planted subgraph model follows immediately from the fact that the posterior in the planted subgraph model is the uniform measure on the copies of H in the observed graph G. □

The following is a standard lemma on the MMSE curve.

Lemma A.2. *For any graph $H = H_n$ and for any $n, p = p_n \in (0,1)$, under the planted subgraph model corresponding to H the function $\mathrm{MMSE}_n(p)$ is a non-decreasing polynomial function of p.*

Proof. By the Nishimori identity,

$$\mathrm{MMSE}_n(p) = 1 - \frac{1}{|H|}\mathbb{E}[|H' \cap H^*|],$$

where H^* is the planted copy of H in G and H' is a uniform random copy of H in the observed G. Expanding a direct argument gives,

$$\mathrm{MMSE}_n(p) = 1 - \frac{M_H}{|H|}\sum_\ell \ell \mathbb{P}(|H" \cap H^*| = \ell)p^{|H|-\ell}, \tag{A.1}$$

where now M_H is the total number of copies of H in K_n, and H^*, H'' are now independent samples from the uniform distribution over *all* copies of H in K_n. Notice that clearly (A.1) implies that $\mathrm{MMSE}_n(p)$ is a polynomial function in p.

The fact that $\mathrm{MMSE}_n(p)$ is non-decreasing in p follows by a simple application of the data processing inequality, see e.g. [MNS+23, Lemma 3.2]. □

We now present a simple identity on the first moment threshold of H.

Lemma A.3. *Suppose H is a weakly dense graph. Then it's first moment threshold $p_{1\mathrm{M}}(H)$ satisfies*

$$p_{1\mathrm{M}}(H) = (1 + o(1))n^{-v(H)/|H|}.$$

We refer the reader to [MNS+23, Lemma 4.2] for the proof.

Finally, we show that, in the context of monotone properties given by graph inclusions, the minimax prior is given by the uniform distribution.

Lemma A.4. *Let $\mathcal{A} \subseteq \{0,1\}^{\binom{n}{2}}$ be a subgraph inclusion property, that is, there exists a graph $H \subseteq \binom{[n]}{2}$ such that $G \in \mathcal{A}$ if and only if G contains an isomorphic copy of H.[7] Then, the maximizer π^* of $\pi \mapsto \mathrm{MMSE}_N(p, \pi)$ in (4.1) is given by the uniform distribution over all copies of H in $K_n = \binom{[n]}{2}$.*

Proof. Let $\sigma \in S_n$ be any permutation of the vertices $[n]$ and, for a prior π, define the action of σ on π as follows:
$$\pi_\sigma(H') = \pi(H'_\sigma),$$
where H'_σ is the copy of H obtained by relabeling the vertices of H' according to σ. It's clear by symmetry that, for all π and σ, we have
$$\mathrm{MMSE}_N(p, \pi) = \mathrm{MMSE}_N(p, \pi_\sigma).$$
Now suppose we choose σ according to the uniform distribution in S_n. Since we have a Markov chain
$$\sigma \to \pi_\sigma \to X_p \cup H', \; H' \sim \pi_\sigma,$$
by the data processing inequality, for any π, we have
$$\mathrm{MMSE}_N(p, \pi) = \mathbb{E}_\sigma \mathrm{MMSE}_N(p, \pi_\sigma) \leqslant \mathrm{MMSE}_N(p, \mathbb{E}_\sigma \pi_\sigma).$$
Finally, since S_n acts transitively on all copies of H in K_n, we have that $\mathbb{E}_\sigma \pi_\sigma$ is the uniform distribution for all π, and hence π^* is uniform. □

---

[7]Isomorphic here is in the sense of *graph* isomorphism, i.e., relabeling of the vertices.